

Doc Type: ISO/IEC JTC1/SC2/WG2/IRG  
 Title: IRG N2798 Update of Vietnamese Glyphs with 戎 (nhung)  
 Source: Vietnam  
 Status: National Body Proposal  
 Authors: Eiso Chan, Lee Collins, Ngô Trung Việt  
 Action: For consideration by the IRG  
 Date: 2025-2-28

The Vietnamese reference font has used a non-standard design for the character and component 戎 (U+620E) from the beginning. Compare the design of V1-5635 to that of the glyphs from other sources shown below:

620E  
 戈 62.2

戎	戎	戎	戎	戎	戎
G0-4856	HB1-A6A5	T1-4826	J0-3D3F	K0-6B54	V1-5635

This is not actually the preferred form in Vietnamese usage. We believe now that the current shape evolved from some confusion with two similar characters, 戎 (U+620C) and 戎 (U+620D), as found in some texts. An example is in line 2454 of the 1866 edition of Kiều, shown below.



The context and other editions clearly indicate that 戎 is the correct character, since the reference is to fighting.

便宜扒勦役外董戎

Tiền nghi bát tiểu, việc ngoài đổng **nhung**.

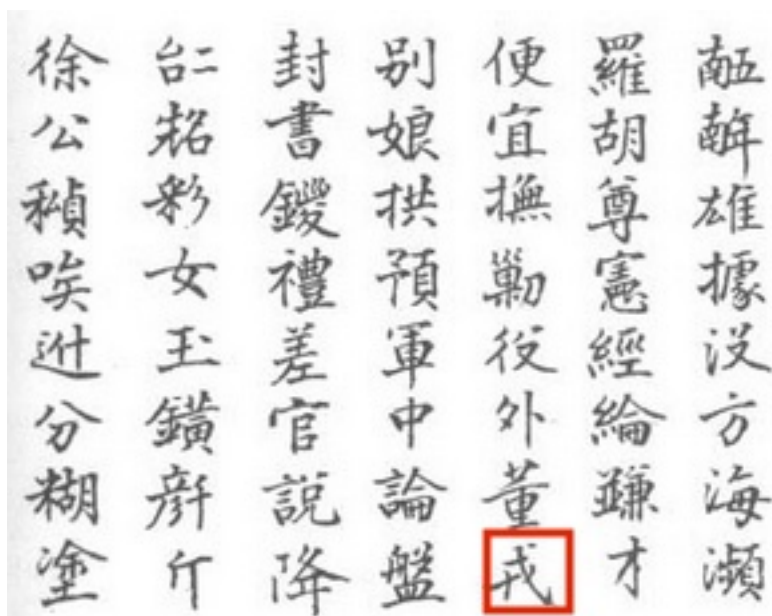
We know this because the Kiều is written in a meter known as *lục bát* which requires that the last character in this 8 word line should rhyme with the last character in the following 6 word line. In this case 雄 (hùng):

別徐罍等英雄

Biết Từ là đấng anh hùng

If 戎 were correct, we would expect a rime related to its reading, *tuất*.

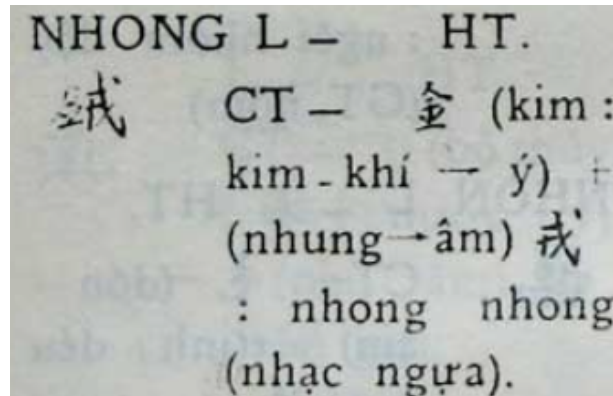
Indeed, we see the correct image in the 1870 version of Kiều shown below:



The phonetic element is a key part of compound Nôm characters, so it's important that reasonable phonetic values match the Vietnamese word that is being represented. The phonetics are typically related to the original Sino-Vietnamese reading of the character being used as phonetic. In the case of 戎, the Sino-Vietnamese reading is *nhung*. In composite characters, it can be used to represent similar sounds including *chông*, *nhong*, *ngông*, etc.

Problems arise when the Nôm characters are mapped to Unicode characters with the wrong phonetic. In the case of the phonetic 戎, we have found two such mismappings: notably U+928A and U+241F3.

In the case of U+928A, the original source for Vietnamese V0-4633 was Vũ Văn Kính's *Tự Điển Chữ Nôm*, p. 582. That work shows that the character should be composed of 金 and phonetic 戎, with a reading, *nhong*.



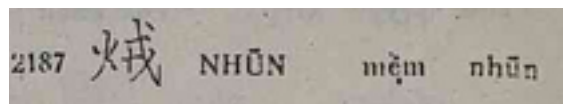
However, it was mapped in error to a character with the Sino-Vietnamese reading, *tuất*.



It should have been mapped to U+2CACD, which has the correct phonetic.



Similarly, the original source for V2-7B48 was Bảng Tra Chữ Nôm, # 2187, read *nhũn*, with phonetic 戎 (*nhung*) shown below.



This was mapped to U+241F3, which is composed of 火 and 戎.

241F3  
火 86.6

烱 烱 烱

UCS2003 GKX-0670.14 TF-306A

烱

V2-7B48

In fact, we have yet to find an example of U+241F3 used in a Vietnamese document, and therefore no corresponding reading. The correct mapping should have been U+241FD. This was somewhat rectified later by mapping the correct shape to U+241FD.

241FD  
火 86.6

烱 烱 烱

UCS2003 T5-2E4B VU-241FD

There are a number of cases where the semantic is mapped correctly, but the V-Source image is not appropriate. For example, U+8CCA.

8CCA  
貝 154.6

賊 賊 賊 賊 賊 賊

G1-5474 HB1-B8E9 T1-6670 J0-4231 K0-6E64 V1-682E

As shown below, the Vietnamese form should not differ significantly from that of the reference glyphs for other regions.

賊	Giặc, bellum ; hostes.
者	— giã, id.
吏	— lại, subito impetu trahere.
昆 碎	} tòi loạn con —, rebellis subditus, filius
— 亂	
— 打	đánh —, bellum facere.

Altogether, we have identified ten characters in which the design of the phonetic component 戎 should be changed to clarify the reading and the semantic. This includes U+620E itself and two characters from IRG Working Set 2024, [WS2024-03139](#) and [WS2024-04542](#). We propose to change the glyphs for the following encoded characters.

Code	TCVN	Old	New
U+620E	V1-5635	戎	戎
U+27295	V2-827D	𧈧	𧈧
U+8319	V1-6473	𧈩	𧈩
U+8CCA	V1-682E	𧈫	𧈫
U+2020B	V2-6E4A	𧈭	𧈭
U+27EEA	V2-8438	𧈯	𧈯
U+2803B	V0-4460	𧈱	𧈱
U+2A01A	V2-8836	𧈳	𧈳

Along with the glyph changes, we propose several changes to the Unihan data for four characters.

code	tag	current	new
U+928A	kIRG_VSource	V0-4633	VN-0928A
U+928A	kVietnamese	nhong	tuất
U+2CACD	kIRG_VSource		V0-4633
U+2CACD	kVietnamese		nhong
U+241F3	kIRG_VSource	V2-7B48	VN-241F3
U+241F3	kVietnamese	nhũn	
U+241FD	kIRG_VSource	VN-241FD	V2-7B48
U+241FD	kVietnamese		nhũn