

# Proposal to Encode the Zou Script and Numerals in Unicode

**Dr. T Marvin Zou**

**4 August 2026**

## 1: Short Summary of the Proposal

This proposal requests the formal encoding of the **Zou Script and Numerals** into the Universal Character Set (ISO/IEC 10646) and the Unicode Standard as a new script block.

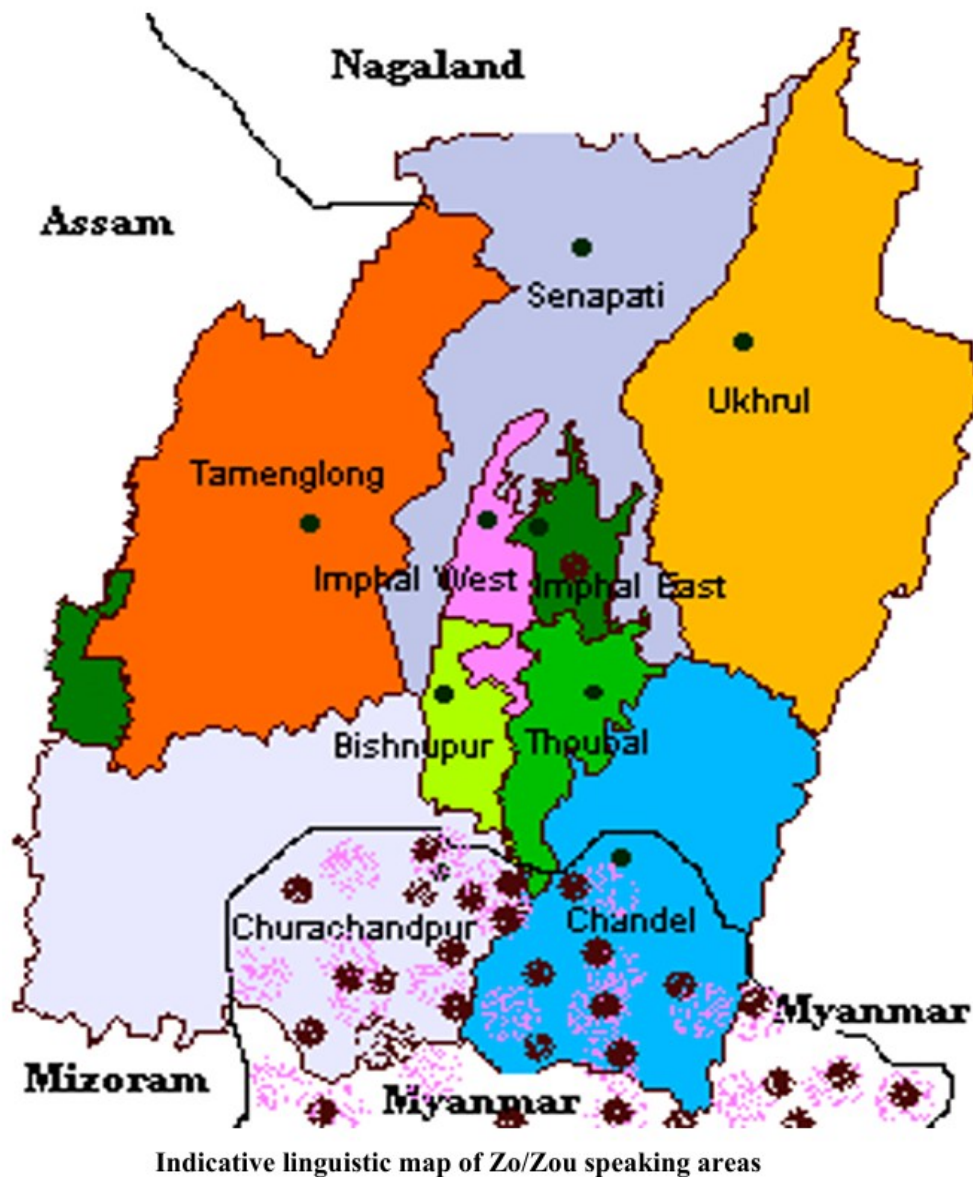
- **Scope:** The proposal defines the complete repertoire of letters, combining marks, and numerical digits required for the accurate writing and linguistic representation of the Zou language.
- **Supporting Assets:** It includes representative glyph shapes, authoritative metrics, character naming conventions, and proper Unicode properties (such as directional, combining, and line-breaking behaviors) to facilitate seamless digital text processing.

## 2 Background

### 2.1 Overview

The term Zou literally translates to "lofty hilly ranges". Classified officially as a Scheduled Tribe in India, the Zou community historically coexists alongside the Paite population, and their native tongue is similarly designated as the Zou language. The ethnonym encompasses both the language and the members of this distinct tribe, who are also referred to as Zoumi (derived from zou-, meaning "lofty hilly ranges," and -mi, meaning "people," which translates contextually to "highlanders" or "people of the lofty hilly ranges"). Geographically, the core population of this ethnic group is concentrated across the frontier highland corridors of Northeast India—predominantly spanning Manipur, Mizoram, Nagaland, Arunachal Pradesh, and Assam—as well as the contiguous border zones of Myanmar and Bangladesh. Within the scope of this morphosyntactic documentation, any topical reference to "Zou" explicitly denotes the language officially recognized by the Government of Manipur.

Genetically, Zou belongs to the Sino-Tibetan (Tibeto-Burman) language family, falling squarely within the Kuki-Chin sub-branch. It shares deep, immediate genetic, structural, and historical affinities with the Mizo macro-community of Northeast India and the various Chin language varieties of northwestern Myanmar. This reflects a shared ancestral lineage despite modern geopolitical boundaries, positioning Zou within a continuous dialect and culture chain that spans the frontier hills of South and Southeast Asia.



MAP OF MANIPUR & ADJOINING BOUNDARIES OF MYANMAR (BURMA)

figure 1 : Linguistic map of Zou speaking areas

### 3: Evolution and History of the Zou Script and its Usage

- **Origin:** The script was developed by **M. Siehzathang** of Churachandpur, Manipur in 1952 for writing the Zou language.



figure 2 & 3 : It shows “Zolai Simpat Bu”, which translates to “The first reading of Zou book”.

