#### **ORIGINAL PAPER**



# From form to sound 自形至聲: visual and aural representations of premodern Chinese phonology and phonorhetoric with applications for phonetic scripts

Jeffrey R. Tharsen<sup>1</sup>

Received: 26 July 2021 / Accepted: 15 September 2022 © The Author(s), under exclusive licence to Springer Nature Switzerland AG 2022

#### Abstract

Recent technological advancements have made it possible to employ databases and custom algorithms to identify and visualize phonetic values for units of language in digital texts represented as words, or in the case of Chinese, as characters. These results can then be calculated into statistics, and more importantly, visualized at the level of the sentence or line, the paragraph, or even the entire work, revealing subtle phonetic patterns and correlations that have generally gone unnoticed by readers over the centuries (or millennia) since their composition. Based on successful applications of these methods to logographic scripts (drawn primarily from premodern Chinese), this paper provides example of how algorithms like these can be applied to languages written in phonetic scripts (specifically premodern and modern English prose and poetry) and demonstrates approaches for using computational methods to calculate and visualize the results of phonological and phonorhetorical analyses at any scale. The algorithms can be used in concert with sound output to generate an immersive multimodal environment from any passage or text by combining the source with visualizations of phonological and phonorhetorical analyses and then linking these directly to the audio, providing readers with a rich multimedia experience. The author draws upon his experience designing the Digital Etymological Dictionary of Old Chinese《古漢語詞源字典》and visualization toolkits like his Visual Text Explorer (and similar endeavors like Poemage) to demonstrate how anyone can employ similar approaches to create intuitive and compelling visualizations of phonological and phonorhetorical patterning for any language or script, ancient or modern, revealing and elucidating the sounds of bygone eras and/or obscure dialects to modern readers.

**Keywords** Old Chinese · Middle Chinese · Phonorhetoric · Text-to-speech · Phonological visualization

Jeffrey R. Tharsen tharsen@uchicago.edu

<sup>&</sup>lt;sup>1</sup> Digital Studies of Language, Culture, and History, Division of the Humanities, The University of Chicago, Chicago, IL, USA

What would it mean to be able to hear classic works of poetry and prose from premodern eras recited in an approximation of the dialect of their composers? How can we employ modern technologies and data visualization methods to reveal subtle correspondences (assonance, consonance, alliteration) and phonorhetoric (rhetorical strategies based on sound patterns) in these works? This article presents a series of case studies using a variety of primary sources, from early Chinese works of poetry and prose to premodern English prose to a 20<sup>th</sup>-century American poem, as examples of how one might use algorithm-based visual and auditory methods to elucidate phonological patterns. If one can generate the phonemes, the methods outlined here should be able to provide similar results for any written composition in any language or dialect, at any scale.<sup>1</sup>

Scholars of the Chinese classics have known for centuries that the pronunciations of Chinese characters have changed so radically from their original sounds that much of ancient Chinese rhyming poetry no longer rhymes and many subtle uses of puns and wordplay in these works are no longer evident to modern scholars. During the Qing 清 dynasty (1636–1912) a few select groups of scholars turned their attention to this question and began to devise classifications of characters based on the "ancient rhymes" evidenced in early Chinese poetry. The first and most famous of these figures was Duan Yucai 段玉裁, who in 1775 produced his "Qun jing yun fen shiqi bu biao" 羣經韻分十七部表 in which he proposes a total of seventeen "rhyme groups" 韻部 for premodern Chinese. Shortly thereafter, Wang Niansun 王念孫 produced his list of twenty-one rhyme groups, adding to and clarifying Duan's work. In the following decades, Jiang Yougao 江有誥 would produce two analytical masterworks detailing patterns of rhyme and cross-rhyme in works of premodern Chinese, the Qun jing yun du 群經韻讀 and Xian Qin yun du先秦韻讀; these remain among the very few substantive attempts to document phonological correspondences in premodern Chinese at the scale of the corpus (Duan, 1829; Wang, 1924; Jiang, 1928).<sup>2</sup>

Over the past century, modern historical linguists of Chinese have followed in Duan, Wang and Jiang's footsteps, greatly refining our understanding of the sounds of various versions of Chinese during different historical eras and in different regions.<sup>3</sup> The life's work of these great scholars, combined with database technologies and parsing algorithms, has allowed us for the first time to begin to devise toolkits that automate much of the arduous lexical spadework and in mere seconds generate output from premodern and modern secondary sources to permit the close examination of the phonological representation of each graph within any premodern Chinese text. The *Digital Etymological Dictionary of Old Chinese (Digital EDOC)* 古漢語詞源字 典 (edoc.uchicago.edu) is the primary toolkit that we will employ throughout much

<sup>&</sup>lt;sup>1</sup> In this study, approximations of the phonology of these texts will be represented by systems generally based on the phonemic symbols used in the International Phonetic Alphabet. For Old and Middle Chinese, these are phonological reconstructions based in historical linguistics and as such do not necessarily represent the actual phonetics that would have been spoken by the authors or transmitters of these works; how well these reconstructions represent human speech acts remains an open question that we hope further in-depth study and examination will help to address. See Baxter and Sargart, *Old Chinese*, p.1–8.

