1. Introduction. The Wancho language has some 51,000 speakers and is used chiefly in the southeast of Arunachal Pradesh, as well as in Assam and Nagaland. There are 64 villages in Longding, 4 villages in the Tirap district, 128 villages in Nagaland, over 23 villages in Myanmar, 3 in Bhutan and 4 in Assam, all where the language is spoken. Wancho is a Sino-Tibetan language of the Konyak subfamily of the Sal family spoken in eastern India, parts of Bangladesh, and Myanmar. The script was devised between 2001 and 2012 by Banwang Losu, a teacher at a government middle school in his home village in Longding District, Arunachal Pradesh. The Wancho script is taught today in schools.

2. Processing. Wancho is a simple alphabetic script comprised of letters which represent both consonants and vowels. Diacritical marks are not used, and there are no conjunct characters. Digraphs and trigraphs are used to express various distinctions. A few characters are similar to other characters, distinguished by a diacritical mark, but this dot is not productive; while three of them appear in pairs in alphabetical order (�atforms of both ����RA and ����U, ����AU), the dotted forms of both ����RA and ����U (����TRA and ����UN) are sorted at some distance from them.

3. Orthography. Tones are marked by repeating the final character to indicate low, mid, or high tone:

<table>
<thead>
<tr>
<th>Characters</th>
<th>Pronunciation</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>���� GI gi</td>
<td>‘cane’</td>
<td>�������� GI</td>
</tr>
<tr>
<td>���� SA sa</td>
<td>‘blunt’</td>
<td>�������� SA</td>
</tr>
<tr>
<td>���� SHA śa</td>
<td>‘iron’</td>
<td>�������� SHA</td>
</tr>
<tr>
<td>���� SHA śa</td>
<td>‘iron’</td>
<td>�������� SHA</td>
</tr>
<tr>
<td>���� ZA za</td>
<td>‘dog’</td>
<td>�������� ZA</td>
</tr>
<tr>
<td>���� O o</td>
<td>‘noise’</td>
<td>�������� O</td>
</tr>
<tr>
<td>���� O o</td>
<td>‘noise’</td>
<td>�������� O</td>
</tr>
</tbody>
</table>

The diphthongs /ai/ and /au/ may be short or long; /au/ is spelt somewhat differently when short. These are spelt əiəi and əuəu here, though another transcription might be aïai and aïäu.
There is apparently an inherent vowel when two consonants come together in two-character mono-

sylables (as in the last line below); nasalized vowels, rimes in the velar-nasal ɲ and in -m also participate in

this orthographic practice:

 Guantanamo ɓi rọ ɗo ‘basket’
 ɬa ɗa ɗa ɗa ‘meet’
 ɗa ɓa ɗa ɗa ‘water’
 ɓa ɡa ɗa ɗa ‘bone’
 ɓa ɬa ɗa ɗa ‘we’
 ɗa ɗa ɗa ‘they’
 ɗa ɗa ɗa ‘we’
 ɗa ɗa ɗa ‘he’

4. **Non-alphabetic characters** are used in Wancho. Ordinary generic punctuation marks like COMMA and

FULL STOP and QUESTION MARK are used in Wancho and can be represented by existing UCS characters.

5. **Digits.** U+1E2F0–U+1E2F9 are the decimal digits 0–9. Common punctuation and operators are used

for mathematical operations. See Figure 3.

6. **Ordering.** The ordering of the character set follows that of the code chart. The first four characters are

given in an order which resembles the Aramaic order used by many alphabets (aa ba ca ga da, sort of a

hybrid of Latin and Greek ΑΒΓΔ) but this seems to be incidental and there is no particular similarity to

that or any other order otherwise. Dotted letters are sorted as separate from their undotted counterparts.

