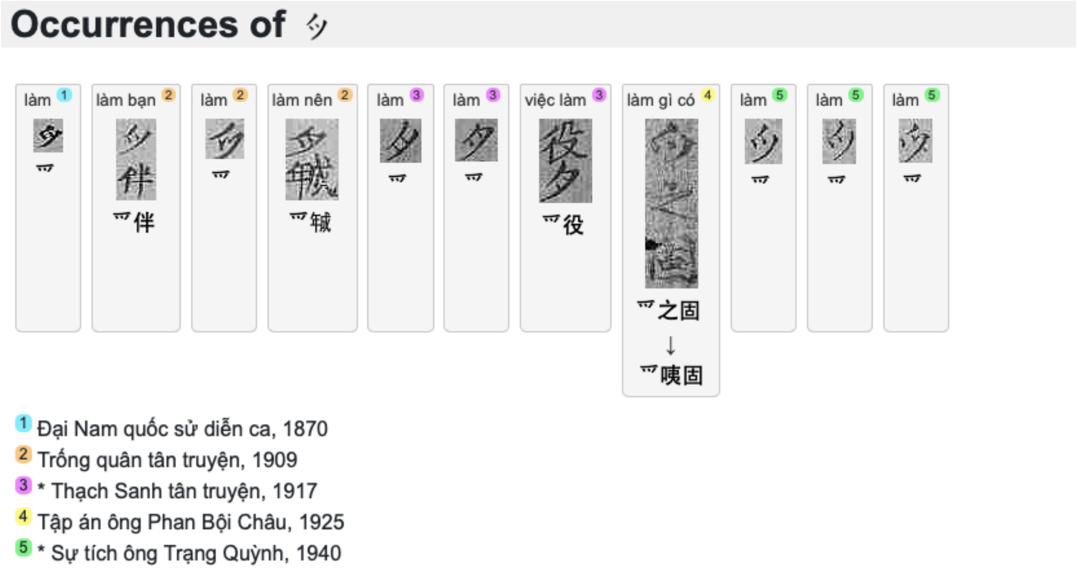
Universal Multiple-Octet Coded Character Set International Organization for Standardization

Doc Type:	ISO/IEC JTC1/SC2/WG2/IRG
Title:	Request to dis-unify U+722B
Source:	Vietnam
Status:	National Body Proposal
Authors:	Eiso Chan, Lee Collins, Ngô Thanh Nhàn
Action:	For consideration by the IRG
Date:	2020-06-10

This paper proposes dis-unification of URO character U+722B 爫 into 2 characters: The original U+722B, which would retain it shape and semantic as Kang Xi radical no. 87, and a new character designed to be centered within the ideographic bounding box, and to occupy the greater part of the bounding box space. The new character will be semantically different from U+722B, representing the Vietnamese verb *làm* and other words written with a full form of U+722B, such as Zhuang *lamz*.

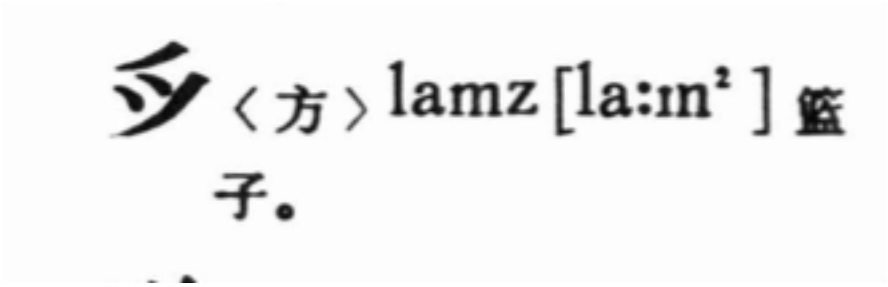
Làm, meaning “make, do, consist of”, etc. is one of the most frequently used Vietnamese verbs. It has many variant forms in chữ Nôm, including: 濫, 𬈋, 滥, and 爫, but the most common forms are those shown in the picture below, with examples from various sources, as documented at https://www.chunom.org/pages/722B/. Detailed images are shown in the Appendix.



Occurrences of 爫

1 Đại Nam quốc sử diễn ca, 1870
2 Trống quân tân truyện, 1909
3 * Thạch Sanh tân truyện, 1917
4 Tập án ông Phan Bội Châu, 1925
5 * Sự tích ông Trạng Quỳnh, 1940