<sup>&</sup>lt;sup>2</sup> See Duan, Liu shu yin yun biao 六書音韻表; Wang, Mao shi qun jing Chu ci gu yun pu毛詩羣經楚辭 古韻譜; Jiang, Qun jing yun du 群經韻讀 and Xian Qin yun du 先秦韻讀.

<sup>&</sup>lt;sup>3</sup> For example, the works of Bernhard Karlgren, Pan Wuyun 潘悟云, William Baxter, Laurent Sagart, Li Fang-kuei 李方桂, Axel Schuessler and Johan-Mattis List.

of this article, but it bears mentioning that there are other renderings of premodern Chinese phonology that could be equally productive and a similar toolkit could provide similar but potentially different answers to the questions posed above.

### 1 Visualizing the phonology of an ancient Chinese poem

The focus of the initial stages when developing the *Digital EDOC* was to examine the phonological structures of some of the most famous poems from the ancient Chinese canon, as these works have served as a focal point for generations of writers, scholars, and poets. The general expectation was that the overwhelming amount of scholarship over the millennia meant that every aspect of these works had already been examined and a phonological analysis would be relatively straightforward, revealing no aspects of these works that previous scholars had not already covered in great detail. But as the following figures attest, it seems that a number of subtle nuances within these works may remain to be uncovered.

The Chinese poem "Guan ju" 開雎 (translated by Arthur Waley as "The Osprey's Cry") is the poem presented first in the anthology *The Classic of Poetry (Shi jing* 詩經), widely considered one of if not *the* most important collection of Chinese poems from antiquity. "Guan ju" has thus received an enormous amount of attention over the centuries, with much ink spilled in debates about whether it should be read as a depiction of a courtship ritual, as a political allegory, as a poem praising an ancient queen, as an expression of sexual desire, or following any of the other interpretations that can be found in the numerous premodern exegeses of the poem.

In terms of phonological structure, the rhyme words in the poem (the graphs with phonetics in boxes in Fig. 1) have been well documented for centuries (including the formulation of a rhyme word with a following particle  $zh\bar{z}$ , a construction found in twenty-seven of the 305 poems in the corpus), and the matching rhyme schemes in stanzas one and three and in stanzas two, four and five are obvious to even the most casual reader. However, a detailed analysis of the phonological reconstructions (primarily using Schuessler's OCM with Baxter's OCB also provided when it diverged significantly from Schuessler) for all the graphs in the poem led to a few small discoveries, none of which seem to have been documented by previous scholars (Baxter, 1992; Schuessler, 2007).<sup>4</sup> When one color-codes the poem as in Fig. 1, the stark divergence between the phonetics in stanzas one and three versus stanzas two, four and five becomes readily apparent, except for the repeated third line that falls in stanza one but not in stanza three (in the third stanza this line is replaced by the famous emotional outburst "悠哉悠哉," "Long thoughts, oh, long unhappy thoughts" in Waley's translation; Waley, 1996).<sup>5</sup> Even more interesting are the double rhymes that become apparent in the first and third stanzas. The repeated finals \*-a and \*-u at the end of the second line of the first stanza of the poem then crop up again in the fourth line, and most strikingly, in the third stanza the ends of the first and second lines form a double rhyme which only would have been apparent in the pre-Qin (pre-221 B.C.E.) period: "

<sup>&</sup>lt;sup>4</sup> See Schuessler, *Etymological Dictionary of Old Chinese* and Baxter, *Handbook of Old Chinese Pho*nology.

<sup>&</sup>lt;sup>5</sup> See Waley, *Book of Songs*, p.81.



Fig. 1 "Guan Ju" phonology in reconstructed old Chinese (Schuessler OCM and Baxter OCB)

不得" (\*pə \*t<sup>\$</sup>ək) would have rhymed perfectly with "思服" (\*sə \*bək) in Old Chinese, while in modern Mandarin *pinyin* these are now pronounced *bù dé* and *sī fú*. There are a couple of other subtle phonological characteristics that color-coding the poem in this manner reveals, but generally speaking, it seems very possible that "Guan ju" may well have been chosen for the preeminent position in the anthology not simply because of its important subject matter and finely wrought poetic constructions, but also due to the significant number and variety of intricate phonological patterns that it employs.

# 2 Visualizing rhyme schemes and phonorhetoric in ancient Chinese prose

One of the dominant features of much premodern Chinese poetry and prose is the use of end-rhyme (either perfect rhyming or cross-rhyming/consonance) in regular patterns. The forms and uses of these patterns in the corpora of bronze inscriptions dating to the Western Zhou (1045–771 B.C.E.) was first described by Wang Guowei 王國維 and Guo Moruo 郭沫若, and more recently examined in great detail by scholars like Chen Shihui 陳世輝 and Wolfgang Behr (Wang, 1936; Guo, 1931; Chen, 1981; Behr, 1997).<sup>6</sup> However, when one uses

<sup>&</sup>lt;sup>6</sup> See Wang, 兩周金石文韻讀; Guo, 金文韻讀補遺; Chen, 金文韻讀續輯; and Behr, *Reimende bronzeinschriften*.

the *Digital EDOC* to algorithmically color-code the inscriptions based on phonological correlations between the graphs in Old Chinese, a number of additional patterns beyond those adduced by Wang, Guo, Chen and Behr become evident (Figs. 2, 3, 4, 5, and 6).