7. **Character names.** The character names here use the usual UCS transcription for the script of India:

AA a, ɓ BA b, ɗ CA c, ɗ DA d, ɠ GA g, ɬ YA j, ɭ PHA p, ɬ LA l, ɭ NA n, ɭ PA p, ɭ TA t, ɭ THA θ, ɭ FA f, ɭ SA s,

ɗ SHA f ɗ SHA f, ɗ ɗ DA d ɗ ɗ DA d, ɗ WA w, ɗ VA v, ɗ KA k, ɗ O O, ɗ AU a, ɗ RA r, ɗ MA m, ɗ KHA k, ɗ HA h, ɗ E e, ɗ I i,

ɗ NGA ɗ NGA ɗ NGA, ɗ U u, ɗ LLHA t, ɗ TSA ts, ɗ TRA tr, ɗ ONG en, ɗ ANɡ an, ɗ ING in, ɗ ON â, ɗ EN ê, ɗ AN â, ɗ NYA ja,

ɗ UN un, ɗ YIH ï.

8. **Unicode Character Properties.**

1E2C0:WANCHO LETTER AA;Lo;0;L;;;;;N;;;;;
1E2C1:WANCHO LETTER BA;Lo;0;L;;;;;N;;;;;
1E2C2:WANCHO LETTER CA;Lo;0;L;;;;;N;;;;;
1E2C3:WANCHO LETTER DA;Lo;0;L;;;;;N;;;;;
1E2C4:WANCHO LETTER GA;Lo;0;L;;;;;N;;;;;
1E2C5:WANCHO LETTER YA;Lo;0;L;;;;;N;;;;;
1E2C6:WANCHO LETTER PHA;Lo;0;L;;;;;N;;;;;
1E2C7:WANCHO LETTER LA;Lo;0;L;;;;;N;;;;;
1E2C8:WANCHO LETTER NA;Lo;0;L;;;;;N;;;;;
1E2C9:WANCHO LETTER PA;Lo;0;L;;;;;N;;;;;
1E2CA:WANCHO LETTER TA;Lo;0;L;;;;;N;;;;;
1E2CB:WANCHO LETTER TRA;Lo;0;L;;;;;N;;;;;
1E2CC:WANCHO LETTER FA;Lo;0;L;;;;;N;;;;;
1E2CD:WANCHO LETTER SA;Lo;0;L;;;;;N;;;;;
1E2CE:WANCHO LETTER SHA;Lo;0;L;;;;;N;;;;;
1E2CF:WANCHO LETTER JA;Lo;0;L;;;;;N;;;;;
1E2D0:WANCHO LETTER Za;Lo;0;L;;;;;N;;;;;
1E2D1:WANCHO LETTER WA;Lo;0;L;;;;;N;;;;;
1E2D2:WANCHO LETTER VA;Lo;0;L;;;;;N;;;;;
1E2D3:WANCHO LETTER RA;Lo;0;L;;;;;N;;;;;
1E2D4:WANCHO LETTER O;Lo;0;L;;;;;N;;;;;
1E2D5:WANCHO LETTER AU;Lo;0;L;;;;;N;;;;;
1E2D6:WANCHO LETTER RA;Lo;0;L;;;;;N;;;;;
1E2D7:WANCHO LETTER NA;Lo;0;L;;;;;N;;;;;
1E2D8:WANCHO LETTER KHA;Lo;0;L;;;;;N;;;;;
1E2D9:WANCHO LETTER HA;Lo;0;L;;;;;N;;;;;
1E2DA:WANCHO LETTER E;Lo;0;L;;;;;N;;;;;
1E2DB:WANCHO LETTER I;Lo;0;L;;;;;N;;;;;
1E2DC:WANCHO LETTER NGA;Lo;0;L;;;;;N;;;;;
1E2DD:WANCHO LETTER U;Lo;0;L;;;;;N;;;;;
1E2DE:WANCHO LETTER LHA;Lo;0;L;;;;;N;;;;;
1E2DF:WANCHO LETTER TSA;Lo;0;L;;;;;N;;;;;
1E2E0:WANCHO LETTER TRA;Lo;0;L;;;;;N;;;;;
1E2E1:WANCHO LETTER ONG;Lo;0;L;;;;;N;;;;;
1E2E2:WANCHO LETTER ANG;Lo;0;L;;;;;N;;;;;
1E2E3:WANCHO LETTER ING;Lo;0;L;;;;;N;;;;;
1E2E4:WANCHO LETTER ON;Lo;0;L;;;;;N;;;;;
1E2E5:WANCHO LETTER EN;Lo;0;L;;;;;N;;;;;
9. References

10. Acknowledgements. This project was made possible in part by a grant from the U.S. National Endowment for the Humanities, which funded the Universal Scripts Project (part of the Script Encoding Initiative at UC Berkeley) in respect of the Wancho encoding. Any views, findings, conclusions or recommendations expressed in this publication do not necessarily reflect those of the National Endowment for the Humanities.

