

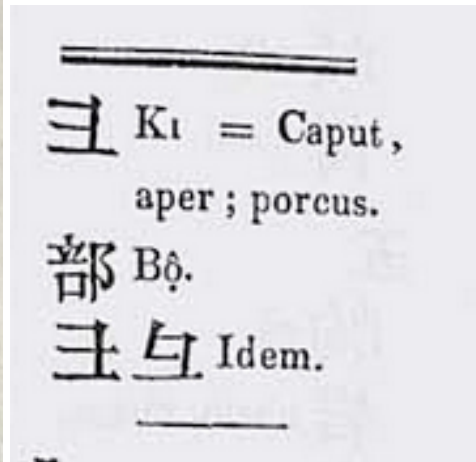
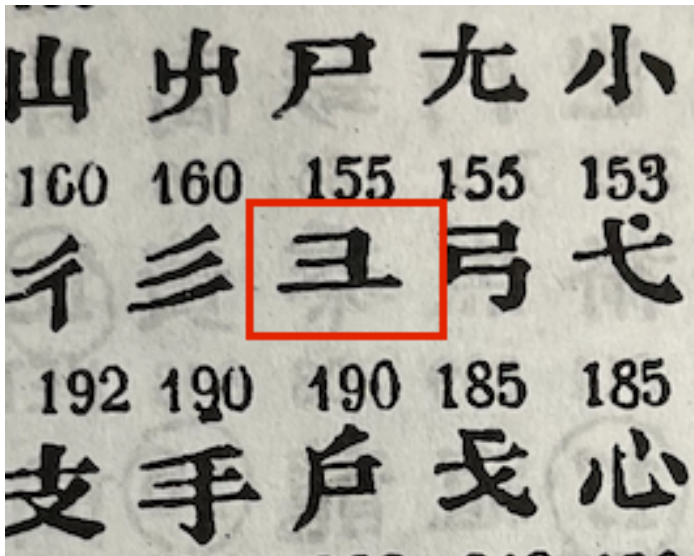
Universal Multiple-Octet Coded Character Set International Organization for Standardization

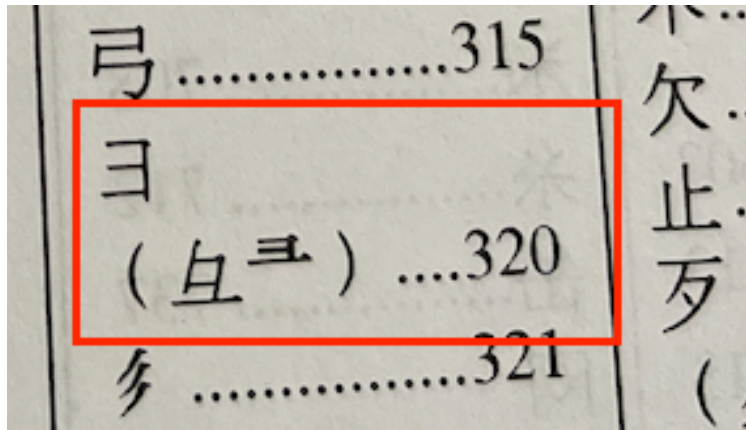
Doc Type: ISO/IEC JTC1/SC2/WG2/IRG
Title: Vietnam Response to IRGN2509
Source: Vietnam
Status: National Body Proposal
Authors: Eiso Chan, Lee Collins, Ngô Trung Việt
Action: For consideration by the IRG
Date: 2022-03-10

Two of the changes in IRGN2509 relate to characters used in Vietnam. Item #1 proposes changing the glyphic representation of U+5F50 / V1-5454 and item #3 proposes encoding two new characters with the representative glyphs 𠂇 and 𠂈.