The Liang Qi *zhong* 梁其鍾 are a set of six bells discovered in 1940 in Fufeng county, Shaanxi province. Roughly half of the 137-graph inscription was cast into the upper central and lower left faces on each of the largest bells, and shorter portions of the inscription were inscribed around the central motif on the front of the smaller bells. There is no year notation recorded on any of the vessels found bearing the name Liang Qi; based on art historical criteria and correlations with other vessels they can be dated to the rein of King Yi 夷王 (r. 865–858 B.C.E.) or King Li 厲  $\pm$  (r. 857–842 B.C.E.).

Color-coding highlights the extraordinary number of rhymed and cross-rhymed lines in the inscription: in section one an initial perfectly-rhymed couplet is followed by several cross-rhyming couplets, section two features a rhymed couplet in \*-ak and three lines cross-rhyming with \*-ŋ finals, then sections three and four are comprised of five lines that all rhyme perfectly (the first couplet in \*-aŋ and the three subsequent lines, including an onomatopoeic couplet describing the sounds of the bells, rhyme in \*-oŋ), and then after a short non-rhyming passage (section five), the final section is comprised of a couplet rhyming perfectly in \*-iŋ, a couplet rhyming perfectly in \*-aŋ, and a final line of four graphs in which the second and final characters



Fig. 2 Phonology and rhymes from the Liang Qi zhong inscription, Section 1

·····································								
y a right) speak with the set of	Key							
Transfer on the second s								
No. 6, A. A	Section 2 : I	Liang Qi Extols hi	s Father, Grar	dfather and th	e King			
	汧(梁)	其 肇(肇	) 倉山	井(型	1)	息	目(祖)	考 /
(18) N. 9 (19) N. 19 (19) St. 19 (10) St. 19 (10) T. 7 / (19) St. (19) (19) (19) (19) (19) (19) (19) (19)	(*ran)	*gə (*drau	12 (S) ) *s-r	ut-s *C.tser	1? (*s <sup>c</sup> en)	*Gwsan	*tsha? ( *ts'a? )	*khu?(S)
Paradi Namagi Nitak John proceed Funds tellant,	Liang Qi thu	is follows the mod	lels of his aug	ust deceased-g	grandfather an	d father,		
R R. P. "Sam "sak sak processing sakja	-		/++					
W (R (P)) 'Yek 'Ye 'we' registration in theorem 's has,	未	明	德	,				
27 (F)	"praŋ? [who] posse	*mraŋ ssed bright virtue	"t'ək					
用 / 名 大 田 / Yana - Tenga Na balak din sara and da balak din	[] posse		,					
M (M // 19 M / 1998) Philippe Pite And Append Philippe And Append	虔	夙	タ	,				
■ ((ii) M / (iii) / "vis (*vis) *µs (*vis) fant ing (*unov) in the interment"	*gran	*suk	*s-N-rak					
(R) R B / X P (Na) 1ga Nado Naga - ya Nado Kang Okan Ingana Han of Rana	pious morni	ng to night,						
W (20) W (20) We (20)	辟	天	子	;				
H Tr(G) (H) H ////	*pek	*l'in	*tsə?					
B((B)) B(B) You's (You's) - Makatob - Mag Avenued qualitative and exact of other tamous and table.	royal officia	ls to Heaven's So	n;					
W ( M ) W ( M ) ( M ) ( M ) 1					ATT COM S			
M M [B3] [B] I Schuk Schuk Schuck Schucht	大	• 月[月	] //	事	初(梁)	其	身	,
	*pm *t	sə? [*ŋ™at]		*m-s-rə?-s	(*raŋ)	*gə	*դո	
(16 bil), all for units a summa for prote bound interview in our any get for bounders. (0, 10, (a), 10, 10, (b), (b), (b), (b), (c), (c), (c), (c), (c), (c), (c), (c	Heaven's So	on appointed to see	rvice Liang Q	i himself,				
(a)         (a) <td>邦</td> <td>/</td> <td>君</td> <td>大</td> <td>正</td> <td></td> <td>;</td> <td></td>	邦	/	君	大	正		;	
다마는 - control of pointering formation 요. (요.) (국) · (주) · (7) ·	*p <sup>s</sup> roŋ		*C.qur	*lsat-s	*teŋ			
Aquet dermol gemeilichte end dermeid fellen. H H ≠ ≠ 17 (6) ≵: - <sup>1</sup> / <sub>2</sub> 5 <sup>2</sup> / <sub>2</sub> (m) → <sup>1</sup> / <sub>2</sub> (a) → <sup>1</sup> / <sub>2</sub> mgke	the lords of	the state and the h	igh officials;					
(m) (m) (m) (m) (m) (m) (m) (m) (m) (m) (m) (m) (m) (m) (m) (m) (m) (m)	Ħ	Ŧ	/	子	朝日		,	
■ ★ 大 ★ ■ C // (%) · Vrone To Toto Toto Toto Toto (Toto/Toto(C))	*m.lon-s	*lsin	· ·	*tsə?	*ron?			
M (A) and (	by Heaven's	Son's favor			· · ·			
(18) ↑ ↑ 8 / ↑ ¶(2) ÷ 50 kt / Toukh (1) / Toug / Touge ( face adge ) when an and toug Ma.	-	STL come			177	and )		
(a) H · H R · · · · · · · · · · · · · · · ·	晟	初(梁)	共	/	) 作 (	暦)	0	
-U	~m·et	(~raŋ) Di suggagad in his s	~gə	"	(*re	к)		
10) B & B - Strag - Syst	uses Llang	21 succeed in his a	iemevements.					