Figures.

![Wancho Alphabet Chart](image)

**Figure 1.** Chart of the Wancho alphabet. The gap appears between the letters which are used for sounds common to other languages of India, and sounds which are considered unique to Wancho. From Losu 2013:31.
# 11 SIMILARITIES

(Similar in sound, different in meaning)

<table>
<thead>
<tr>
<th>Short Vowel Sounds</th>
<th>Long Vowel Sounds</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Wancho</strong></td>
<td><strong>English</strong></td>
</tr>
<tr>
<td>££</td>
<td>Ba</td>
</tr>
<tr>
<td>££</td>
<td>Ban</td>
</tr>
<tr>
<td>££</td>
<td>Bea</td>
</tr>
<tr>
<td>££</td>
<td>Bu</td>
</tr>
<tr>
<td>££</td>
<td>Bo*</td>
</tr>
<tr>
<td>££</td>
<td>Cha</td>
</tr>
<tr>
<td>££</td>
<td>Cho</td>
</tr>
<tr>
<td>££</td>
<td>Chm</td>
</tr>
<tr>
<td>££</td>
<td>Chi</td>
</tr>
<tr>
<td>££</td>
<td>Da</td>
</tr>
<tr>
<td>££</td>
<td>Dam</td>
</tr>
<tr>
<td>££</td>
<td>Di</td>
</tr>
<tr>
<td>££</td>
<td>Ga</td>
</tr>
<tr>
<td>££</td>
<td>Gan</td>
</tr>
<tr>
<td>££</td>
<td>Gi</td>
</tr>
<tr>
<td>££</td>
<td>Gong</td>
</tr>
<tr>
<td>££</td>
<td>Go*</td>
</tr>
</tbody>
</table>

* symbol shows incomplete pronunciation
Glotal stops, IPA symbol ː is used in English and Colon jj is used in Hindi

<table>
<thead>
<tr>
<th>Wancho</th>
<th>English</th>
<th>Wancho</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>£££</td>
<td>No*</td>
<td>Basket</td>
<td>£££</td>
</tr>
<tr>
<td>£££</td>
<td>Ni</td>
<td>Smile</td>
<td>£££</td>
</tr>
<tr>
<td>£££</td>
<td>Pan</td>
<td>Tree</td>
<td>£££</td>
</tr>
<tr>
<td>£££</td>
<td>Pi</td>
<td>Fly</td>
<td>£££</td>
</tr>
<tr>
<td>£££</td>
<td>Tu</td>
<td>Not sharp</td>
<td>£££</td>
</tr>
<tr>
<td>£££</td>
<td>Than</td>
<td>Face</td>
<td>£££</td>
</tr>
<tr>
<td>£££</td>
<td>Ki</td>
<td>Quarrel</td>
<td>££££</td>
</tr>
<tr>
<td>£££</td>
<td>O</td>
<td>Noise</td>
<td>£</td>
</tr>
<tr>
<td>£££</td>
<td>Man</td>
<td>No</td>
<td>££££</td>
</tr>
<tr>
<td>££££</td>
<td>Khan</td>
<td>Yam</td>
<td>££££</td>
</tr>
<tr>
<td>££££</td>
<td>Mai</td>
<td>Good</td>
<td>£££££</td>
</tr>
<tr>
<td>££££</td>
<td>Ja*</td>
<td>Water</td>
<td>£££££</td>
</tr>
<tr>
<td>££££</td>
<td>Nga</td>
<td>Tare</td>
<td>£££££</td>
</tr>
<tr>
<td>££££</td>
<td>Khi</td>
<td>Red</td>
<td>£££££</td>
</tr>
<tr>
<td>££££</td>
<td>Kan</td>
<td>To carry</td>
<td>£££££</td>
</tr>
<tr>
<td>£££££</td>
<td>Kom</td>
<td>They</td>
<td>££££££</td>
</tr>
<tr>
<td>££££</td>
<td>Ju</td>
<td>Beg</td>
<td>£££££</td>
</tr>
<tr>
<td>£££££</td>
<td>Pong</td>
<td>Meet</td>
<td>£££££</td>
</tr>
<tr>
<td>£££££</td>
<td>Lam</td>
<td>Way</td>
<td>£££££</td>
</tr>
<tr>
<td>£££££</td>
<td>Shan</td>
<td>Face</td>
<td>£££££</td>
</tr>
<tr>
<td>£££££</td>
<td>Chh</td>
<td>Bite</td>
<td>£££££</td>
</tr>
</tbody>
</table>

