*Cang Jie pian* Rhyming Data\_Documentation

The *Cang Jie pian* 蒼頡篇 was an important scribal treatise (*shishu* 史書) in early imperial China. Through mastery of this text, students were able to attain coveted government positions (Foster 2017). Although the *Cang Jie pian* failed to be transmitted among our received corpus of ancient Chinese texts, manuscript discoveries over the past century shed new light on the nature of its content. Indeed, the *Cang Jie pian* is ubiquitous among recently unearthed caches of Han period manuscripts, appearing now in sixteen different collections. Furthermore, it appears in diverse archaeological contexts, from an aristocratic burial in Anhui, to the military installations in Gansu, and even among the remains of the Jingjue 精絕 kingdom in Niya 尼雅, Xinjiang. The Han dynasty (206 BCE-220 CE) witnessed the construction of China’s first enduring empire. Accompanying this political reform were linguistic changes and shifting manuscript culture norms. The central role played by the *Cang Jie pian* in scribal education, alongside its widespread distribution during the Han dynasty, recommend itas a potentially fruitful subject for research into the standardization of Chinese language and writing at this time.

The accompanying Excel file is an initial organization of the rhyming data in the *Cang Jie pian*. It is offered as a first step towards the systematic analysis of the language of the *Cang Jie pian*, to aid in the reconstruction of Han Chinese and the articulation of linguistic changes that occurred during this period. The data is presented using the annotation framework established by List, Hill, & Foster 2019, in order to facilitate access and allow for analysis with computational tools. The first sheet, “CJP Rhyming Data,” presents a sortable table filled in with the pertinent information from our manuscript sources. Additional explanations for the columns may be found below. The second sheet, “Bibliography,” provides standard references for the primary source publications and the secondary scholarship cited in the previous table.

Fortunately, the *Cang Jie pian* is a tightly structured text and is organized around rhymes. Every line is four characters in length, with a rhyme position falling at the conclusion of every second line (e.g., every eighth character). Each chapter, moreover, participates in a single overarching rhyme scheme. Knowing these rules greatly eases our adjudication of where rhyming positions should occur. “CJP Rhyming Data” only presents these structural rhymes, and the variants found in those positions. It is possible that irregular internal rhyming exists in the *Cang Jie pian* as well. Proper analysis of this will require a listing of all variants in the text (e.g., JYX EPT 50.1 風 -> 諷 in 諷誦), which must await for a future version of the table.

“CJP Rhyming Data” utilizes two manuscripts as it base sources for organizing the data. The first is the Peking University *Cang Jie pian*, as published in *Beijing daxue cang Xi Han zhushu yi* 北京大學藏西漢竹書壹 (Beijing daxue chutu wenxian yanjiusuo 2015; abbreviated PKU). The second is JYX EPT 50.1 from among the “new Juyan strips 居延新簡” case, which writes out a nearly complete version of the *Cang Jie pian*’s so-called “opening chapter” (Zhang 2016). These two manuscripts cover most of the known content of the *Cang Jie pian*, but are not exhaustive. Additional rhymes may be found among the other, more fragmentary, manuscript data. Future updates to this Excel file are anticipated, to incorporate this information.

Neither the PKU nor the JYX EPT 50.1 mss are complete in their content. Only the PKU “Zhuan Xu 顓頊 chapter” (strips #46-52) remains intact. Every other chapter in the PKU and JYX EPT 50.1 mss includes some measure of textual loss. At times, these gaps can be supplemented by other manuscript witnesses. These supplements are made in the “LINE” column, and explained in the “NOTES” column. For example, the final line on PKU 1 reads “□□獲得,” with the first two characters missing. These characters may be supplemented by the Fuyang Shuanggudui *Cang Jie pian* witness, strip C009, which has “益就獲得” (Hu & Han 1983; Zhongguo jiandu jicheng bianji weiyuanhui 2001+; abbreviated FY). In “CJP Rhyming Data,” I give the “LINE\_IN\_SOURCE” as “□□獲得” and the “LINE” as “益就獲得.” Since some of the content for the line is still found on PKU 1, it remains the representative source in “SOURCE.” All of this is documented in “NOTES.”

If the supplemented text extends beyond lines found in the PKU or JYX EPT 50.1 mss, then a representative witness is cited in the “SOURCE” column. The extension of *Cang Jie pian* content can be speculative, particularly when it relies on chaining together a number of short fragments. Caution is advised in such instances. For example, before PKU 1 we may surmise that there was a line written “高囂平夷” now lost on the PKU ms, but found on other manuscripts. Because none of this content is present in the PKU ms, I use FY C046 as a representative base source. FY C046 is then given in the “SOURCE” column.