Fig. 3 Phonology and rhymes from the Liang Qi zhong inscription, Section 2



Fig. 4 Phonology and rhymes from the Liang Qi zhong inscription, Sections 3 and 4

Lang Q Along MonOption #2028           100: 日 1: 0: 0: 0: 0: 0: 0: 0: 0: 0: 0: 0: 0: 0:	<b>通</b>
Hold Res (Hold Res (H	Кеу
ingelie sum and defenses. B = B → B = View None View View View Viewenik provide ministrate in the former kings. B = C (B) = V ⊂ B (B) = ministrate in the former kings.	-ŋ 合韻 cross-rhyme
100.00     100.00     100.00       100.	-ŋ 押韻 perfect rhyme
H     H	-n 合観 cross rhyme
B         X: P         1           Yet         Yet         Yet         Yet           Yet         Yet         Yet         Yet           B         Yet         Yet         Yet           Yet         Yet         Yet         Yet           Yet         Yet         Yet         Yet	-n 押韻 perfect rhyme
Televis See approach at a series 1 sans (2). Samuell, M / M R X H 1 Verson - No are - Verson - No -	-k 掃韻 nerfect rhyme
by Barren's Bart Advance ■ (18) ■ (20) →	上 合語 exess thume
(*sa) *sa Yana Yana Yana Yana Yana Turan Lang Oshara sequence til te an of Horace #(6/8) # # # # # # # %sa (fyre) . Synch Age . Ang a thranismity headwordy and this maniferenci	· Kuse
B (16) (B) (B) (B) (B) (B) (B) (B) (B) (B) (B	- j swaa periet rayme
M (M)         M (M)         (M)	- ] Fail cross-rayme Section 5 : Intended Use of the Bells
B         B         B         C <thc< th=""> <thc< th=""> <thc< th=""> <thc< th=""></thc<></thc<></thc<></thc<>	用 卲 各(格) 喜 / 侃 前 文 人 : *m.loŋ-s *dauh (S) *křak (*kŕak) *d <sup>+</sup> b? *khǎn? *d×en *mən *niŋ [The bells] will be used to summon the prior decorated men to come enjoy the harmonies;
$\label{eq:constraints} \begin{array}{c} \mathbf{X} & $	用 旂(祈) 匄 康 絕(娱) 屯(纯) ∕ 右(祐), *m.lop-s *C.cər *kât(S) *k-['aŋ / *k*aŋ (*ŋwa (S)) *d <sup>*</sup> un (*dun) *c <sup>*</sup> ə? (*c*ə?) [the bells] will be used to pray for health, happiness and pure blessings,
Bit 1 - Bit 2 -	繛〔綽〕 绾〔官〕 通   余(徐) 。 (*tʰawk) 〔*kʷʿan〕 *[ˤoŋ   *pə.tʿok (*tʿok) ample, extensive and penetrating fortune.

Fig. 5 Phonology and rhymes from the Liang Qi zhong inscription, Section 5



Fig. 6 Phonology and rhymes from the Liang Qi zhong inscription, Section 6

rhyme perfectly: 壽 \*du?-s/\*N-tu? and 寶 \*p<sup>S</sup>u?. Documentation of the widespread use of rhyme and cross-rhyme in this inscription (and the way that rhymed sections are interspersed with short sections featuring no discernable rhyming) provides direct evidence of the complexity of compositional practices used during the middle period of the Western Zhou dynasty, adding levels of depth and nuance both to our understanding of the inscription itself and to the literary culture and environment in which it was produced.<sup>7</sup>

#### 3 Visual and aural rendering of tang-dynasty Chinese verse

When first building the *Digital EDOC*, we used it to analyze a number of poems from the period often referred to as the "golden age" of Chinese verse: the Tang 唐 dynasty (618–907). In the Tang, exchanges of rhymed poems, especially in social settings, and the use of the poetic form called "regulated verse" (*lü shi*律詩) became widespread cultural phenomena (Owen, 1981).<sup>8</sup> When the full reconstructed Middle Chinese phonology of a Tang-era poem is visualized using a similar technique to that used for "Guan ju" above, the output looks something like this (the IPA in Fig. 7 below comes from Edwin Pulleyblank's IPA for Middle Chinese provided in his *Lexicon of Reconstructed Pronunciation*; Pulleyblank, 1991):

Du Fu 杜甫 (712–770) remains arguably the most famous of all the Tang poets, and "Chun wang" ("Spring Scene") is generally considered his most famous single work; it depicts the aftermath of the war that devastated his home city and much of the country during the early Tang, evoking deep feelings of destruction and loss from the point of view of the now-agèd poet. The correspondences in "Chun wang" are less striking than the ones we found in "Guan ju" but the phonological parallels reveal intricate patterns: the -k finals on the first characters in the first and last lines, for example, and the ranges of final sonorants (-n, -ŋ, and -m) used in parallel positions in the second, third and fourth lines. One potentially interesting result of use a toolkit like the *Digital EDOC* by modern readers is that it permits close examination of the ways that the poem no longer rhymes in modern Chinese (here we'll use Mandarin as the main example, highlighted in light blue); the four lines of the poem each end in a perfect -əm rhyme in Middle Chinese, but the finals in modern Mandarin are -en, -in, -in and -an respectively (Pulleyblank 1991).<sup>9</sup>

Over the past three decades, innovations in algorithmically-derived speech synthesis methods have proved useful to people with various speech- and hearingrelated disabilities and recently as a primary method for device interaction (as in the Amazon Alexa, Google Home Assistant & Apple HomePod). The initial stage of development of computer-based speech synthesis was in the late 1970's and early 1980's, and methods remained relatively consistent until the last five years, as new

<sup>&</sup>lt;sup>7</sup> For a discussion of aural-mnemonic devices in Chinese philosophical texts, see Grebnev (2020). See also Tharsen (2015) and (2019).