**Figure 2.** Some minimal pairs in Wancho. From Losu 2013:86–87.
Figure 3. Examples of arithmetical operations in Wancho. From Losu 2013:96–97.

Figure 4. Example of the Wancho ngun sign, used to indicate rupees. From Losu 2013:119.
Figure 5. Examples of words beginning the alphabet in Wancho script, Latin script, Devanagari script, with English and Hindi translations. Variant readings like घन n chahnu [tʃaʔnù] / घन thahnu [θaʔnù] indicate dialect variation. From Losu 2013:44–45.
Figure 6. A Wancho poem. From Losu 2013:124.
Figure 7. A Wancho song with transliteration. From Losu 2013:125.
Figure 8. A page from one of the Middle School examinations in Wancho.
Figure 9. Students taking an examination in Wancho script.
### Letters
1E20 -zA\U1E20  WANCHO LETTER AA
1E21 -zA\U1E21  WANCHO LETTER BA
1E22 -zA\U1E22  WANCHO LETTER CA
1E23 -zA\U1E23  WANCHO LETTER DA
1E24 -zA\U1E24  WANCHO LETTER GA
1E25 -zA\U1E25  WANCHO LETTER YA
1E26 -zA\U1E26  WANCHO LETTER PHA
1E27 -zA\U1E27  WANCHO LETTER LA
1E28 -zA\U1E28  WANCHO LETTER NA
1E29 -zA\U1E29  WANCHO LETTER PA
1E2A -zA\U1E2A  WANCHO LETTER TA
1E2B -zA\U1E2B  WANCHO LETTER THA
1E2C -zA\U1E2C  WANCHO LETTER FA
1E2D -zA\U1E2D  WANCHO LETTER SA
1E2E -zA\U1E2E  WANCHO LETTER SHA
1E2F -zA\U1E2F  WANCHO LETTER JA
1E30 -zA\U1E30  WANCHO LETTER ZA
1E31 -zA\U1E31  WANCHO LETTER WA
1E32 -zA\U1E32  WANCHO LETTER VA
1E33 -zA\U1E33  WANCHO LETTER KA
1E34 -zA\U1E34  WANCHO LETTER O
1E35 -zA\U1E35  WANCHO LETTER AU
1E36 -zA\U1E36  WANCHO LETTER RA
1E37 -zA\U1E37  WANCHO LETTER MA
1E38 -zA\U1E38  WANCHO LETTER KHA
1E39 -zA\U1E39  WANCHO LETTER HA
1E3A -zA\U1E3A  WANCHO LETTER E
1E3B -zA\U1E3B  WANCHO LETTER I
1E3C -zA\U1E3C  WANCHO LETTER NGA
1E3D -zA\U1E3D  WANCHO LETTER U
1E3E -zA\U1E3E  WANCHO LETTER LLHA
1E3F -zA\U1E3F  WANCHO LETTER TSA
1E40 -zA\U1E40  WANCHO LETTER TRA
1E41 -zA\U1E41  WANCHO LETTER ONG
1E42 -zA\U1E42  WANCHO LETTER ANG
1E43 -zA\U1E43  WANCHO LETTER ING
1E44 -zA\U1E44  WANCHO LETTER ON
1E45 -zA\U1E45  WANCHO LETTER EN
1E46 -zA\U1E46  WANCHO LETTER AN
1E47 -zA\U1E47  WANCHO LETTER NYA
1E48 -zA\U1E48  WANCHO LETTER UN
1E49 -zA\U1E49  WANCHO LETTER YIH