- **Early Recognition & Adoption:** It apparently took two decades for the script to attract much interest, but in the 1970s Zou organizations began expressing support for Siahzathang’s script, and the **United Zou Organization** adopted the Zou script as its official writing system for the Zou language in **1976**.

## PREFACE

The First & Second Edition of the Zou Primary Book of 1986 was out of stock. Hence, the Zou Script Development Society in its meeting held on 9.6.2007 at the residence of Pu T. Gouzadou, Ex-Minister, Zoveng, Churachandpur, unanimously resolved to publish the Third Edition of the book. After thorough and proper revision and adding some acceptable lessons, the book is now published in the form of revised and enlarged Edition.

For the preservation and Development of the Zou Script, the Society deputed Pu M. Siahzathang and Pu M. Nengchinlam to Central Institute of Indian Languages, Mysore vide their letter No. F. No. CEL/Zou/2007 - 08 dated 3<sup>rd</sup> July, 2007 and the Zou Script has been recorded as one of the Indian Scripts.

For the promotion of learning of the Zou Script, the Society sought permission to the authorities of the Govt. Primary Schools of Zou Area and Mission Schools to prescribe the Book as an additional subject. The Society is grateful for their co-operation and positive response.

Pu Ginsuanhau @ G.S. Haupu a benevolent Special Contractor selflessly agreed to publish the Book, when approached. The Society is very much indebted to Pu G.S. Haupu. "KALAI SAI"

Last but not the least the Society members are very much thankful to the Chairman Pu M. Siahzathang, who had been working heart and soul in the research work of the ZOLAI and nurturing it properly and laughed away the pains and hardships in preparing the manuscript of the book.

Suggestions for its improvement and correction of mistakes if any from readers are always welcome.

Thank you all.

The Zou Script Development Society,  
Churachandpur, Manipur, India.  
Regd. No. 316 of 2007

figure 4 : Zou Script early development

- **State Government Support:** Crucially, the Zou script received recognition and support from the Manipur state government, and the **Manipur State Kala Academy** acknowledged the work of Siahzathang by presenting him with awards in **1984** and **1991**.

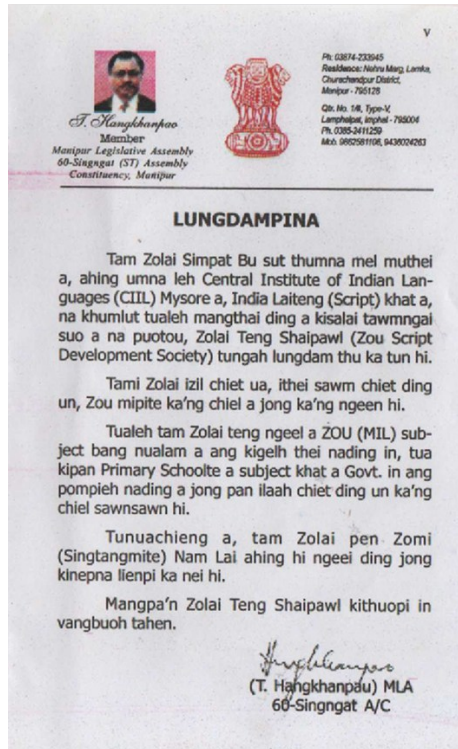


figure 5 : Government support

- **Promotion & Literature:** In the 1990s, literary organizations instituted working groups to manage the development and promotion of the script. The **Zou Literature Society (ZLS)** in Churachandpur began to promote the script by forming the **Zou Script Development Society (ZSDS)**, which in turn has published books to provide instruction in the script, including the *Zou Script Self-Instructor* and the *Zou Script Primary Book*.
- **Digital Integration:** The first Zou digital keyboard was invented by **Dr. T Marvin Zou** and published on **14 May 2026** in the official Keyman repository.

**Keyboard Details**

|                     |                                                                  |                        |                                             |
|---------------------|------------------------------------------------------------------|------------------------|---------------------------------------------|
| Keyboard ID         | zou_script_and_numerals                                          | Status                 | Experimental                                |
| Supported Platforms | Windows, macOS, Linux, Web, iPhone and iPad, Android, Mobile web | Package Download       | <a href="#">zou_script_and_numerals.kmp</a> |
| Author              | Dr. T Marvin Zou                                                 | Monthly Downloads      | 60                                          |
| License             | MIT                                                              | Total Downloads        | 194                                         |
| Documentation       | <a href="#">Keyboard help</a>                                    | Encoding               | Unicode                                     |
| Source              | <a href="#">experimental/zou_script_and_numerals</a>             | Minimum Keyman Version | 10.0                                        |
| Keyboard Version    | 11                                                               | Supported Languages    | Zou                                         |
| Last Updated        | 2026-05-26 15:52                                                 |                        |                                             |

Permanent link to this keyboard: [https://keyman.com/keyboards/zou\\_script\\_and\\_numerals](https://keyman.com/keyboards/zou_script_and_numerals)

figure 6 : Digital Integration in Keyman SIL

## 3.1 Usage

- Educational Materials:** The primary educational materials of the Zou language includes books from class 9,10, 11 & 12 also know as MIL. Under class 9 & 10 it contain books like “Chinthu Zaila” (Zou Literature), “Zou Course” (ZouLai) 9 & 10, and “Zou Grammar” 9 & 10.