<sup>&</sup>lt;sup>8</sup> See Owen, *High T'ang*, 1981.

<sup>&</sup>lt;sup>9</sup> See Pulleyblank, Lexicon, 1991.

	Du Fu 杜甫 : "Spring Scene"											
〈春望〉"Spring	Scene" by Du F	u 杜甫 (712-720)				(2-3)						
 guó kwək The country	 破 p∂ <b>p<sup>a</sup>uâ</b> destroyed,	一 山 shān şam its hills a	一 河 Yâ Nd rivers	 在 zài dzâi remain	, ;	一 城 chéng źjäŋ the city	一 春 tŵjuen in Spring,	 草 Cảo tś <sup>s</sup> âu plants and	 木 mù muk trees	一 深 shēn śjəm grown deep	•	tone 平仄 character 中文 pinyin 拼音 rhymes English translation (col Over 1996)
( ) 蔵 gǎn kâm Movad	ー 時 shí ži	一 花 huā xwa	) 潤 jiàn dzjān	 涙 liii ljwi	•	   恨   Yon   Narmed at	 別 bié pjāt	 鳥 niǎo tieu birds	– ش jīng kjuŋ	一 心 xīn sjəm	•	modern
(一) 烽 fěng	」 火 huǒ	_ 連 lián	= = sān	 月 yuè	, ,	一 家 jiǎ	一 書 shū	 抵 dī	」 英 wàn	一 金 jin		Tang 唐 sonorants
p <sup>h</sup> jwoŋ War's beacon	xuâ fires	ljän have gone on	sâm three	njwet months		ka letters fro	śjwo m home	tiei are worth	mjweŋ (ten) thousands	kjəm in gold		no longer rhyme
(1) 自	_ 剪	_ 搔	। स्र	 短		 渾	一欲	 不	<> 勝	-		
bâi buk White	tóu dəu hair,	são são scratched	gèng kuŋ ever	duăn tuân thinner		hùn Yuən about te	yû jiwok	bù pjəu no longer	shèng śjəŋ hold	zān tsjəm [my] cap-pin		modern euphony

Fig. 7 Phonological features of Du Fu's "Spring Scene" based on Pulleyblank's Middle Chinese IPA

computational approaches based on deep neural networks have begun to show promise in reliably replicating human speech.<sup>10</sup>

If a reliable and sophisticated method for reproducing speech from text can indeed be developed, then the auralization of premodern works could make for a valuable object of study and bring to modern audiences the sounds of works from a variety of cultures and historical eras. It must be acknowledged that the hurdles to such an ambitious undertaking are significant; as an initial attempt at proof of concept, we designed an algorithm to convert Pulleyblank's Middle Chinese IPA to a traditional framework for generating speech from a phonemic script: Apple's built-in customizable speech synthesis system, originally called "MacinTalk", now part of "PlainTalk".<sup>11</sup>

The main benefit of the system created by Apple is its ability to use a custom alphabet to represent phonemes as the input for the voice synthesizer; while the voices are still very rudimentary, it remains an easy-to-use and widely accessible method to represent IPA as algorithmically generated speech (and thus can represent any phonemes in the PlainTalk framework). The phonological representation of the IPA for "Chun wang" in PlainTalk is in Fig. 8 (note that the original rhymes are preserved with the repeated final "-AXm"), and the file can be "spoken" by any macOS operating system using the commands provided.<sup>12</sup> When presented to a modern

<sup>&</sup>lt;sup>10</sup> The MacinTalk framework used in this article was originally launched in 1990 and has remained relatively similar to this day. Amazon Polly, based on a similar codebase to the Alexa engine, was released on Amazon Web Services in 2017. Most modern neural network-based methods (DNNs) are based on Google's WaveNet vocoder or Fastspeech2. For overviews and commentary on these approaches, see Van Den Oord et al., "WaveNet: A generative model for raw audio," 2016; Tobing et al., "An Evaluation of Voice Conversion with Neural Network Spectral Mapping Models and WaveNet Vocoder," 2020; Ren et al., "FastSpeech 2: Fast and High-Quality End-to-End Text to Speech," 2020; and Zhang et al., "Revisiting IPA-based Cross-lingual Text-to-speech," 2021.

<sup>&</sup>lt;sup>11</sup> See the MacinTalk/PlainTalk "Speech Synthesis Programming Guide": https://developer.apple.com/ library/archive/documentation/UserExperience/Conceptual/SpeechSynthesisProgrammingGuide/Phone mes/Phonemes.html

<sup>&</sup>lt;sup>12</sup> We have made the files containing the MacinTalk phonetics presented in this article available on Github: https://github.com/thars3n/IJDH\_phonetics.