Only the content for lines of the base text are recorded in “CJP Rhyming Data.” Paratextual features, such as the titles and character counts seen on the PKU ms, have been removed. Similarly, the Shuiquanzi 水泉子 ms adds “extensions of meanings” (a sort of commentary) to the four-character base lines of the *Cang Jie pian* (Zhang 2015; abbreviated SQZ). These are not yet recorded, but will be in a future version. As a primer, the *Cang Jie pian* was studied and often used in practice writing. The PKU ms is a model text, written in a careful and legible hand, and faithfully formatted. This is not always the case on other manuscripts. Numerous strips or strip fragments have been discovered where content from the *Cang Jie pian* is scribbled out in disjointed lines, mixed with other content, or full of repeated characters and mistaken orthography. Although JYX EPT 50.1 is composed in a confident hand, it is unclear if it is a model text, or was the result of practice writing. The content at the very end appears to be unrelated to the *Cang Jie pian*. It has thus been removed as well.

For our purposes here, we may differentiate between two editions of the *Cang Jie pian*: (1) An early version in which chapters vary in length, but contain over 100 characters; (2) a later “Village Teachers” version in which chapters have precisely 60 characters in length. The PKU ms is of the first variety and I suspect that the JYX EPT 50.1 ms is of the second variety. Due to the fragmented state of many of our *Cang Jie pian* manuscripts, it is not always feasible to determine from which edition the content derives. Since edition-level variation in content is possible, a degree of caution is warranted when drawing across different manuscript witnesses to supplement or reconstruct hypothetical base text. (See especially PKU 65 for this issue). Comparing these two editions, however, it appears that the text of the Village Teachers’ edition rather mechanically divides the longer chapters of its predecessor into shorter 60-character long segments, without significant further alteration to the content (Foster forthcoming). For this reason, despite the fragmentary nature of our sources, I feel more confident about proposing supplements across manuscript witnesses. A more extensive discussion of the textual history of the *Cang Jie pian* and the methodologies that have been employed in its reconstruction may be found in Foster 2017.

The transcriptions given in “CJPRhyming Data” follow those given in Beijing daxue chutu wenxian yanjiusuo 2015 and Zhang 2016. Future versions of “CJP Rhyming Data” will take into account *all* proposed transcriptions for each word as debated in the scholarly literature. Be aware that these are strict transcriptions, not interpretative. For example, PKU 1 writes 㝡, but this is best read as 冣 instead. Strict transcriptions do not communicate when scribal errors, alternative forms, or loaning effect our reading. When this information potentially impacts the pronunciation of a word in a rhyming position, it is discussed in the “NOTES” column. For content outside of rhyming positions, no comment is given.

Similarly, there are times when the published transcription for a character is difficult or impossible to type. In these cases, I revert to the more interpretative readings given in the publication’s annotations. For example, the second character on PKU 2 is given a strict transcription of ![A picture containing light, sitting, traffic, city

Description automatically generated]() , which is not yet encoded. The Peking University editors identify this as the word 勦, which is what I then use for “CJP Rhyming Data.” Alternatively, if it is important to retain the spelling of the strict transcription, or when the writing is only partial, I describe the character with ⿰ symbols and regular *kaiti* 楷體 form components, but try to relegate this to the “NOTES” column. These details are discussed only when they pertain to rhyming positions, otherwise no comment is given.

The columns employed in the “CJP Rhyming Data” table follow those proposed in List et al. 2019 (see especially p.31), with the addition of RHYME, RHYME\_WORD, SOURCE, and NOTES. Further specifications pertinent to this particular case study are as follows:

POEM = CJP stands for *Cang Jie pian* 蒼頡篇.

STANZA = Continuous units of text are designated by the range of strips on which they appear in our two base sources, the PKU and JYX EPT 50.1 mss. This is necessary both because explicit titles (e.g., “Zhuan Xu 顓頊” on PKU 46-47) are not available for every chapter of the *Cang Jie pian*, while the very structure of a chapter differs between the two editions of the text (100+ characters vs. 60 characters). The grouping of strips in the PKU ms is based on content, rhyming, textual formatting, and the verso lines. Alternative arrangements to the manuscript are possible, and future discoveries could necessitate a reorganization of the stanza delineations.

LINE\_IN\_SOURCE = This gives only the base text found on the manuscript cited in the corresponding SOURCE column. Supplements to that base text drawn from parallels on other manuscripts are not given.

LINE = This gives the reconstructed text for the line, including any supplements to the base text drawn from parallels on other manuscripts. In cases where there is loss in the SOURCE manuscript, the line is hypothetical.

LINE\_ORDER = As described in List et al. 2019, LINE\_ORDER is “A numerical value that provides the order of the lines of a poem in a given stanza.” We can exploit LINE\_ORDER to document variants found on other manuscripts (List et al. 2019, p.39). This is accomplished by designating the same LINE\_ORDER to multiple entries under the same STANZA. For example, entries with the ID of 27 and 28 both belong to STANZA “PKU1-7,” and have the same LINE\_ORDER of 5. Yet they differ in their LINE content: “寬 + 恵 + 善 + 志” vs. “寬 + 恵 + 善 + 記” respectively. This tells us that 志 and 記 are potentially variants for the same line of the *Cang Jie pian,* as found on different manuscripts. For PKU 65, two alternative line series have been proposed. To reflect this, I give designate the lines as a and b. The material is then repeated in the “CJP Rhyming Data” table.