figure 7 & 8 : Chinthu Zaila ( Zou Literature) for class 9 & 10

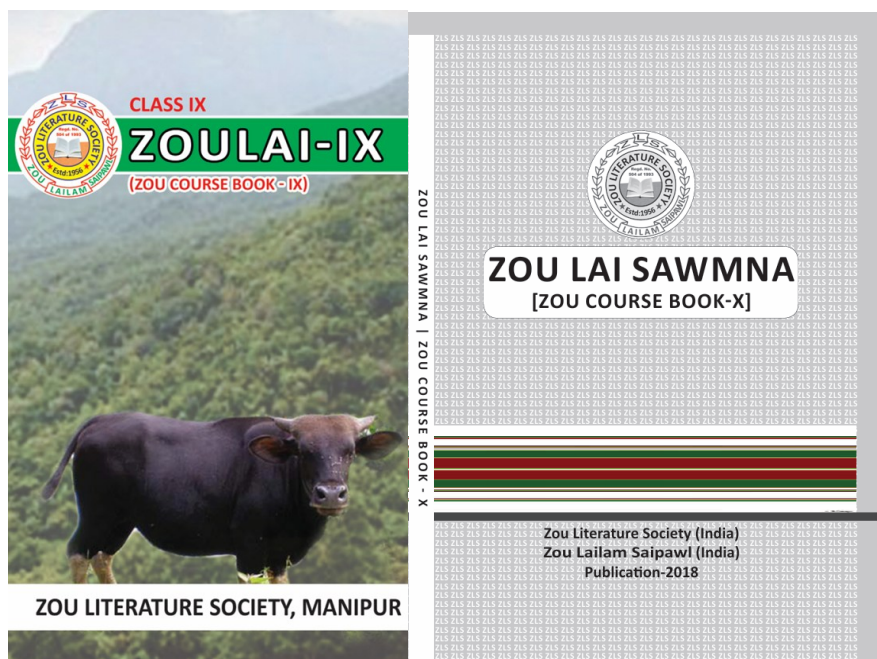


figure 9 & 10 : Zou course (Zoulai) for class 9 & 10

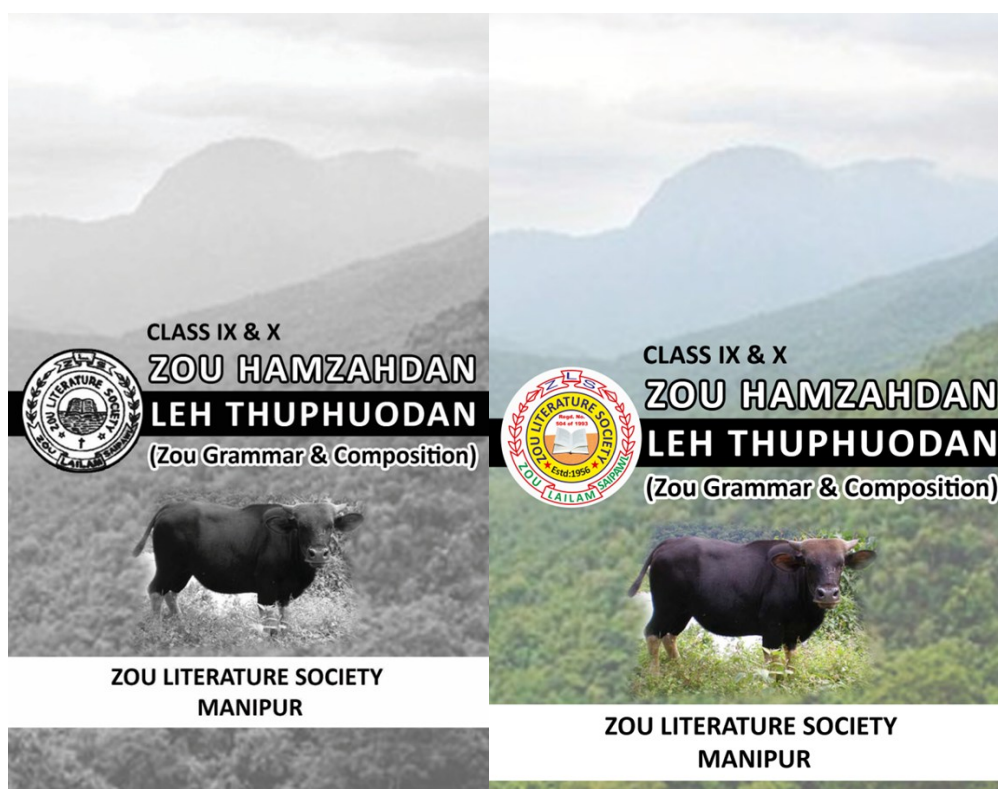


figure 11 & 12 : Zou Grammar and Composition for class 9 & 10

In class 11 and 12 it include books like “CHINTHU-ZAILA (Prose & Poetry)”, and “Zouham zahdan (Grammar) XI & XII”.

Prescribed as text book for class XI & XII by the Council of Higher Education,  
Manipur Vide its Order No.3/60/2004-HSC Dated th 9<sup>th</sup> June, 2006

**CHINTHU-ZAILA**  
(Prose & Poetry)

**ZLS Text Book Committee**  
T.C. Tungung, B.A., B.T., Chairman,  
James Khamchinsuon M.A., Secretary  
Liensuonkap, B.A (Hons), Member  
Aloysius T. Nehkhohang, B.A., B.D., Member  
Mannamaching, MCS, Member  
Cecilia Maman, M.A, Member,  
T. Lamkothawng, B.A., Member  
T. Chinzhakup, B.A., Member

Revised & Edited by  
Liensuonkap & T. Nehkhohang

Zou Literature Society  
Publications - 2012

**Chinthu-Zaila**  
(Prose & Poetry)

First Edition - 2005  
Copy - 500  
©ZLS - 2005 All Rights Reserved

Type Setting at  
Lamka Job Centre, Churachandpur, manipur  
Printed at  
Diamond Offset, Tedim Road, Churachandpur.