### Digits
1E2F0 -zA\U1E2F0  WANCHO DIGIT ZERO
1E2F1 -zA\U1E2F1  WANCHO DIGIT ONE
1E2F2 -zA\U1E2F2  WANCHO DIGIT TWO
1E2F3 -zA\U1E2F3  WANCHO DIGIT THREE
1E2F4 -zA\U1E2F4  WANCHO DIGIT FOUR
1E2F5 -zA\U1E2F5  WANCHO DIGIT FIVE
1E2F6 -zA\U1E2F6  WANCHO DIGIT SIX
1E2F7 -zA\U1E2F7  WANCHO DIGIT SEVEN
1E2F8 -zA\U1E2F8  WANCHO DIGIT EIGHT
1E2F9 -zA\U1E2F9  WANCHO DIGIT NINE

### Symbol
1E2FA -zA\U1E2FA  WANCHO NGUN SIGN

\(\Rightarrow 20B9 ₹\)  \(\text{indian rupee sign}\)
A. Administrative
1. Title
Proposal to encode the Wancho script in the UCS
2. Requester’s name
UC Berkeley Script Encoding Initiative (Universal Scripts Project)
3. Requester type (Member body/Liaison/Individual contribution)
Liaison contribution.
4. Submission date
2016-03-09
5. Requester’s reference (if applicable)
6. Choose one of the following:
6a. This is a complete proposal
Yes.
6b. More information will be provided later
No.

B. Technical – General
1. Choose one of the following:
1a. This proposal is for a new script (set of characters)
Yes.
1b. Proposed name of script
Wancho.
1c. The proposal is for addition of character(s) to an existing block
No.
1d. Name of the existing block
2. Number of characters in proposal
53.
3. Proposed category (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)
Category A.
4a. Is a repertoire including character names provided?
Yes.
4b. If YES, are the names in accordance with the “character naming guidelines” in Annex L of P&P document?
Yes.
4c. Are the character shapes attached in a legible form suitable for review?
Yes.
5a. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?
Michael Everson via Banwang Losu
5b. If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:
Michael Everson, Fontographer.
6a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?
Yes.
6b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?
Yes.
7. 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)?
Yes.
8. 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.
See above.

C. Technical – Justification
1. Has this proposal for addition of character(s) been submitted before? If YES, explain.
Yes. N-4785.
2a. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?
Yes.
2b. If YES, with whom?
Banwang Losu, Stephen Morey
2c. 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?
Contemporary use by Wancho speakers.
4a. The context of use for the proposed characters (type of use; common or rare)
Common.
4b. Reference
5a. Are the proposed characters in current use by the user community?
Yes.
5b. If YES, where?
In Arunachal Pradesh.

6a. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?
No.
6b. If YES, is a rationale provided?
6c. If YES, reference

7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?
Yes.
8a. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?
No.
8b. If YES, is a rationale for its inclusion provided?
8c. If YES, reference

9a. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?
No.
9b. If YES, is a rationale for its inclusion provided?
9c. If YES, reference

10a. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?
No.
10b. If YES, is a rationale for its inclusion provided?
10c. If YES, reference

11a. Does the proposal include use of combining characters and/or use of composite sequences (see clauses 4.12 and 4.14 in ISO/IEC 10646-1: 2000)?
No.
11b. If YES, is a rationale for such use provided?
11c. If YES, reference

12a. Does the proposal contain characters with any special properties such as control function or similar semantics?
No.
12b. If YES, describe in detail (include attachment if necessary)

13a. Does the proposal contain any Ideographic compatibility character(s)?
No.
13b. If YES, is the equivalent corresponding unified ideographic character(s) identified?