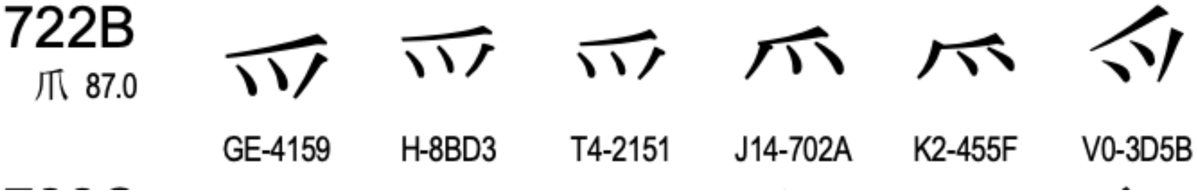
There are two source characters from which this form of *làm* appears to have been derived. One is U+7232 爲. The semantic of *làm* is clearly related to the sense of U+7232, but the shape has been reduced to the top component of U+7232. The resulting simplification facilitates writing of this high frequency character and marks the usage as Vietnamese. The second is the character U+6FEB 濫, which has the Sino-Vietnamese reading *lạm*. This derivation is purely phonetic. Nguyễn Quang Hồng's *Tự Điện Chữ Nôm Dẫn Giải* shows these steps for the alternate derivation of *làm*: 濫 → 濫 → 氵爫 → 爫. The second source of *làm* is interesting because it resembles the derivation of the Zhuang character *lamz*, the sense and form of which appear to be derived from U+7C43, 籃 (basket). Both 濫 and 籃 share the phonetic component 監.



爫〈方〉lamz [laːm²] 篮子。

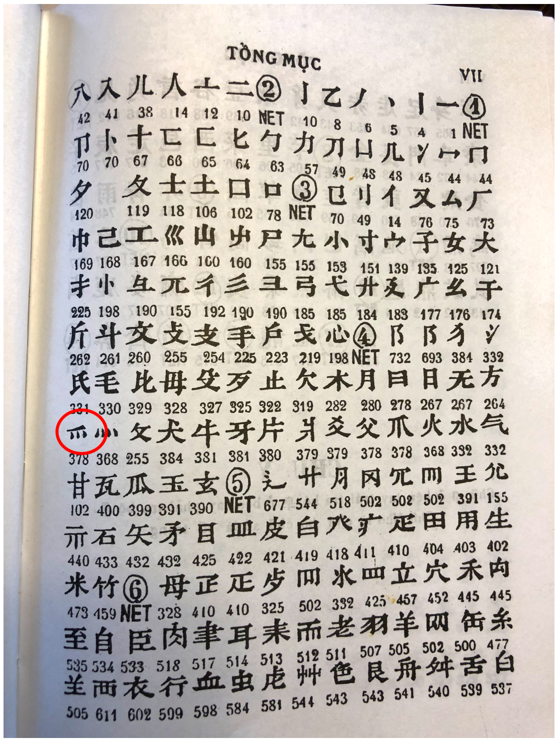
It is also possible that there is some kind of borrowing relationship between Vietnamese and Zhuang.

The 爫 variant of *Làm* was first encoded in TCVN 5773:1993 as V0-3D5B. When chữ Nôm was standardized in Unicode, V0-3D5B was unified with U+722B on the basis of the first source derivation described above.



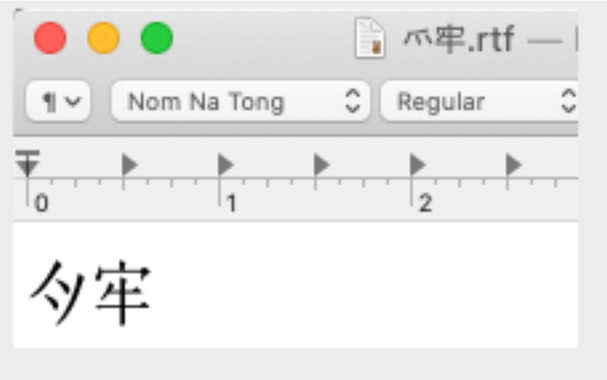
Based on our experience with the use of U+722B, this was a mistake.

U+722B is properly a variant of U+722A 爪, read *trảo,* meaning claw, talon, etc. It is used as the component form of Kangxi radical 87, and not cognate with the sense of *làm.* The problem arising from the lack of cognateness manifests itself in several ways. One is that Vietnamese also requires U+722B as a the component form of *trảo* for use in radical look up tables, illustrating character formation, etc. This is the proper use of U+722B and the form used in these cases matches the shape of U+722B as drawn in most CJK fonts, as illustrated below.


TỔNG MỤC

VII

Làm, should be drawn to occupy the full ideographic character box, so as to be in proportion with other Nôm characters. There are many situations where it is impossible to represent both forms without resorting styled text. And, since virtually every CJK font implements the component form of U+722B, where an application uses text that falls back to the system font, which is universally CJK, not Vietnamese, *làm* is rendered incorrectly with the component form of *trảo*. Below is an example from the Mac OS, where *làm* is rendered correctly in the document, but appears as the *trảo* component in the title.