Background

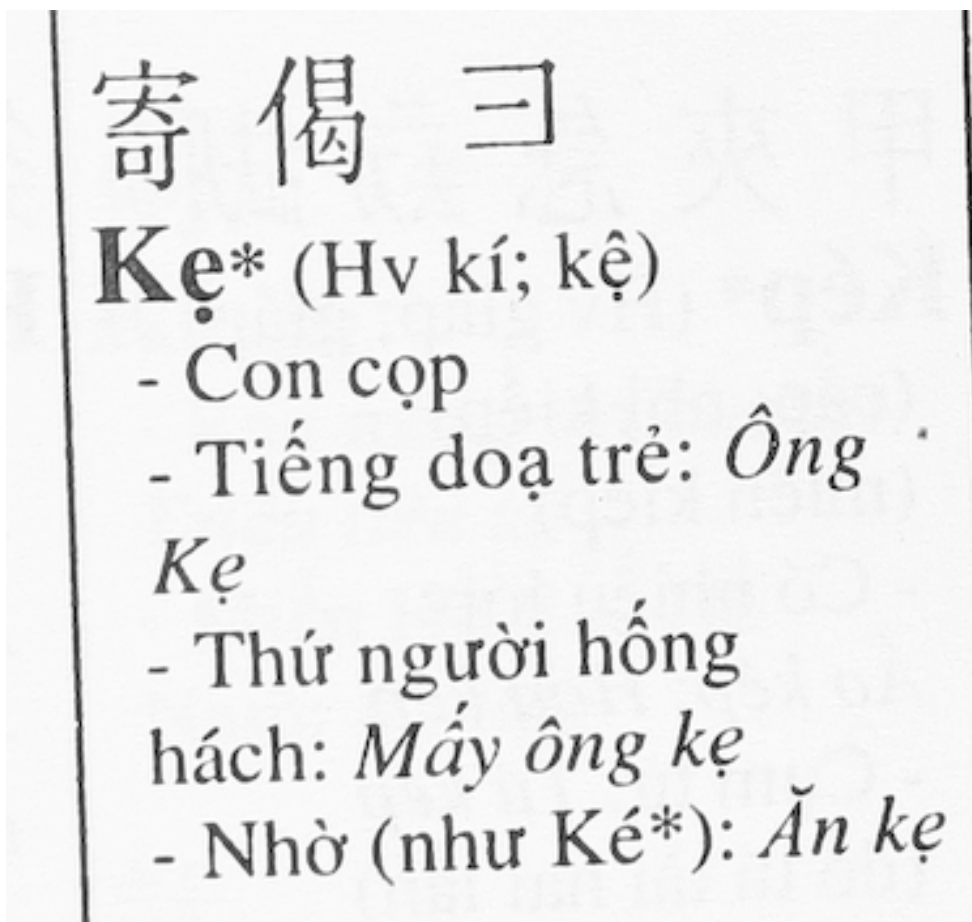
U+5F50 / V1-5454 is currently heavily overloaded in Vietnamese usage as found in various texts and dictionaries. It is used both as Kangxi radical #58, and as an independent character. As a radical, the shapes 𠂇 and 𠂈 are common, as seen here in Thiều Chừu's *Hán-Việt Tự-Điển*, Taber's *Dictionarium Anamitico-Latinum*, and Trần Văn Chánh's *Tự-Điển Hán-Việt Hán ngữ cổ đại và hiện đại*.





As an independent character there are 3 primary usages:

1. The "snow" form, 𠃉, is typically used for the Vietnamese words *kẹ* "tiger" and *kệ* "not pay attention", but as the explanation in the dictionary *Giúp Đọc Nôm và Hán Việt* shows below, the character is derived from the Sino-Vietnamese "kí", corresponding to Mandarin "jì", with the sense "head of a boar."



𠄎 偈 喝

Kê* (Hv kí; kê)

(khâu ½ kê)

Không đếm xỉa:

Mặc kê; Thấy kê

2. The form 𠄎, is is used in Nôm Tày for the word *mự*, meaning “turn” or “time”.

𠄎 MỰ²

Lân. ĐN. Vừa.

許 免 嗜 双 𠄎 三 𠄎

*Hầu mìn khấn soong mự
sam mự (Kh.T.Q.S)*

(Để cho nó gáy hai lần ba
lần).

𠄎 MÙA¹

Note, the above distinctions are based on our current data. It is possible that both of the above forms are interchangeable in meaning.

3. Finally, the form 𠄎 can also be found defined as a simplification of the character "snow":



Conclusion

Given the situation above, adoption of the changes proposed in IRGN2509 will require a long period for users to migrate to the new codes and shapes. However, we believe that long term separation of U+5F50 / V1-5454 into 𠄎 and 𠄏 is a positive move and support that. However, we consider that U+2E95 is already sufficient to represent the “long hand form” for our usage and will change the glyph in our font for that purpose. In sum, we propose the following changes:

1. Change the representative V-Source glyph of U+5F50 to 𠄎 and make the corresponding changes to the data:

U+5F50	kIRG_VSource	VN-05F50
U+5F50	kVietnamese	kí, mự

2. Encode a new character, tentatively at U+2B739, with the shape: 𠄏

*U+2B739	kIRG_VSource	V1-5454
*U+2B739	kVietnamese	kẹ, kệ

3. Additional data changes in UCD:

NamesList.txt.

2E95 CJK RADICAL SNOOT TWO

* actually a form of the radical for hand, despite its resemblance in shape

to the radical for snout

x 5F50

x 2B739

EquivalentUnifiedIdeograph.txt

2E95 ; 2B739 # CJK RADICAL SNOOT TWO