Published by  
The Zou Literature Society (ZLS)  
Post Box-48  
Zomi Colony, Churachandpur, Manipur.  
Pin-795128  
Email: zouzls@mail.com

figure 13 : CHINTHU-ZAILA (Prose & Poetry) for class 11 & 12



figure 14 : Zouham zahdan (Grammar) for class 11 & 12

- **Research Paper : “Zo Tonology”** by Philip Thangliemang Túngdim, Center for Linguistics, Jawaharlal Nehru University, New Delhi, India

## Zo Tonology

Philip Thangliemang Túngdim  
Center for Linguistics, Jawaharlal Nehru University, New Delhi, India  
philipzo\_64@rediffmail.com

**Abstract:** Zo<sup>1</sup> is a Northern Kuki-Chin language spoken in Manipur state, India and the Sagaing sub-division and, in Chin Hills of Myanmar respectively. It is a monosyllabic, agglutinating, nonrestricted tone language having *two minimal tone pairs* for most of its main verbs and adjectives, commonly referred to in Kuki-Chin literature as Stem1 and Stem2. Majority of the Zo<sup>2</sup> verbs undergo morphophonemic changes called *Lax-Tense alternations* or *Tone Alternations* or *Morpho-tonemic process*. Since, it was first noted, some useful studies have been done. This paper describes the types of tones attested in Zo in particular, with the appearances of Stem2 verb tones, and it attempts to exhibit the pragmatic tone rules of this language in general.

**Keywords:** minimal tone pairs, Stem1 and Stem2, morpho-tonemic process.

<sup>1</sup> The name of the language is also variously written as Zou, Jou, Yo, Yoe, Zome, Zomi, Kuki-Chin.

<sup>2</sup> The terminology is used by Khoi Lam Thang (2001); Button (n.d). Here, I am using the practice followed by renowned linguists. Zo is a close cognate of Tiddim Chin. In India, Zou is used for the same language.

<sup>3</sup> Tiddim is loconym of Tedim as explained by Khoi Lam Thang (2001) is a language spoken in Tedim and other areas of the Chin Hills of Myanmar (Burma).

figure 15 : Research Paper on the topic “**Zo Tonology**” by Dr. Philip Thangliènmang  
Túngdìm IAS (Retired)

## 4: Need for Public Interchange

To successfully present a case for encoding the Zou script and numerals within the Unicode Standard and ISO/IEC 10646, it must be demonstrated that character encoding is required for the public interchange of information in plain text.

- **Requirement for Plain Text Interchange:** The Zou script and its associated numerals are actively used by the Zou-speaking community for modern communication, education, documentation, and publishing. Without standard Unicode encoding, text data cannot be reliably exchanged, searched, indexed, or processed across different operating systems, web browsers, databases, and digital communication platforms without relying on proprietary or non-standard font hacks (such as private use area mappings).
- **Community and Technological Integration:** Standardizing the script ensures that the Zou community can fully participate in the digital ecosystem, enabling plain text messaging, digital archiving, web publishing, and multi-platform interoperability.

## 5: Writing System

Zou (also known as Zo or Zomi) primarily uses a **Latin alphabet**, which was introduced by Christian missionary J.H. Cope around 1910 and remains the standard for schools and churches. Additionally, there is an indigenous alternative script called **Zoulai** (or Zolai), created in 1952 by M. Siahzathang of Churachandpur, Manipur.

### 5.1 Structure

Key characteristics of Zou writing systems include:

- **Latin Script:** Developed in the early 20th century, this orthography is widely used for religious texts, literature, and education, though it lacks standardized capitalization rules in some contexts.
- **Zoulai Script:** This is a hybrid system combining alphabetic and alphasyllabic features, written from **left to right**. It was originally developed with 25 characters, later expanded to 27, and includes numerals. While adopted by cultural organizations in the 1970s, its usage has fluctuated, and it remains unencoded in Unicode as of 2025.
- **Phonology:** The Zou language is tonal, featuring three contrastive tones (level, low-rising, and falling) that distinguish meaning, along with 22 consonant and seven vowel phonemes. The Zoulai script itself does not inherently mark tonal information.

### 5.2 Zou Script Consonants, Vowels and Numerals. Also its IPA and Latin Script

| Zou Script |           | Latin Script | IPA Symbols |
|------------|-----------|--------------|-------------|
| Uppercase  | Lowercase |              |             |
| ᐃ          | ᐃ         | A            | [a]         |
| ᐅ          | ᐅ         | Ba           | [b]         |
| ᐇ          | ᐇ         | Ou           | [o]         |
| ᐉ          | ᐉ         | Da           | [d]         |
| ᐋ          | ᐋ         | Ri           | [r]         |
| ᐍ          | ᐍ         | Ga           | [g]         |
| ᐏ          | ᐏ         | Ha           | [h/?]       |
| ᐑ          | ᐑ         | Fi           | [f]         |
| ᐓ          | ᐓ         | Na           | [n]         |
| ᐕ          | ᐕ         | I            | [i~j]       |
| ᐗ          | ᐗ         | Ji           | [ʃ]         |
| ᐙ          | ᐙ         | Chi          | [ç]         |
| ᐛ          | ᐛ         | Shi          | [ʃ]         |
| ᐝ          | ᐝ         | Ka           | [k]         |

|      |    |     |                   |
|------|----|-----|-------------------|
| စ    | ဆ  | O   | [o]               |
| ပြ   | ပေ | Pa  | [p]               |
| ဧ    | အ  | E   | [e/ə]             |
| ဝိ   | ဝ  | Khe | [k <sup>h</sup> ] |
| မ္မိ | မ  | Mo  | [m]               |
| ဗျ   | ဗ  | Nga | [ŋ]               |
| ဗ    | လ  | Sa  | [s]               |
| တံ   | တ  | Ta  | [t]               |
| လှ   | လ  | U   | [u~w]             |
| ဝေ   | ဝ  | Va  | [v]               |
| လ    | လ  | Za  | [z]               |
| ဗျ   | ဗ  | Y   | [j]               |
| လှ   | လ  | La  | [l]               |
|      | ဝ  | 0   |                   |
|      | ၇  | 1   |                   |
|      | ၉  | 2   |                   |

|  |   |   |  |
|--|---|---|--|
|  | ၁ | 3 |  |
|  | ၂ | 4 |  |
|  | ၃ | 5 |  |
|  | ၄ | 6 |  |
|  | ၅ | 7 |  |
|  | ၆ | 8 |  |
|  | ၇ | 9 |  |

