

Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation Internationale de Normalisation
Международная организация по стандартизации

Doc Type: Working Group Document

Title: Request to rename Jurchen to Jurchen Large Script

Source: Viacheslav Zaytsev and Andrew West

Status: Individual Contribution

Action: For consideration by JTC1/SC2/WG2 and UTC

Date: 2025-06-24

1. Introduction

As discussed in [WG2 N5309](#) = [L2/25-152](#), historical sources mention the existence of two different Jurchen scripts during the Jin 金 dynasty (1115–1234), referred to as Jurchen Large Script (*Nǚzhēn dàzì* 女真大字) and Jurchen Small Script (*Nǚzhēn xiǎozì* 女真小字), mirroring the Khitan Large Script (*Qidān dàzì* 契丹大字) and Khitan Small Script (*Qidān xiǎozì* 契丹小字) that were devised during the preceding Liao 遼 dynasty (907–1125). However, almost all surviving examples of Jurchen text from the Jin dynasty through to the Ming dynasty (1368–1644) are written in a single script. This script is similar in form and structure to the Khitan Large Script, so can be assumed to be the Jurchen Large Script. However, as it seemed to be the only surviving Jurchen script, it has generally been referred to simply as ‘Jurchen script’ (Chinese: *Nǚzhēnwén* 女真文) in academic literature. For example, the influential Jurchen dictionary compiled by Jīn Qǐzōng (1918–2004) is titled *Nǚzhēnwén Cídiǎn* 女真文辞典 [Dictionary of Jurchen script] = 秉岑使勇并 *dzuçien bitxə bunə-ku* [Mirror on Jurchen writing] (Běijīng: Wénwù chūbǎnshè, 1984).

For this reason the proposal to encode the surviving Jurchen script ([WG2 N5261R](#) = [L2/24-139](#)) submitted in 2024 by West and Sun asked for the block name and script name to be both simply called ‘Jurchen’, and its characters to be algorithmically named ‘JURCHEN IDEOGRAPH-XXXXX’.

However, during the 1970s and the 1980s several examples of a single short inscription separately engraved on gold and silver *páizǐ* 牌子 (travel passes or symbols of authority) were unearthed in northeast China. The short inscription closely resembles the Khitan Small Script, but the characters are not attested in other Khitan Small Script sources, and Aisin-Gioro Ulhicun, daughter of Jīn Qǐzōng, and a well-known scholar of Khitan and Jurchen studies in her own right, has convincingly argued that the inscription actually represents the Jurchen Small Script.

In 2025 Zaytsev and West submitted a proposal to encode the supposed Jurchen Small Script characters mentioned above ([WG2 N5309](#) = [L2/25-152](#)). The proposed encoding of these characters in the UCS may cause confusion for users of the standard as to what the ‘Jurchen’ script is. We propose to clarify that the script proposed by West and Sun in WG2 N5261R is actually the Jurchen Large Script, and the characters proposed for addition to the Khitan Small Script block by Zaytsev and West in WG2 N5309 is the Jurchen Small Script. This aligns with the block names for the currently encoded Khitan Small Script and the proposed Khitan Large Script ([WG2 N5319](#)).

2. Proposed Changes from WG2 N5261R

We propose the following changes with respect to WG2 N5261R:

Block name ⇒ Jurchen Large Script

Script name ⇒ Jurchen_Large_Script

Character names ⇒ Jurchen LARGE SCRIPT CHARACTER-XXXXX

Note 1:

WG2 N5261R proposed the algorithmic character names Jurchen IDEOGRAPH-XXXXX which was an indirect way of indicating that these are Jurchen Large Script characters. As we now propose to include ‘Jurchen Large Script’ in the character name, and as ‘ideographic’ is not an accurate description of the repertoire of logographic and phonographic Jurchen characters, we instead propose to use the algorithmic names Jurchen LARGE SCRIPT CHARACTER-XXXXX (cf. Khitan Small Script and Nüshu which both use the neutral term ‘character’ in their algorithmic character names).

Note 2:

The Unicode script code for Jurchen Large Script should be ‘Jurc’, which is the same as the allocated [ISO 15924](#) code for ‘Jurchen’. The English and French names associated with the ISO 15924 script code ‘Jurc’ could be revised to ‘Jurchen large script’ and ‘grande écriture jurchen’ respectively. There is no need to assign an ISO 15924 script code for Jurchen Small Script as Jurchen Small Script text could be identified using the language/script tag juc-Kits (‘juc’ is the ISO 639-3 language code for Jurchen).