爫牢.rtf —

Nom Na Tong

Regular

爫牢

Our proposal then is to add a new character representing *làm* in one of the CJK Unified Ideograph Ranges. It would have the same properties as U+722B, but the Unihan data would reflect the new semantics

U+xxxxx	
kIRG_VSource	V0-3D5B
kVietnamese	làm
kRSKangXi	87.0
kRSUnicode	4.3
kTotalStrokes	4

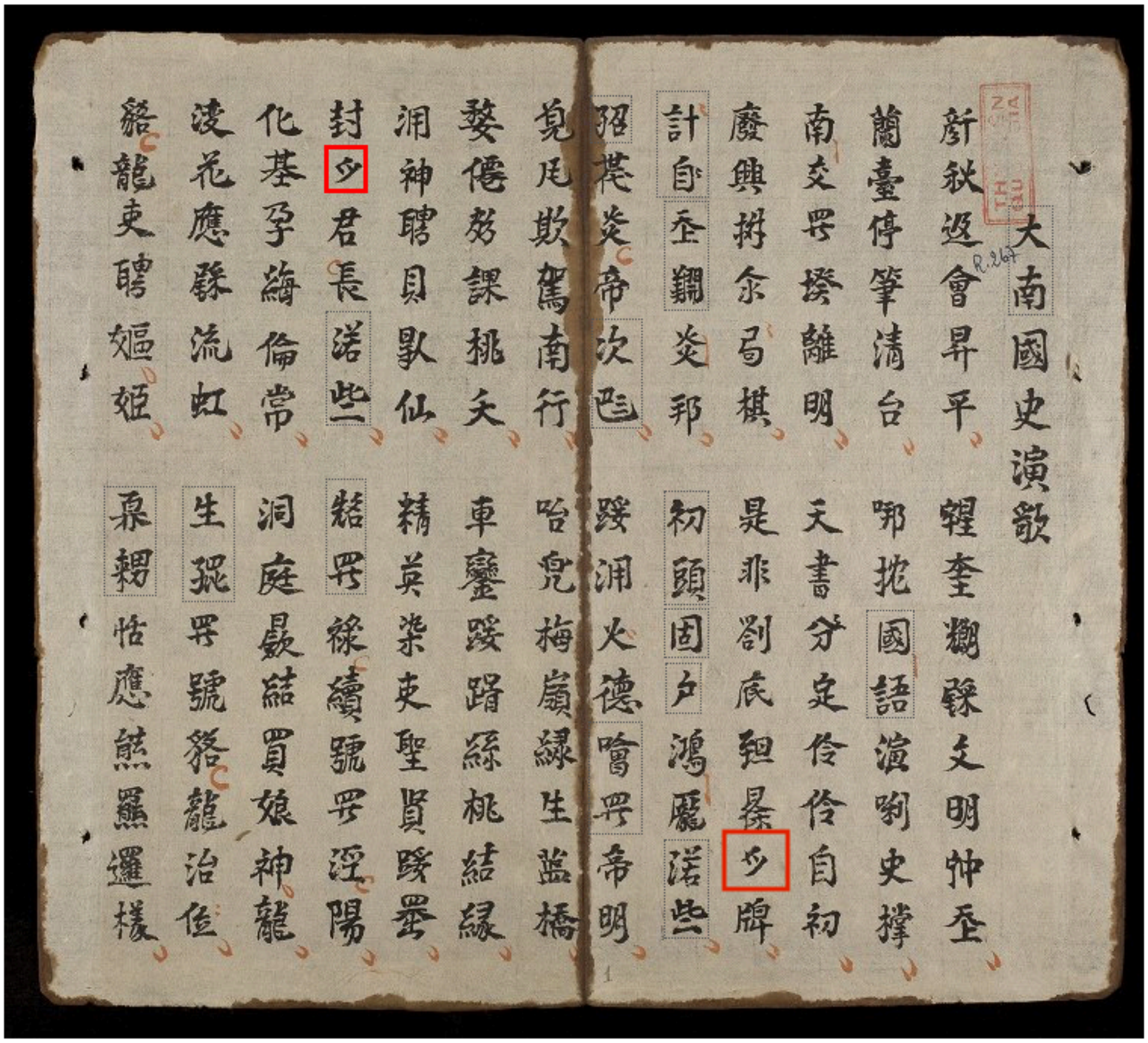
The V-Source glyph for U+722B would change to resemble the component forms shown above for the other sources. Also, two fields of the data for U+722B would likewise change :