1. Modern Chinese : "sayrate=120voice=Ting-Ting -f DuFu_pinyin.txt"					
2. Middle Chinese : "sayrate=180voice=Victoria -f DuFu_PHON.txt"					
	<pre>[[inpt PHON]] kwAXk, phUHAA, sAAm, yAA dzAY. syAYng, SyUWn, tsAW, mUHk, SyAXm. kAAm, zzIH, hwAA, dsyAYn, lYIY. yUHn, pyAEt, tIYOW, kyEHng, SyAXm. phwAWng, hUHAO, lyAYng, sAAm, nywEHt. kAA, sYwAO, tIYAE, mywAAng, kyAXm. bEHk, dOW, sAW, kEHng, tUHAAn. yUHAXn, ywAOk, pyUH, sYAXn, tsyAXm.</pre>				

Fig. 8 Commands for and phonology of Du Fu's "Spring Scene" in MacinTalk

audience the effects of hearing a 9<sup>th</sup>-century Chinese poem "recited" by a machine has tangible effects; one scholar of Chinese who is a native speaker of Korean noted the Tang-era pronunciations of the graphs were closer to her native language than she had realized, and others mentioned that the rhythms and sounds of the poem seemed more accurate and compelling when rendered in even such a rough machine-based approximation of Middle Chinese than in any modern dialect.<sup>13</sup>

## 4 Aural rendering of premodern english prose

Once it was determined that the phonology of a logographic script like Chinese (rendered in reconstructed Old or Middle Chinese) could be algorithmically approximated at scale, the next step was to determine whether one could do the same for a premodern language represented in a phonetic script. While Chinese characters generally have multiple possible pronunciations the use of characters makes it relatively easy for algorithms to choose and deploy a specific sound for each graph; phonetic writing systems often display a higher degree of complexity in the ways that graphemes represent sound, and combinations of sounds in particular.

As a test case for premodern English, the famous opening section of Geoffrey Chaucer's Prologue to the *Canterbury Tales* seemed an obvious choice as it remains today one of the most widely studied and recognizable Middle English texts. The IPA used for the Prologue here is from Albert Baugh and Thomas Cable's 2002 *History of the English Language*, as using the work of expert linguists to determine the phonology of a specific work is preferable over a dictionary or algorithm-based deep learning system (this approach is discussed in the next section).<sup>14</sup>

A similar script to the one used to convert Pulleyblank's IPA to the PlainTalk format and the same "say" command with the specific parameters given in the figure above was used (the "Alex" voice sounded better to our ear than the "Victoria" voice

<sup>&</sup>lt;sup>13</sup> Special thanks go to the organizers of Oxford University's Early Text Cultures project who hosted an early presentation of this article in their 2019–2020 "Writing Orality" series. https://www.earlytextcultur es.org/events/past-events/writing-orality/performances

<sup>&</sup>lt;sup>14</sup> See Cable and Baugh, *Companion*, 2002.

Cable, Thomas and Albert C. Baugh. A Companion to Baugh and Cable's A History of the English Language. Third edition. London: Routledge. 2002. Middle English Version Listen to the following text by clicking Prote Vididic English Version 2 Letter to the follow Whan that Agnit with his shownes sorte The drophu of March hath parced to the rote. And bathed every wyne in swoli licour Of which vertu engendred is the flour: Whan Zephruse ek with his sweete breth Inspired path in every holf and heeth The indrig corpose, and the vonge some Hei and the forwales maken workpoint. That siepen at the vorget work of the (So prixeth hem nature in hir corages).— Thame longen forks to good no prigmages. And pathmess for to selen situange strondes; And pathmess for to selen situange strondes; And pathmess for to selen situange strondes; Co Eristeh hem nature in hir corages).— Tame longen for to selen situange strondes; And pathmess for to selen situange strondes; Co Eristeh hem nature in hir corages).— Tame longen for the selen. The hooly biptic martif for to seles. The hooly biptic martif for to seles. 1. "Modern English" (Scottish) : "say --rate=150 --voice=Fiona -f Prologue\_eng.txt" 2. Middle English : "say --rate=150 --voice=Alex -f Prologue Part1 PHON.txt" That hem ham notiper writer use we were seeke. Biff that in that seson on a day. In Southwest a the Tabut as I lay to Caruther and the tabut as I lay to Carutherbury with ful devoid conge. A rught was come into that hostelye Wel ryne and twenty in a compaigne. Of sondry 10k, by aventure yiller in felaweshipe, and pligitmes were they alle. In ta toward Carutherbury wolden ryde. The chambres and the stables weren wyde. And well we weren esed atte base. [[inpt PHON]] hwdAn TAAt AAprIXl wIXT HIXS SUWrAXs SOWtAX. TAX drUbu UXF mAArc hAAr pEHrsAXd tUW TAX rOWtAX. And bAADAXd EHVrIX VEHIXN IXN swIXC LIXkUHr. UXF hwIXC VEHrIIXU HENJEHndrAXd IXS TAX fUHr. hwdAn 2EHfIXrUHS IYK wIXT hIXS swEYtAX brEHT. IXnspIXrAXd hAAT IXN EHVrIX hUHL AANd hEHT. TAX tEHndrAX krUXpDAXS AANG TAX yUHNQAX sUHnNAX. AANd smAAlAX TUWIAXS mAA kAXN mEYLOWGIHAX. AANd smAAlAX TUWIAXS mAA kAXN mEYLOWGIHAX. SUX pIXkEHT hEHm nAALIXUHr IXN hIXr kUHrAAJAXS. The chambres and the stables weren wyde, And well we were meased atte beste And shortly, whan the sonne was to reste. So hadde I spoken with hem evention That I was of hir felaveshipe anon, And made forward enly for to ryse. To take ours wery ther as I ywo deryse. But nathlenes, whil I have tyme and space, But nathlenes, whil I have tyme and space. Me thrythesh it acordawn to resour To take conditionum Of ech of hem, so as its sened me. And whiche the yweren, and of what degree. SUX prIXKEHT hEHm nAAIIXUHT IXn hIXr kUHrAAJAXs. TAAn lOWNgXn HUXLk UOW gUXn UXN nplgrimAAJAXs. AAnd pAAlmEHrs FUXr tOW SEYKAXN strAWnJAX strUXndAXs. tOW fEHrnAX hAAlwAXs kUWT IXn sUHndrIX lUXndAXs. AAnd spEHsIXAAIIX frOWm EHvrIH SIXrAXs EHndAX. OWF EHNgAXlUXnd tOW kAAWntAXrbrIX TEY WEHndAX. TAX hUXIL bIIXsfUH MAArtIXr HUXr tUX sEYKAX. TAAL HEHm hAAT hUXlpAXn hwAAN TAAT TEHIX wEHr sEYKAX. And whiche they weren, and of what degree, And eek in what array that they were inne; And at a knyght than wol I first bigynne.