### 5.3 Vowel Signs, Special Signs, Akaihuoi (Germination of Vowels, Vowel Modifiers), and Punctuations

|                                          |               |                |           |
|------------------------------------------|---------------|----------------|-----------|
| Vowel signs                              |               |                |           |
| ၁                                        | ၂             | ၃              | ၄         |
| /a/                                      | /o/           | /ɔ/            | /e/       |
| • Small letters also used as vowel signs |               |                |           |
| ၅                                        | ၆             | ၇              | ၈         |
| /i/                                      | /ɔ/           | /e/            |           |
| Special signs                            |               |                |           |
| ၉                                        | ၁၀            | ၁၁             | ၁၂        |
| Akaihuoi                                 | Low tone Mark | High tone Mark | Nuhdong   |
| • Akaihuoi:                              |               |                |           |
| • Germination of Vowels:                 |               |                |           |
| ၁၃                                       | ၁၄            | ၁၅             | ၁၆        |
| aa                                       | ou            | ii             | ee        |
| • Vowel modifier:                        |               |                |           |
| ၁၇                                       | ၁၈            |                |           |
| aw                                       | w             |                |           |
| Punctuations:                            |               |                |           |
| ၁၉                                       | ၂၀            | ၂၁             | ၂၂        |
| Question                                 | Exclamatory   | Comma          | Full stop |
| ?                                        | !             | ,              | .         |

figure 16 : Zou Vowels Special Signs, Akaihuoi (Germination of Vowels, Vowel Modifiers), and Punctuations

## 5.4 Not Casing, Carry Vowel Signs, Diphthongs, Conjunctions, and Consonant Vowel Ligature

| <p>         ຊຽງ ລຸ່ມ ລຸ່ມ ລຸ່ມ<br/>         ຊຽງ ລຸ່ມ ລຸ່ມ ລຸ່ມ       </p> |     |     |     |     |     |
|---------------------------------------------------------------------------|-----|-----|-----|-----|-----|
| ຊຽງ                                                                       | ຊຽງ | ຊຽງ | ຊຽງ | ຊຽງ | ຊຽງ |
| ie                                                                        | uo  | ph  | th  | ge  | vo  |
| ຊຽງ                                                                       | ຊຽງ | ຊຽງ | ຊຽງ | ຊຽງ | ຊຽງ |
| lh                                                                        | fu  | tl  | ke  | hr  | tia |
| ຊຽງ                                                                       | ຊຽງ | ຊຽງ | ຊຽງ | ຊຽງ | ຊຽງ |
| ua                                                                        | ai  | ia  | be  | jh  | ei  |

- Not casing, carry vowel signs except (ຊຽງ)
- Diphthongs (red)
- Conjunctions (blue)
- Consonant vowel ligature (green)

figure 17 : Not Casing, Carry Vowel Signs, Diphthongs, Conjunctions, and Consonant Vowel Ligature of the Zou language

## 6 Script Images / Representative Glyphs (Keyboard Layout Mappings) and its Stability

The following layout references display the character arrangements and representative glyphs utilized across standard and shifted keyboard input states for the Zou script:

- **Unshifted and Shifted Keyman Layout:**

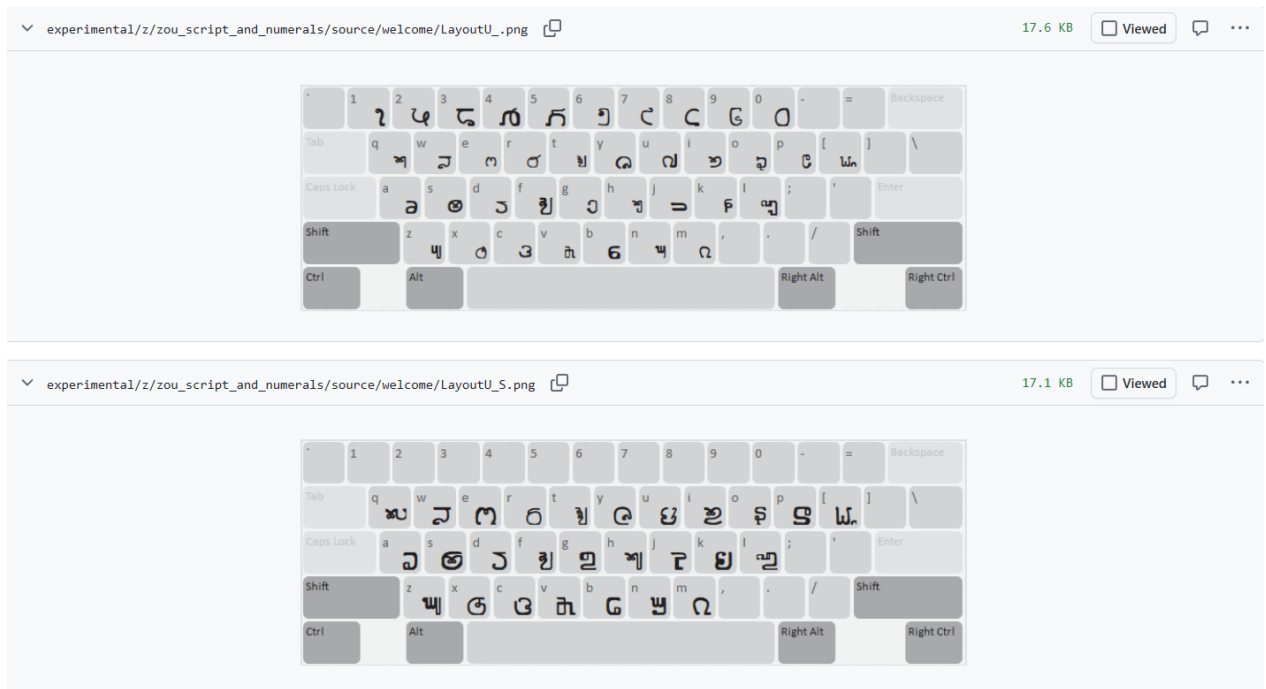


figure 18 : Unshifted and Shifted Keyman Layout of Zou

- **Stability**



figure 19 : Foundation Stone of Zou Script and Numerals in Behiang (Place name)