U+722B	
kIRG_VSource	VU-0722B
kVietnamese	trảo

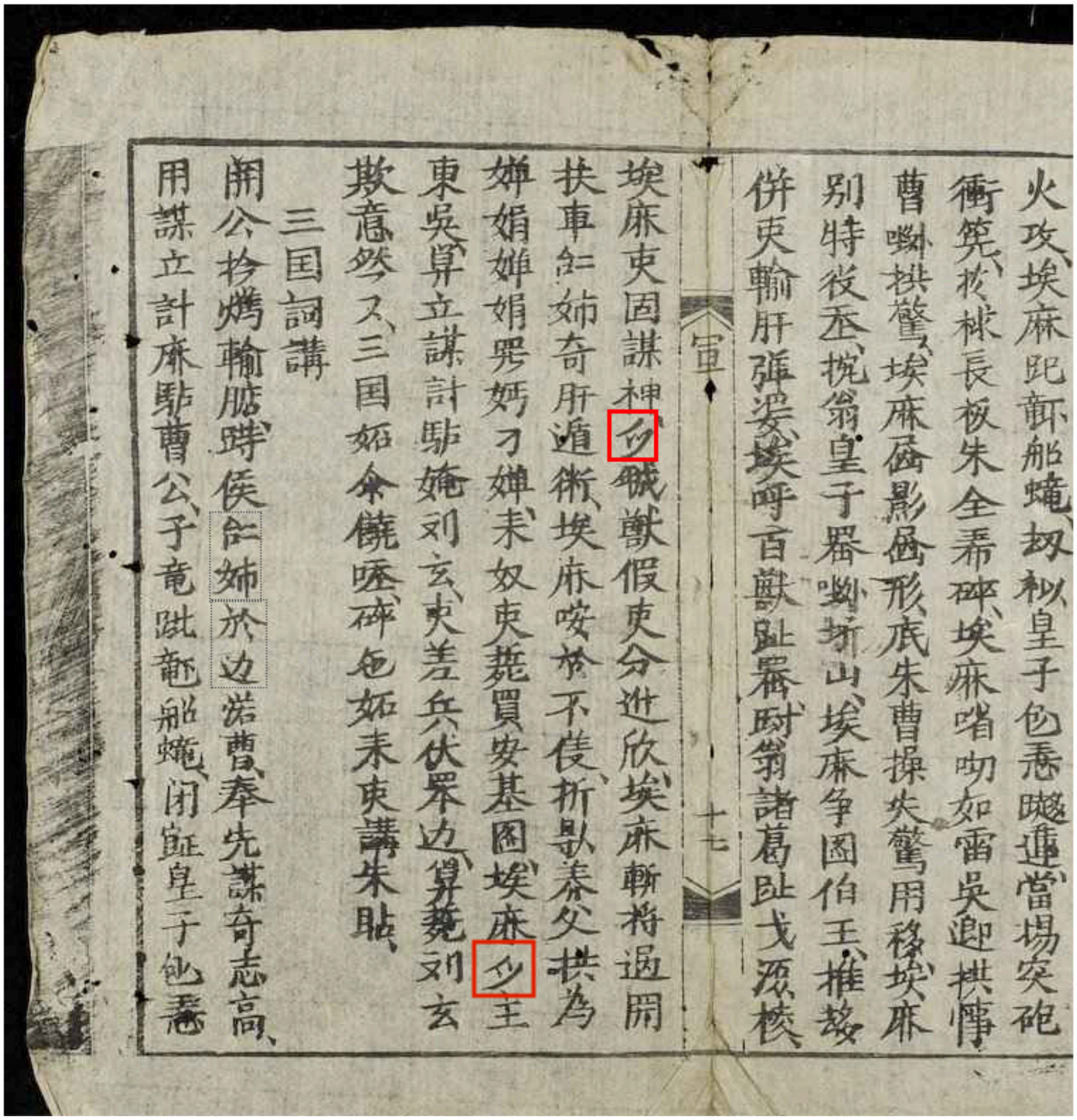
If China decides to standardize the Zhuang character *lamz*, they may also want to consider unification with the new code point. Although the semantics are different, there does appear to a relationship of borrowing or common derivation, that would justify unification. WS2017-03573:USAT05603, is a different matter. Although it resembles some of the forms of *làm*, the IRG might not consider this to be a valid candidate for unification based on the semantic. WS2017-03573:USAT05603 is a variant of the character, 肉 *nhục* (flesh), and has no relationship other than the similarity of some forms.

Appendix: Detailed examples of the *làm* glyph (source: https://www.chunom.org/pages/722B/)

1. *Đại Nam quốc sử diễn ca* (大南國史演歌), p1



2. *Trống quân tân truyện* (巈軍新傳), p9



火攻、埃麻𧿨𨔾船𦨭、㧅𥿗皇子包𢆥𩙻遊當塲突砲
衝𥯉、𢬣𥿗長𡋂朱全𢆥碎、埃麻𠺥吻如雷吳迎拱傳
曹𠰘拱驚、埃麻𡗶影𡗶形、底朱曹操失驚用移埃麻
別特𢧚歪、𢭸翁皇子𠱊𠰘𢬣山、埃麻爭圖伯王、推𢴒
併吏輸肝彈婆、埃呼百𠬠趾𨅹𢂞翁諸葛趾戈源核、