Fig. 9 Baugh & cable's phonology of Chaucer's prologue with MacinTalk commands and phonemes

used above, though it's somewhat crude; it reminds one of the voice of the supercomputer in the movie *WarGames*). These allowed an auditory rendering of the text in a rough general approximation of how Chaucer might have pronounced the words in Middle English to be produced (see Fig. 9).<sup>15</sup>

# 5 Visual rendering of the phonology of a modern American poem

To test a modern phonetic script, the well-known "Stopping by Woods on a Snowy Evening" by the American poet Robert Frost seemed an ideal choice given its relative popularity and the intricacy of its phonological structure.<sup>16</sup> Unlike with Chaucer, there was no readily available IPA version of this poem, so the PhoTransEdit website (photransedit.com/Online/Text2Phonetics.aspx) was employed to generate the IPA for a "General American" dialect (Figs. 10 and 11).<sup>17</sup>

<sup>&</sup>lt;sup>15</sup> One could equally use the phonology to produce a color-coded visual depiction of the rhymes in the Prologue, similar to the charts for early Chinese prose above.

<sup>&</sup>lt;sup>16</sup> Frost and Lankes, New Hampshire (1923).

<sup>&</sup>lt;sup>17</sup> There also exist recordings of Frost himself reading the poem that one could use to refine the IPA, if one wished to hear it in its author's own voice. See, for example: youtube.com/watch?v=rebVUgCg-SAU.



Fig. 10 PhoTransEdit website user interface



Fig. 11 Frost's "Stopping By Woods on a Snowy Evening" with "General American" dialect IPA generated by Photransedit

To visualize the phonology at the level of the grapheme, we used the Visual Text Explorer (the VTE was originally developed in 2016 to be able to visually highlight character interactions in novels; in training it for Chinese we found it worked at the single-graph level equally well, and thus could represent phonemes in color gradients) to provide the visual representation of the rhyme scheme in Figs. 12, 13, 14, and 15.



Figs. 12–15 Phonology of the four stanzas of Frost's "Stopping By Woods on a Snowy Evening" as rendered in the Visual Text Explorer

Frost employs a clever rhyme scheme in this poem: in all but the final stanza, lines one, two and four end in perfect rhymes, and the non-rhyming word at the end of each third line becomes the dominant rhyme for the following stanza (in the final stanza then all four lines rhyme perfectly), thus AABA BBCB CCDC DDDD. The Visual Text Explorer highlights this pattern with a colored bar for each phoneme; for example, the yellow-orange pattern for -19 at the end of the third line in the first stanza is then visually repeated in the rhyme words at the ends of lines one, two and four in the next stanza, and so on.<sup>18</sup> Similar patterns can be generated for any phonological rendering of any text.

# 6 Tentative conclusions and future steps

The methods for representing the sound patterns in these eminent literary works presented here are simply initial proof-of-concept test examples for what one hopes will become far more refined strategies in the years to come. Most importantly, the tools and methods are simple enough that they can be readily employed by any scholar or student and readily repurposed for any language or script. Future work creating algorithm-based IPA generators for other languages and dialects, more accurate methods of speech synthesis, and algorithmic production of the phonemes and inflections in premodern languages (not simply Old or Middle Chinese, as in the examples above, but potentially dialects of ancient Egyptian, Akkadian, Athenian Greek or the Latin of the early Roman Empire – essentially any source that one could render in a phone/phoneme-based script like the International Phonetic Alphabet) will all contribute to our ability to more precisely identify and analyze the subtle nuances, phonorhetorical devices and intricate wordplay within these texts, hopefully providing new insights into the methods employed by their authors and the literary and cultural practices and traditions that informed their creation and transmission.<sup>19</sup>

Data availability via Github repository https://github.com/thars3n/IJDH\_phonetics

Code availability Not applicable

#### Declarations

Conflicts of interest/Competing interests Not applicable

# References

Baxter, W. H. (1992). A Handbook of Old Chinese Phonology. Mouton de Gruyter. Baxter, W. H., & Sagart, L. (2014). Old Chinese: A New Reconstruction. Oxford University Press.