## 7. Structure, Unicode PUA Mapping & Character Inventory

To support digital rendering and provisional testing, the typeface maps individual character glyphs across the standard Unicode Private Use Area (PUA) range, incorporating dedicated buffer zones (U+E04A to U+E07F) for future script expansions, punctuation marks, and rendering ligatures:

- U+E000 – U+E01A: Core Consonants and Base Letterforms
- U+E020 – U+E03A: Vowel Signs, Medials, and Combining Marks
- U+E040 – U+E049: Native Zou Numerals (0 through 9)
- U+E04A – U+E07F: **PUA Buffer Zone** (Reserved for future character additions, punctuation, context-specific glyph variants, and rendering ligatures)

| Zou Letter (Phonetic) | Type                | PUA Code        | Unicode Name / Description                  |
|-----------------------|---------------------|-----------------|---------------------------------------------|
| A – La (Core Letters) | Consonants / Vowels | U+E000 – U+E03A | Uppercase & Lowercase Letterforms           |
| 0 – 9 (Numerals)      | Native Numerals     | U+E040 – U+E049 | Zou Digits Zero through Nine                |
| Buffer Zone           | Reserved PUA        | U+E04A – U+E07F | Future Expansions, Punctuation, & Ligatures |

## 8 Character Properties

- **Unicodedata.txt**  
E000;ZOU CAPITAL LETTER A;Lu;0;L;;;;N;;;E040;E040;  
E001;ZOU CAPITAL LETTER B;Lu;0;L;;;;N;;;E041;E041;  
E002;ZOU CAPITAL LETTER C;Lu;0;L;;;;N;;;E042;E042;  
E003;ZOU CAPITAL LETTER D;Lu;0;L;;;;N;;;E043;E043;

E004;ZOU CAPITAL LETTER E;Lu;0;L;;;;N;;;E044;E044;  
 E005;ZOU CAPITAL LETTER F;Lu;0;L;;;;N;;;E045;E045;  
 E006;ZOU CAPITAL LETTER G;Lu;0;L;;;;N;;;E046;E046;  
 E007;ZOU CAPITAL LETTER H;Lu;0;L;;;;N;;;E047;E047;  
 E008;ZOU CAPITAL LETTER I;Lu;0;L;;;;N;;;E048;E048;  
 E009;ZOU CAPITAL LETTER J;Lu;0;L;;;;N;;;E049;E049;  
 E040;ZOU SMALL LETTER A;Ll;0;L;;;;N;;E000;;E000;  
 E041;ZOU SMALL LETTER B;Ll;0;L;;;;N;;E001;;E001;  
 E042;ZOU SMALL LETTER C;Ll;0;L;;;;N;;E002;;E002;  
 E043;ZOU SMALL LETTER D;Ll;0;L;;;;N;;E003;;E003;  
 E044;ZOU SMALL LETTER E;Ll;0;L;;;;N;;E004;;E004;  
 E045;ZOU SMALL LETTER F;Ll;0;L;;;;N;;E005;;E005;  
 E046;ZOU SMALL LETTER G;Ll;0;L;;;;N;;E006;;E006;  
 E047;ZOU SMALL LETTER H;Ll;0;L;;;;N;;E007;;E007;  
 E048;ZOU SMALL LETTER I;Ll;0;L;;;;N;;E008;;E008;  
 E049;ZOU SMALL LETTER J;Ll;0;L;;;;N;;E009;;E009;  
 E020;ZOU DIGIT ZERO;Nd;0;L;;0;0;0;N;;;;;  
 E021;ZOU DIGIT ONE;Nd;0;L;;1;1;1;N;;;;;  
 E022;ZOU DIGIT TWO;Nd;0;L;;2;2;2;N;;;;;  
 E023;ZOU DIGIT THREE;Nd;0;L;;3;3;3;N;;;;;  
 E024;ZOU DIGIT FOUR;Nd;0;L;;4;4;4;N;;;;;  
 E025;ZOU DIGIT FIVE;Nd;0;L;;5;5;5;N;;;;;  
 E026;ZOU DIGIT SIX;Nd;0;L;;6;6;6;N;;;;;  
 E027;ZOU DIGIT SEVEN;Nd;0;L;;7;7;7;N;;;;;  
 E028;ZOU DIGIT EIGHT;Nd;0;L;;8;8;8;N;;;;;  
 E029;ZOU DIGIT NINE;Nd;0;L;;9;9;9;N;;;;;

- **Collation**

\uE000 < \uE040 \uE001 < \uE041 \uE002 < \uE042 \uE003 < \uE043 \uE004 <  
 \uE044 \uE005 < \uE045 \uE006 < \uE046 \uE007 < \uE047 \uE008 < \uE048 \uE009  
 < \uE049 \uE00A < \uE04A \uE00B < \uE04B \uE00C < \uE04C \uE00D < \uE04D  
 \uE00E < \uE04E \uE00F < \uE04F \uE010 < \uE050 \uE011 < \uE051 \uE012 <  
 \uE052 \uE013 < \uE053 \uE014 < \uE054 \uE015 < \uE055 \uE016 < \uE056 \uE017  
 < \uE057 \uE018 < \uE058 \uE019 < \uE059 \uE01A < \uE05A  
 \uE020 < \uE021 < \uE022 < \uE023 < \uE024 < \uE025 < \uE026 < \uE027 < \uE028  
 < \uE029

note : U+E040 to U+E049 is for Numerals

## 9 Acknowledgements

The development of this proposal and the associated digital infrastructure for the Zou script and numerals would not have been possible without the invaluable support, insights, and collaboration of many individuals and organizations.

We wish to express our sincere gratitude to the native speakers, academic experts, and members of the Zou-speaking community who provided critical linguistic data and verified textual examples. Special thanks are extended to the Zou Literature Society for their foundational publications and documentation, which served as an essential reference baseline.

We are especially grateful to **Dr. Philip Thanglienmang, IAS (Retired)**, for graciously sharing his valuable data and insights, which significantly contributed to this work. Furthermore, we acknowledge the technical contributions of the broader open-source and typography communities, as well as the implementation framework provided by the Keyman and Unicode ecosystems, which ensured the seamless digital integration of the script.