宣

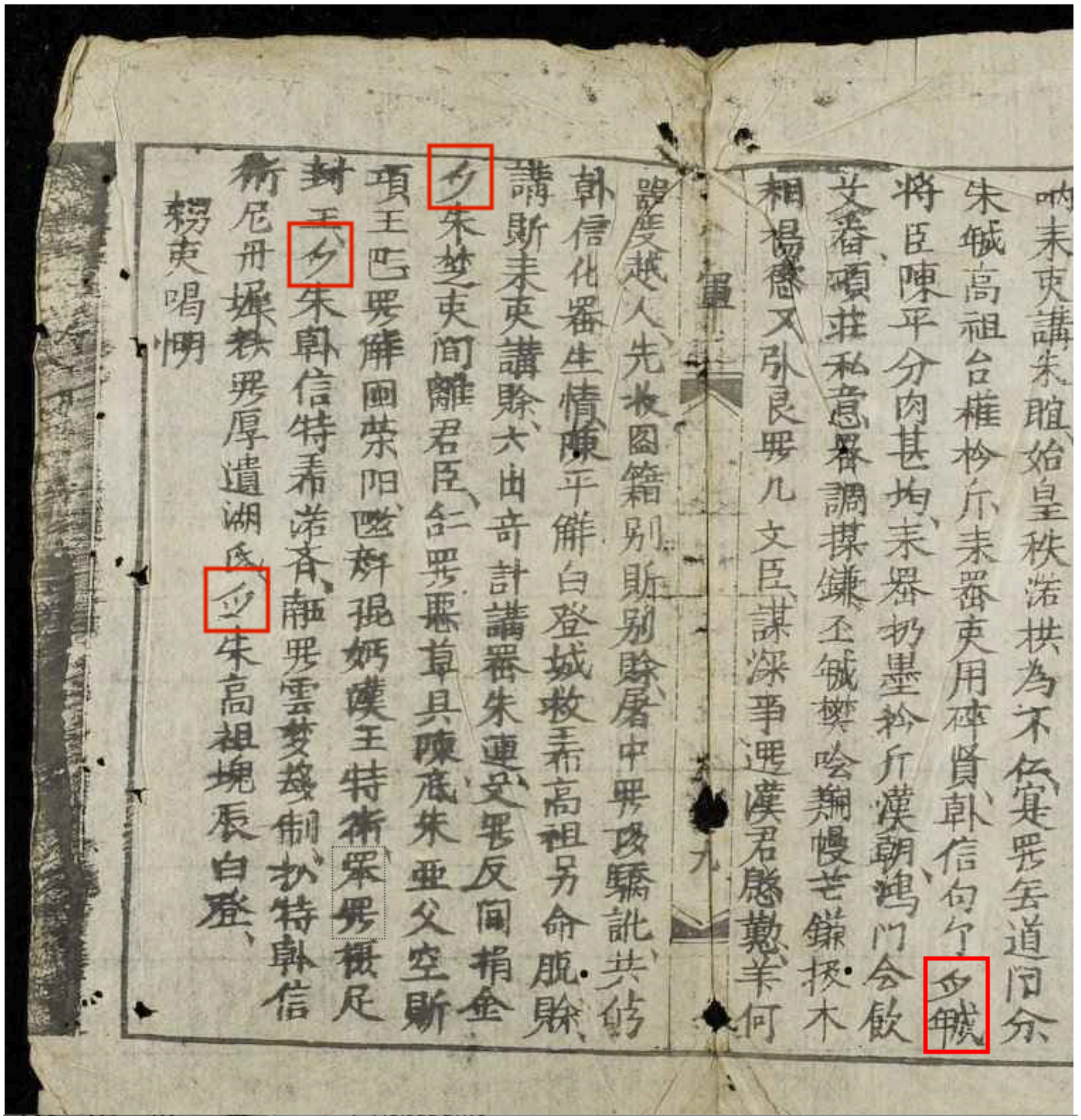
十七

埃麻吏固謀神彡𢧚𤞻假吏分進欣埃麻斬將過關
扶車𢚸姉奇肝遁術、埃麻𠯪於不倿、折𡢐秦父拱為
𡛔娟𡛔娟𤳆𡛔刁𡛔耒奴吏𦹵買安基圖埃麻彡主
東吳算立謀討駙𡢼劉玄、吏差兵伏𦊚边、算𦹵劉玄
撒意㝵ス、三国𡛔𡭧𩟡𠱋碎𠢞𡛔耒吏講朱貼、