ALIGNMENT = For lines with a rhyme present, the Chinese character is replaced with a reconstruction of the word’s pronunciation. This reconstruction follows William Baxter and Laurent Sagart’s 2016 Old Chinese (https://ocbaxtersagart.lsait.lsa.umich.edu). When multiple reconstructions are possible, each is recorded in “NOTES.” In such cases, the pronunciation given in “ALIGNMENT” is a preliminary judgment about the word intended in the linguistic context of the line (including the possibility of loaning), but this can be ambiguous and awaits final analysis. It should not be taken as my definitive statement on the text’s meaning. When a reconstruction is not available for the in Baxter and Sagart for the word as written, an alternative reconstruction is found, often for its phonetic component.

RHYME = This column provides a convenient identity for the rhyme clusters of each stanza. In some stanzas, it is unclear where the rhyming position should fall. These are not designated a rhyme identity. For PKU 65, two alternative rhyme schemes have been proposed. To reflect this, I give y1 and y2 as the two sets. Since this necessitates different line breaks, the material is repeated on the “CJP Rhyming Data” table.

RHYME\_WORD = This column isolates the character used to write the rhyme word in the given line, to ease comparisons.

SOURCE = When dealing with multiple manuscript witnesses, it is necessary to reference specific sources for the content. Labels are provided for individual strips bearing representative base text (found in LINE\_IN\_SOURCE). A key for these labels may be found at the top of the “Bibliography” sheet.

Finally, a brief word about sources not used. At the time that “CJP Rhyming Data” was first compiled (10 June 2020), only a brief report was available for the 2018 discovery of *Cang Jie pian* material at the Chengba 城壩 site in Quxian 渠縣, Sichuan (Sichuan sheng wenwu kaogu yanjiuyuan et al. 2019). The first volume of the Xuanquanzhi 懸泉置 finds has also been released (Gansu sheng jiandu bowuguan 2019), but I have not yet been able to search through it for additional *Cang Jie pian* pieces. Once available to me, these data will be incorporated into “CJP Rhyming Data.” In the fall of 2019, it was announced that a private collector owned another *Cang Jie pian* manuscript, written across numbered wooden boards (Liu 2019). If genuine, it is the longest witness of the *Cang Jie pian* currently extant, and it moreover clarifies the chapter divisions for the Village Teachers edition. Until the manuscript is properly authenticated, however, it will not be included in “CJP Rhyming Data.”

References

Baxter, William H., and Laurent Sagart. *Old Chinese: A New Reconstruction*. New York: Oxford University Press, 2016.

Beijing daxue chutu wenxian yanjiusuo 北京大學出土文獻研究所, ed. *Beijing daxue cang Xi Han zhushu* (*yi*) 北京大學藏西漢竹書(壹). Shanghai: Shanghai guji chubanshe, 2015. (PKU).

Foster, Christopher J. "Study of the *Cang Jie pian*: Past and Present." PhD Dissertation, Harvard University, 2017.

Foster, Christopher J. "The Shape of the Text: *Gu* Prisms and Han Primers." In Susan Blader, Constance Cook, and Christopher J. Foster, eds. *Thinking About Early China with Sarah Allan*, forthcoming.

Gansu sheng jiandu bowuguan 甘肅省簡牘博物館, ed. *Xuanquan Han jian* (*yi*) 懸泉漢簡(壹). Shanghai: Zhongxi shuju, 2019.

Hu Pingsheng 胡平生 and Han Ziqiang 韓自強. "*Cang Jie pian* de chubu yanjiu 蒼頡篇的初步研究." *Wenwu* 文物 2 (1983): 35-40. (FY).

List, Johann-Mattis, Nathan. W. Hill, and Christopher J. Foster. “Towards a Standardized Annotation of Rhyme Judgments in Chinese Historical Phonology (and Beyond).” Journal of Language Relationship 17 (2019): 26-43, https://doi.org/10.31826/jlr-2019-171-207.

Liu Huan 劉桓. *Xinjian Han du Cang Jie pian Shi pian jiaoshi* 新簡漢牘蒼頡篇史篇校釋. Beijing: Zhonghua shuju, 2019.

Sichuan sheng wenwu kaogu yanjiuyuan 四川省文物考古研究院 and Quxian lishi bowuguan 渠縣歷史博物館, ed. “Sichuan Quxian Chengba yizhi 四川渠縣城壩遺址.” *Kaogu* 考古 7 (2019): 60-76.

Zhang Cunliang 張存良. "Shuiquanzi Han jian Cang Jie pian zhengli yu yanjiu 水泉子漢簡蒼頡篇整理與研究." PhD Dissertation, Lanzhou University, 2015.

Zhang Defang 張德芳, et al. *Juyan xin jian jishi* 居延新簡集釋. Lanzhou: Gansu wenhua chubanshe, 2016, 7 vols. (JYX).

Zhongguo jiandu jicheng bianji weiyuanhui 中國簡牘集成編輯委員會, ed. *Zhongguo jiandu jicheng* 中國簡牘集成. Lanzhou: Dunhuang wenyi chubanshe, 2001+, vol.14 & 18. (FY).