 $<sup>^{18}</sup>$  While not necessarily useful for speakers of modern English, one could equally convert the IPA to PlainTalk to "speak" the phonology of the poem, as in the examples from Chaucer and Du Fu杜甫.

<sup>&</sup>lt;sup>19</sup> The now discontinued Poemage project was designed to provide precisely this type of nuanced visual guide to the "sonic topology" of poetic works; see McCurdy et al. (2016).

- Behr, W. (1997). Reimende Bronzeinschriften und die Entstehung der chinesischen Endreimdichtung. Ph.D. diss., J.W. Goethe-Universität, Frankfurt a.M.
- Cable, T., & Baugh, A. C. (2002). A Companion to Baugh and Cable's History of the English Language. (3rd ed). Routledge.
- Chen, S. 陳世輝. (1981). Jinwen yundu xuji 金文韻讀續輯. Gu Wenzi Yanjiu, 5, 169-190.

Duan, Y. 段玉裁. (1829). Liu shu yin yun biao 六書音韻表. Huang Qing jing jie.

- Frost, R., & Lankes, J. J. (1923). New Hampshire: A Poem With Notes and Grace Notes. H. Holt and Co.
- Grebnev, Y. (2020). Aural-mnemonic architectonics of ancient chinese philosophical texts. *Monumenta Serica*, 68(2), 289–314.
- Guo, M. 郭沫若. (1931). Jinwen yundu buyi 金文韻讀補遺. Shinagaku 支那學, 6.1, 1-32 and 6.2, 127-128.
- Jiang, Y. 江有誥. (1928). Yin xue shi shu 音學十書. Zhongguo shudian.
- Karlgren, B. (1957). Grammata Serica Recensa. Museum of Far Eastern Antiquities.
- Li, F.-k.李方桂. (1980). Shanggu yin yanjiu上古音研究. Shangwu yinshuguan.
- List, J.-M., Pathmanathan, J., Hill, N., Bapteste, E., & Lopez, P. (2017). Vowel purity and rhyme evidence in Old Chinese reconstruction. *Lingua Sinica*, 3(1), 1–17.
- McCurdy, N., Lein, J., Coles, K., & Meyer, M. (2016). Poemage: Visualizing the sonic topology of a poem. *IEEE Transactions on Visualization and Computer Graphics*, 22, 439–448.
- Owen, S. (1981). The Great Age of Chinese Poetry: The High T'ang. Yale University Press.
- Pan, W. 潘悟云. (2000). Hanyu lishi yinyunxue 汉语历史音韵学. Shanghai jiaoyu chubanshe.
- Pulleyblank, E. G. (1991). Lexicon of Reconstructed Pronunciation in Early Middle Chinese, Late Middle Chinese, and Early Mandarin. UBC Press.
- Ren, Y., Chenxu, H., Xu, T., Tao, Q., Sheng, Z., Zhao, Z., & Liu, T. Y. (2020). Fastspeech 2: Fast and high-quality end-to-end text to speech. arXiv:2006.04558.
- Schuessler, A. (2007). ABC Etymological Dictionary of Old Chinese. University of Hawai'i Press.
- Tharsen, J. (2015). Chinese Euphonics: Phonetic Patterns, Phonorhetoric and Literary Artistry in Early Chinese Narrative Texts. Ph.D. dissertation, University of Chicago.
- Tharsen, J. (2019). Comparative phonorhetorical analyses of speeches in the *Zuo Commentary* and the *Discourses of the States. Journal of Chinese Literature and Culture*, 6(2), 275–330.
- Tobing, P. L., Wu Y. C., Hayashi T., Kobayashi K., & Toda, T. (2020). An Evaluation of Voice Conversion with Neural Network Spectral Mapping Models and WaveNet Vocoder. APSIPA Transactions on Signal and Information Processing, 9, E26. https://doi.org/10.1017/ATSIP.2020.24
- Van Den Oord, A., Dieleman S., Zen H., Simonyan K., Vinyals O., Graves, A., Kalchbrenner N., Senior A., & Kavukcuoglu K. (2016). WaveNet: A generative model for raw audio. Proc. 9th ISCA Workshop on Speech Synthesis Workshop (SSW 9), 125.
- Waley, A. (1996). The Book of Songs. Grove Press.
- Wang, N. 王念孫. (1924). *Mao shi qun jing Chu ci gu yun pu* 毛詩羣經楚辭古韻譜. Luoshi jiben, dongfang xuehui congkan.
- Wang, G. 王國維. (1936). Liang Zhou jinshiwen yundu 兩周金石文韻讀. Wang Jing'an xiansheng yishu.
- Zhang, H., Zhan H., Zhang, Y., Yu, X., & Lin, Y. (2021). Revisiting IPA-based Cross-lingual Text-tospeech. arXiv:2110.07187.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.