三国詞講

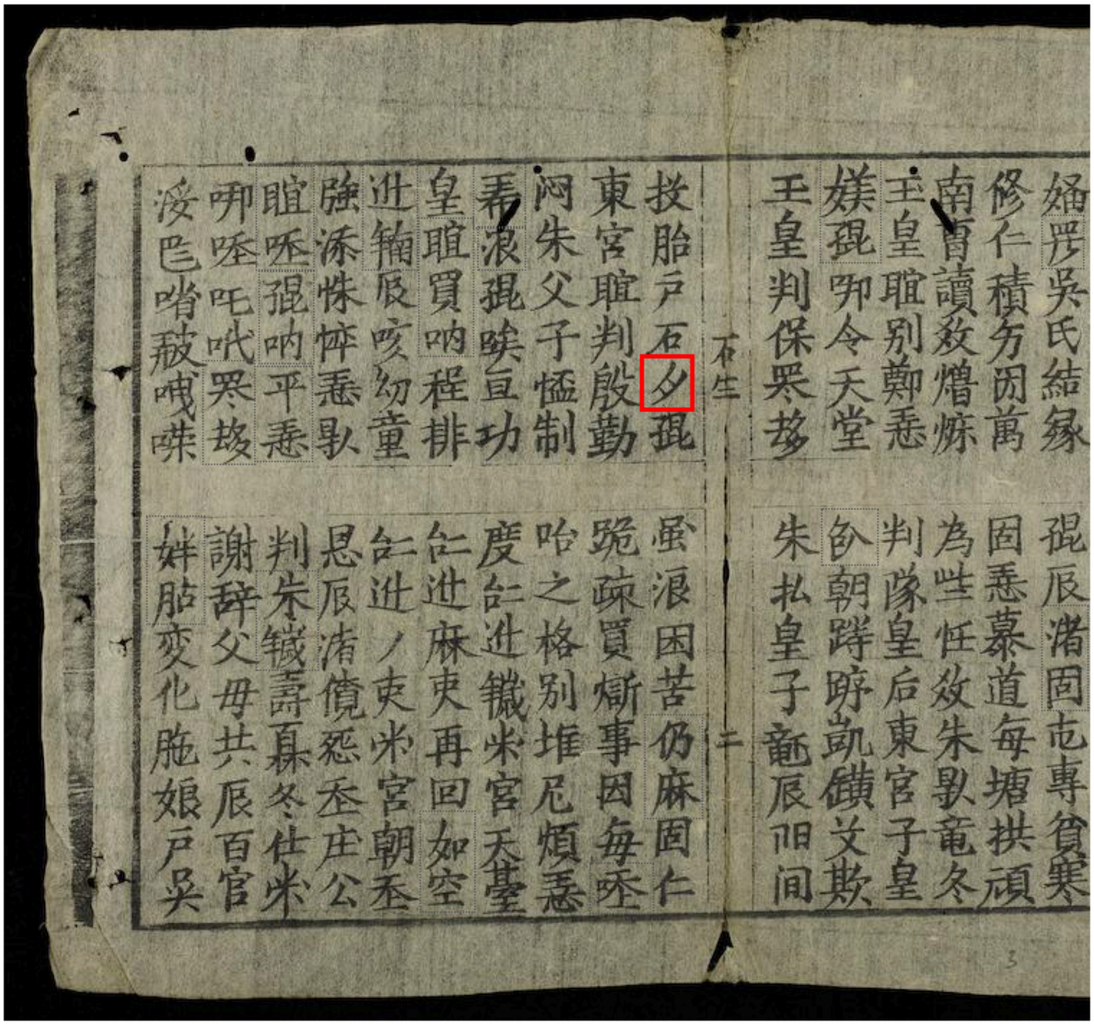
閧公𢬣𤈜輸𦛌𧿆侯𢚸姉於边𡤓曹、奉先謀奇志高、
用謀立計麻駙曹公、子竜𨇉𨗓船𦨭、閉𨁥皇子包𢆥

3. *Trống quân tân truyện* (藏軍新傳), p11



吶耒吏講朱聉、始皇秩渃拱為不仁寔罗缶道門分
朱鋮高祖台權抡斥耒器吏用碑賢、韓信句勺乡鋮
將臣陳平分肉甚均、耒𠸗扐壘衿斥漢朝鴻門會飲
芰畨頑莊私意畨調謀鐮丕鋮樊噲𡘯幔芒鐮捘木
相楊懲又引良罗几文臣、謀深爭遲漢君懇懃羊何
器度越人、先收圖籍別斯別賒居中罗於驕訛、共彷
韓信化器生情陳平𠊛白登城救吝高祖另命朧賒、
講斯耒吏講賒、大由奇計講器朱運受罗反間捐金
乡朱楚吏間離君臣、𠓨罗聽草具陳底朱亞父空斯
項王𠯹罗𠊛國榮陽、鹽斫湜媽漢王特衛寧罗𢬣足
封王、乡朱韓信特吝渃齊、靜罗雲夢𢧚制、扐特韓信
新尼冊𡏦𢧚罗厚遺湖氐、乡朱高祖塊辰白登、
𢪀吏唱𢠩