## References

1. Dr. T Marvin Zou. (2026). *Zou Script and Numerals Keyboard Layout* (v1.1) [Digital Keyboard Layout Software]. Keyman Repository, SIL International. Available at: [github.com/keymanapp/keyboards/pull/3996](https://github.com/keymanapp/keyboards/pull/3996).
2. Crystal, D. (2008). *A Dictionary of Linguistics and Phonetics* (6th ed.). Blackwell Publishing.
3. Grierson, G. A. (1904). *Linguistic Survey of India: Vol. III, Tibeto-Burman Family*. Office of the Superintendent of Government Printing, India.
4. Dr. Philip Thanglienmang,. (2012). *Zo Tonology Paper*. Shillong.
5. Thangneu, L. (2012). *A Descriptive Grammar of Zou*. Unpublished Doctoral Dissertation, Assam University, Silchar.
6. Zou Literature Society (ZLS). (2012). *Chinthu-Zaila (Prose & Poetry)* (Prescribed textbook for Class XI & XII, Council of Higher Education, Manipur, 2nd ed.). Churachandpur: Zou Literature Society Publications.

**ISO/IEC JTC 1/SC 2/WG 2  
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS  
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646<sup>1</sup>**

Please fill all the sections A, B and C below.

Please read Principles and Procedures Document (P & P) from <http://std.dkuug.dk/JTC1/SC2/WG2/docs/principles.html> for guidelines and details before filling this form.

Please ensure you are using the latest Form from <http://std.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html>.

See also <http://std.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html> for latest Roadmaps.

**A. Administrative**

1. **Title:** Proposal to encode the Zou script and numerals in the SMP/BMP
2. Requester's name: Dr. T Marvin Zou
3. Requester type (Member body/Liaison/Individual contribution): Individual contribution
4. Submission date: 4 August 2026
5. Requester's reference (if applicable):
6. Choose one of the following:  
This is a complete proposal: ☐ This is a complete proposal  
(or) More information will be provided later: ☐

**B. Technical – General**

1. Choose one of the following:
  - a. This proposal is for a new script (set of characters):  
Proposed name of script: Zou
  - b. The proposal is for addition of character(s) to an existing block:  
Name of the existing block:
2. Number of characters in proposal: 64 characters
3. Proposed category (select one from below - see section 2.2 of P&P document):  
A-Contemporary ☒ B.1-Specialized (small collection) ☐ B.2-Specialized (large collection) ☐  
C-Major extinct ☐ D-Attested extinct ☐ E-Minor extinct ☐  
F-Archaic Hieroglyphic or Ideographic ☐ G-Obscure or questionable usage symbols ☐
4. Is a repertoire including character names provided? ☒  
a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document? ☒  
b. Are the character shapes attached in a legible form suitable for review? ☒
5. Fonts related:  
a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?  
Dr. T Marvin Zou  
b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):  
OFL & MIT
6. References:  
a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided? ☒  
b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached? ☒
7. Special encoding issues:  
Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)? ☒

**8. Additional Information:**

Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at <http://www.unicode.org> for such information on other scripts. Also see Unicode Character Database ( <http://www.unicode.org/reports/tr44/> ) and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

<sup>1</sup> Form number: N4502-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05, 2009-11, 2011-03, 2012-01)

### C. Technical - Justification

|                                                                                                                                                                |   |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------|---|
| 1. Has this proposal for addition of character(s) been submitted before?                                                                                       | N |
| If YES explain                                                                                                                                                 |   |
| 2. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?          | N |
| If YES, with whom?                                                                                                                                             |   |
| If YES, available relevant documents:                                                                                                                          |   |
| 3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? | Y |
| Reference: This Document                                                                                                                                       |   |
| 4. The context of use for the proposed characters (type of use; common or rare)                                                                                |   |
| Reference:                                                                                                                                                     |   |
| 5. Are the proposed characters in current use by the user community?                                                                                           | Y |
| If YES, where? Reference: This Document                                                                                                                        |   |
| 6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?                                  | N |
| If YES, is a rationale provided?                                                                                                                               |   |
| If YES, reference:                                                                                                                                             |   |
| 7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?                                                        | Y |
| 8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?                                        | Y |
| If YES, is a rationale for its inclusion provided?                                                                                                             |   |
| If YES, reference: This Document                                                                                                                               |   |
| 9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?               | Y |
| If YES, is a rationale for its inclusion provided?                                                                                                             |   |
| If YES, reference: This Document                                                                                                                               |   |
| 10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to, or could be confused with, an existing character?         | Y |
| If YES, is a rationale for its inclusion provided?                                                                                                             |   |
| If YES, reference: This Document                                                                                                                               |   |
| 11. Does the proposal include use of combining characters and/or use of composite sequences?                                                                   | N |
| If YES, is a rationale for such use provided?                                                                                                                  |   |
| If YES, reference:                                                                                                                                             |   |
| Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?                                                              |   |
| If YES, reference:                                                                                                                                             |   |
| 12. Does the proposal contain characters with any special properties such as control function or similar semantics?                                            | N |
| If YES, describe in detail (include attachment if necessary)                                                                                                   |   |
|                                                                                                                                                                |   |
|                                                                                                                                                                |   |
| 13. Does the proposal contain any Ideographic compatibility characters?                                                                                        | N |
| If YES, are the equivalent corresponding unified ideographic characters identified?                                                                            |   |
| If YES, reference:                                                                                                                                             |   |

(Detach the following pages from the filled summary form before submitting).

## Information accompanying submissions

The process of deciding which characters should be included in the repertoire of the standard by WG 2 depends on the availability of accurate and comprehensive information about any proposed additions. WG 2, at its San Francisco meeting 26, designed a form (template) that will assist the submitters in gathering and providing the relevant information, and will assist WG 2 in making more informed decisions.

This form is part of the [WG 2 Principles and Procedures document](http://std.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html) and is available on line at <http://std.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html>.

A duly completed proposal summary form must accompany each new submission. The latest version of this form must be used in submissions. Such a form will assist WG 2 to better evaluate the proposal, and progress the proposal towards a speedier acceptance and inclusion in the standard. Submitters are also requested to ensure that a proposed character does not already exist in the standard.