4. *Thạch Sanh tân truyện* (石生新傳), p2



姶罗吳氏結緣
修仁積务冈萬
南曾讀敎熷𤷱
主皇聀別鄭悉
媄琨卯令天堂
王皇判保哭孝

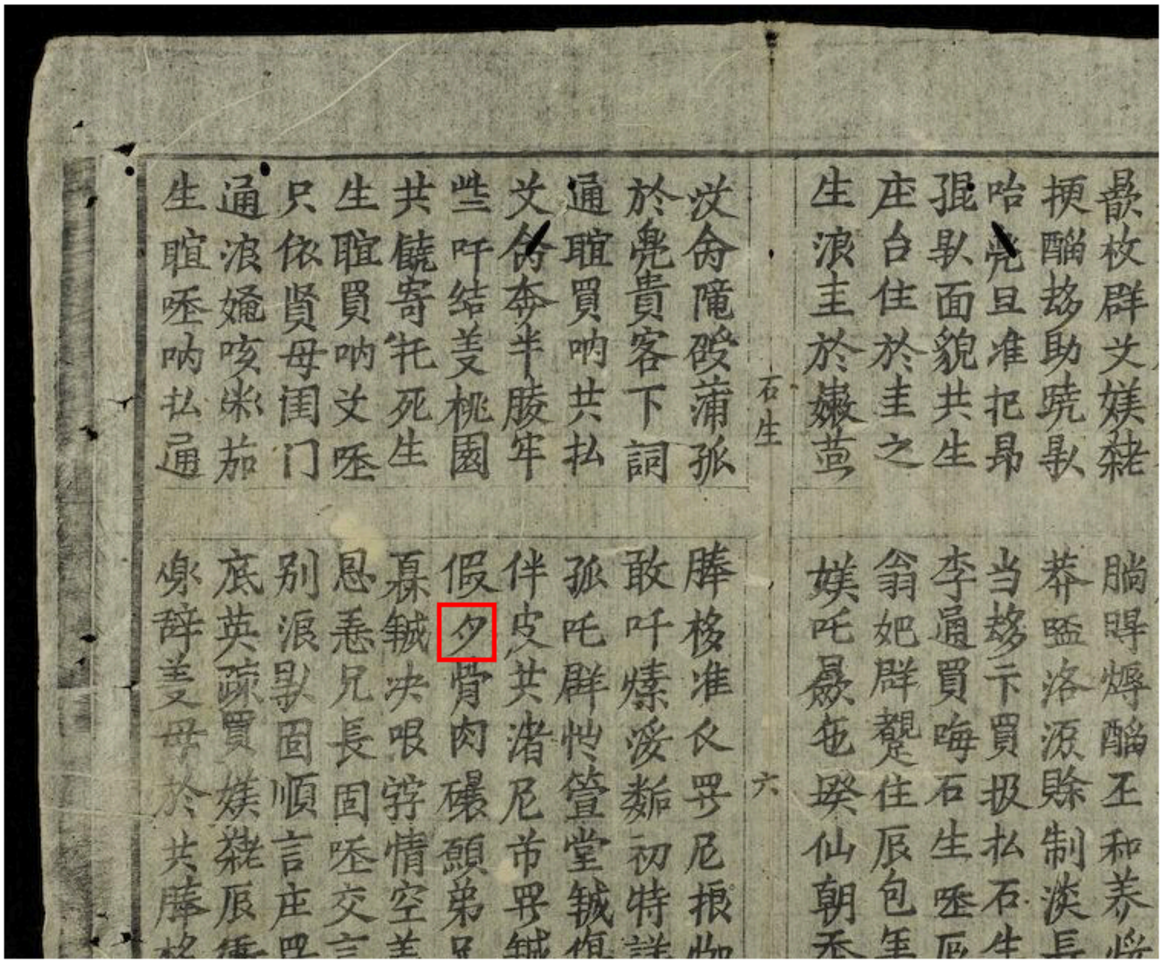
琨辰渚固屯專貧寒
固悉慕道每塘拱頑
為些仼效朱最竜冬
判隊皇后東宮子皇
𠓨朝蹲踈凱鐄爻欺
朱払皇子竜辰阳间

石生　二

扷胎戶石夕琨
東宮聀判般勤
罔朱父子慍制
秊浪琨暎亘功
皇聀買吶程排
进輴辰咳幼童
强添怵悴悉最
聀咥琨吶平悉
唧咥吒呔哭拸
涹邑哨蔑喊喋

蚤浪困苦仍麻固仁
䟫疎買𢢆事因每𡗶
咍之格別堆尼煩悉
度𠄼进鐄米宮天臺
𠄼进麻吏再回如空
𠄼进丿吏米宮朝𡗶
恳辰渚僥惡𡗶庄公
判朱鐵壽槇冬仕米
謝辞父母共辰百官
姅胎变化胣娘戶吳