Submitters are encouraged to visit the "[Where is my Character](#)" page on the Unicode web site for more information on checking if their proposed character or script is already encoded in the standard, or a similar proposal has already been made by someone else. There are also several electronic discussion lists maintained by the Unicode consortium that one could use to discuss with other experts internationally on various subjects related to the standard. Submitters are also encouraged to familiarize themselves with ISO/IEC TR15285 – Character Glyph Model. The latest version of ISO/IEC 10646 as well as ISO/IEC TR15285 are available on line under 'Freely Available Standards' at: <http://standards.iso.org/ittf/PubliclyAvailableStandards/>.

National bodies should take note of the guidelines in section *C.4 Some guidelines on proposing new material as ballot comments* in the [P&P document](#).

In addition to text extracted from the standard in *Annex L: Character-naming guidelines* in the [P&P document](#), the following definitions from the standard are also referenced in the proposal summary form:

*Clause 4.12 Combining character:*

A member of an identified subset of the coded character set of ISO/IEC 10646 intended for combination with the preceding non-combining graphic character, or with a sequence of combining characters preceded by a non-combining character (see also 4.14).

NOTE – ISO/IEC 10646 specifies several subset collections, which include combining characters.

*Clause 4.14 Composite sequence:*

A sequence of graphic characters consisting of a non-combining character followed by one or more combining characters (see also 4.12).

NOTE 1 – A graphic symbol for a composite sequence generally consists of the combination of the graphic symbols of each character in the sequence.

NOTE 2 – A composite sequence is not a character and therefore is not a member of the repertoire of ISO/IEC 10646.

## Submitter's responsibilities

The national body or liaison organization (or any other organization or an individual) proposing new character(s) or a new script shall provide:

1. Proposed category for the script or character(s), character name(s), and description of usage.
2. Justification for the category and name(s).
3. A representative glyph(s) image on paper:  
If the proposed glyph image is similar to a glyph image of a previously encoded ISO/IEC 10646 character, then additional justification for encoding the new character shall be provided.  
**Note:** Any proposal that suggests that one or more of such variant forms is actually a distinct character requiring separate encoding should provide detailed, printed evidence that there is actual, contrastive use of the variant form(s). It is insufficient for a proposal to claim a requirement to encode as characters in the Standard, glyphic forms which happen to occur in another character encoding that did not follow the Character-Glyph Model that guides the choice of appropriate characters for encoding in ISO/IEC 10646.  
**Note:** WG 2 has resolved in Resolution [M38.12](#) not to add any more Arabic presentation forms to the standard and suggests users to employ appropriate input methods, rendering and font technologies to meet the user requirements.
4. Mappings to accepted sources, for example, other standards, dictionaries, accessible published materials.
5. Font resource:  
A font must be provided to the Editor promptly after the characters have been approved by WG2 for use in printing the charts. If a font is not provided, the Editor cannot include the glyphs in the charts and, as a result, the repertoire of characters corresponding to these glyphs will not be included in draft amendments. Fonts must be in one of the following formats (in preferential order): OpenType, TrueType, Postscript Type 1.  
Fonts submitted must have no license restrictions that prevent embedding into PDF documents. Because of synchronization between ISO/IEC 10646 and The Unicode Standard, any grant of license must cover use in publishing both standards and related documents (see resolution [M45.30](#)). For technical reasons, the editors must be able under that license to freely modify or replace glyph outlines in their copies of the fonts. The submitter should be prepared to provide fonts of suitable quality and license conditions, unless equivalent fonts are already available to the editors.
6. List of all the parties consulted:  
Towards ensuring that all significant stakeholders have been consulted, WG 2 requests that submitters solicit feedback from other groups who use the proposed characters, and list these contacts. Submitters are encouraged to provide the email id-s of the submitters as well as other experts who have been consulted to facilitate any clarification queries.
7. Equivalent glyph images:  
If the submission intends using composite sequences of proposed or existing combining and non-combining characters, a list consisting of each composite sequence and its corresponding glyph image shall be provided to better understand the intended use.
8. Compatibility equivalents:  
If the submission includes compatibility ideographic characters, identify (per resolution [M45.29](#)):
  - the source, which contains two distinct code positions that correspond to a single unified CJK Ideograph character of ISO/IEC 10646
  - the ISO 10646 unified CJK Ideograph
  - the code position in the source for the unified CJK ideograph
  - the code position in the source for the proposed compatibility ideograph
9. Properties that may affect the BiDi processing  
Any BiDirectional algorithm related properties associated with the characters should be spelled out (see UAX#9 - <http://www.unicode.org/reports/tr9/>).
10. Any additional information that will assist in correct understanding of the different characteristics and linguistic processing of the proposed character(s) or script.
11. If any of the proposed characters are suitable as *syntax* characters please take note of guideline D.2.5 (on page 24 and of Unicode Standard Annex 31 at <http://www.unicode.org/reports/tr31/>).
12. If the proposal is for a new script, identify the standardized characters that are commonly used directly in the script, or proposed to be unified with the characters of the script, in particular standardized characters allocated in different blocks. Examples include punctuation marks and combining marks. Such information will assist in assigning properties for characters shared across multiple scripts or in identifying character repertoires needed to support particular languages.
13. If you are aware of already standardized characters that are visually close to any of the proposed characters, you are invited to list them in the proposal. This will assist in the analysis of the script for 'visually confusables', towards providing additional guidance on use of the standard from a security perspective (see UTR#36 - Unicode Security Considerations – at <http://www.unicode.org/reports/tr36/>).

### ArabicShaping.txt

```
1042  
1043 # EOF  
1 ; undefined; U; No_Joining_Group  
+1 -2
```

### IndicPositionalCategory.txt

```
; # undefined undefined
```

### IndicSyllabicCategory.txt

```
; # undefined undefined
```

### LineBreak.txt

```
3736  
3737 # EOF  
1 ; XX # undefined undefined  
+1 -2
```

### NamesList.txt

```
undefined
```