5. *Thạch Sanh tân truyện* (石生新傳), p6



歇枚辟艾媄𦓅
挭醞拸助曉最
哈尭旦准把昂
𨀈最面貌共生
座台住於圭之
生浪主於嫌蒿

腨𥋳𡢐醞王和恙
𫏢監洛源賒制淡長
当𢀨下買扱払石生
李逼買晦石生咥辰
翁妃辟覰住辰包年
媄吒最奄娶仙朝香

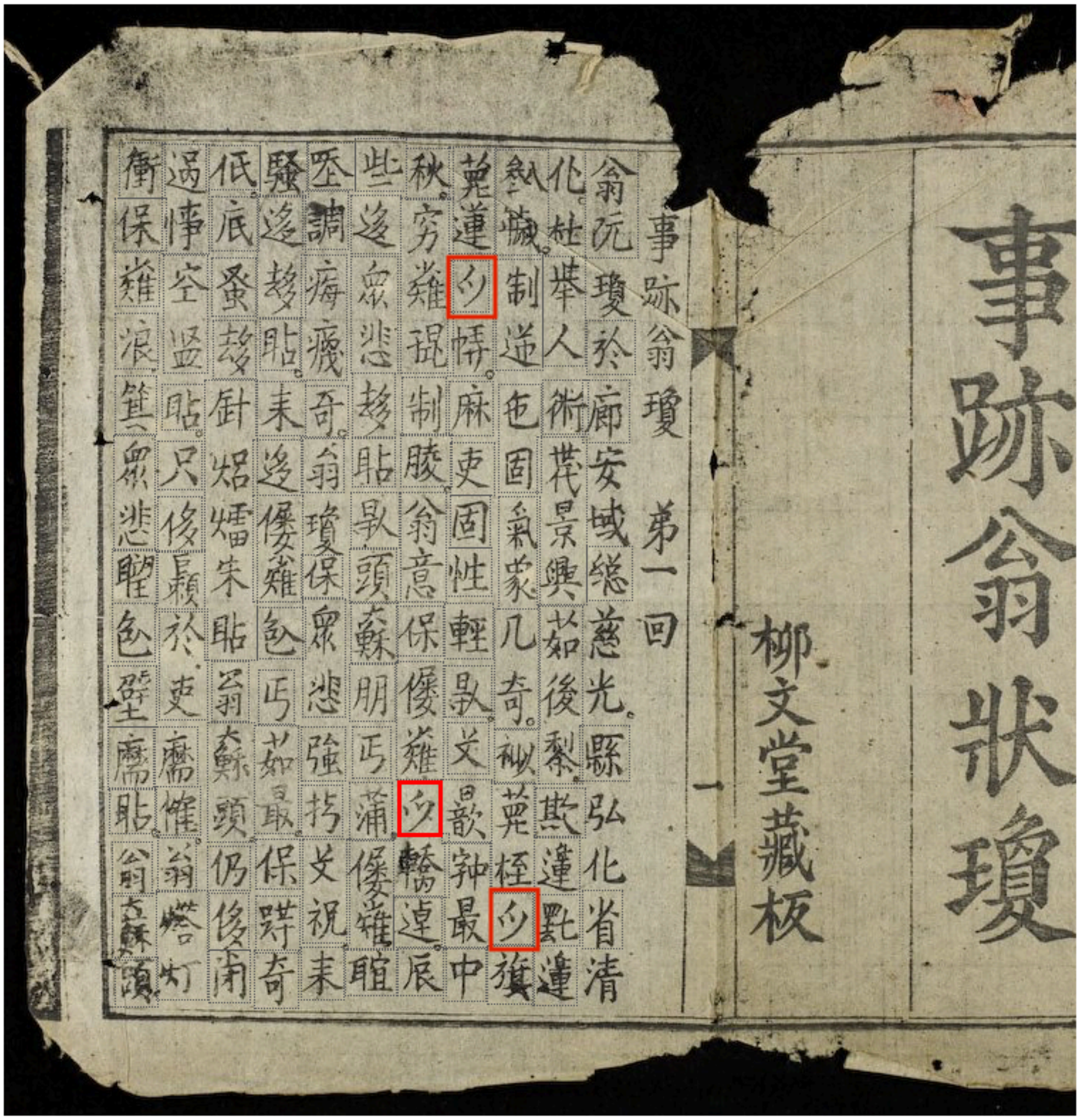
石生

六

汥𠇮庵破蒲孤
於兜貴客下詞
通聀買吶共払
艾𠇮奔半脖年
些吀結羌桃園
共饒寄托死生
生聀買吶艾咥
只依賢母闺门
通浪嬀咳粥茄
生聀咥吶払通

膵核准卮罗尼振
敢吀悰爰嘶初特
孤吒辟快登堂鉞
伴皮共渚尼节罗鉞
假夕骨肉𥗐頭弟
𡗶鉞決哏𥨨情空羌
恩悉兄長固咥交
別浪最固順言庄
底英疎買媄𦝄辰
𡞕辞羌母於共膵

6. Sự tích ông Trạng Quỳnh (事跡翁狀瓊), p.1



事跡翁狀瓊

柳文堂藏板

事跡翁瓊　第一回

翁阮瓊於廊安域総慈光。縣弘化省清化。杜𡾼人術茌景興茹後黎。欺𨖲點𨖲𢆥𢆥。制逆㐌固氣象几奇。祕莞桎𠫾獲莞蓮𠫾悸。麻吏固性輕㝡文歆狆最中秋。穷蘿調制䏾。翁意保優蘿𠫾轎違。辰些迻衆悲迻貼㝡頭蘇朋丐蒲。優蘿䁥丕調瘖癈奇。翁瓊保衆悲強扐安祝耒騷迻𡗊貼。耒迻優蘿色丐茹最保䟢奇低。底蚤𡗊斟焒熠朱貼翁蘇頭仍後鬧過悖空竖貼。只後顙於吏癮權。翁嗒灯衛保蘿浪箕衆悲瞠色壁癮貼。翁蘇頭

Bibliography

Guangxi Zhuangzu Zizhiqu Shaoshu Minzu Guji Zhengli Chuban Guihua Lingdao Xiaozu (ed.). *Sawndip Sawloih* (古壮字字典). Guangxi Minzu Chubanshe. Nanning, 2012.

Nguyễn Quang Hông. *Tự Điển Chữ Nôm Dẫn Giải*. [Dictionary of Chữ Nôm with source citations]. Social Sciences Publishing House, Hà Nội, 2014.