



# Negotiating Neographies in Unicode

Approaches for Encoding Newly-Invented Scripts

Anshuman Pandey

Deborah Anderson

Script Encoding Initiative, UC Berkeley

𐌲𐌿𐌽𐌰

𐌲𐌿𐌽𐌰

𐌲𐌿𐌽𐌰

𐌲𐌿𐌽𐌰

𐌲𐌿𐌽𐌰



What is a  
“neography”?

## Neography:

- = Newly invented script: neo + orthography
- In this talk, refers to a script created in the 20<sup>th</sup> - 21<sup>st</sup> centuries
- Most neographies are from Africa or South Asia



# Script Encoding Initiative

Project working to help communities get eligible characters and scripts into the Unicode Standard

Department of Linguistics  
University of California, Berkeley

- Scripts to Encode
- Progress Overview

- News & Presentations
- Press

- How to Donate
- Donors

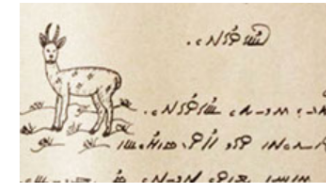
## WELCOME TO THE SCRIPT ENCODING INITIATIVE

The Script Encoding Initiative (SEI), established in the [UC Berkeley Department of Linguistics](#) in April 2002, is a project devoted to the preparation of formal proposals for the encoding of scripts and script elements not yet currently supported in Unicode (ISO/IEC 10646).

[Unicode](#) is the universal computing standard specifying the representation of text in all modern software. To date, Unicode has largely focused on the major modern scripts, particularly those scripts most widely used in business. Some minority and historic scripts have already been encoded, as well as historic characters of the major modern scripts.



The goal of the SEI project is to fund the preparation of script proposals that will be successfully approved by the Unicode Technical Committee and WG2 (ISO/IEC 10646) without requiring extensive revision or involvement of the committee itself.



A secondary goal to encourage the creation of freely-available Unicode-conformant fonts. This will help to promote widespread adoption and implementation of the scripts.

By providing funding for proposal authors, drawn from faculty and graduate students as well as other experts, the Script Encoding Initiative represents a concerted effort to tackle the remaining scripts and remaining script issues. The project will be assisted by a Unicode Vice President to assure that the proposals meet requirements of the Unicode Technical Committee and of the international standards community. To date, the project has helped get over [70 scripts encoded](#).

The Script Encoding Initiative project is of world-wide importance, for minority and historic scripts. For a minority language, having its script included in the universal character set will help to

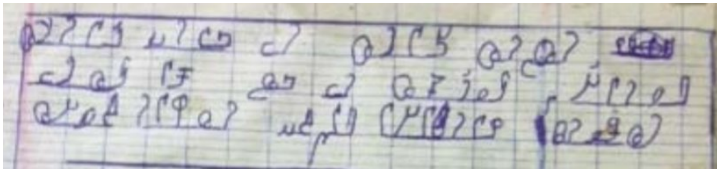
<https://linguistics.berkeley.edu/sei>



# Neographies (New Scripts) in Unicode

General impression of process

New script



published in Unicode



Facebook

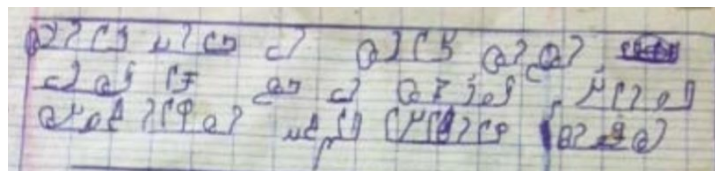


# Neographies (New Scripts) in Unicode

More accurate view of process

[Long proposal review and approval process]

New script



published in Unicode

[need CLDR data, fonts, keyboards, ...]



Facebook





# Script Ad Hoc and Neographies

## Script Ad Hoc

- meets monthly to review new script proposals.
- makes recommendations to the Unicode Technical Committee



- Note: Some new script requests come through other means: Unicode Feedback form or to Script Encoding Initiative



# Script Ad Hoc Questions (for Neography Proposal Authors)

## Users

- Number of people are using the script today? How many are using it daily to write the language?





# Script Ad Hoc Questions (for Neography Proposal Authors)

## Publications

- Are there publications not produced by the script's creator?
- Is it used for newspapers, pamphlets, and in books? How many books in the script have been produced?

First Book in Nsibidi







# Script Ad Hoc Questions (for Neography Proposal Authors)

## Schools / Usage

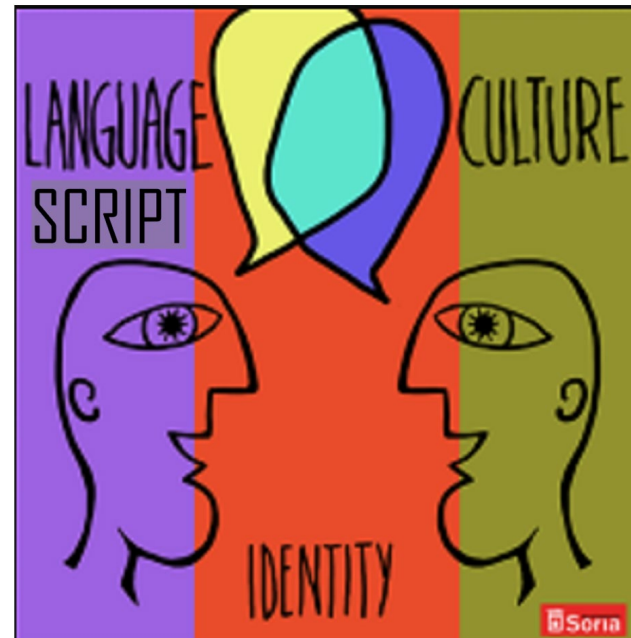
- Number of people learning the script in schools today?  
Any graduates of the school yet and, if so, how many?
- Evidence script's use is growing?





## Reasons for creating new scripts (general)

- Can preserve a group's unique linguistic identity and cultural heritage, differentiating one group from others
- Considered a unique cultural contribution, often viewed with pride, since most scripts are created by a member of the community





# African Neographies

# Reasons for creating new scripts (African scripts)

- Way to write indigenous languages of Africa, not tied to Africa's history of colonization



Map of Africa, 1910  
(image: Wikipedia)





## Reasons for creating new scripts (African scripts)

- Often considered by users to be better suited to represent the sounds of their language (over the Latin or Arabic scripts)
- Viewed as a way to improve literacy; not dependent upon knowing French, English (etc.)





## Reasons for creating new scripts (African scripts)

- In some cases, the script's creator may have a vision for the script to be extended to write all (or many of) languages of Africa



# Features of African neographies

- Most African neographies are created to write languages in West Africa; a few scripts are connected with religious groups

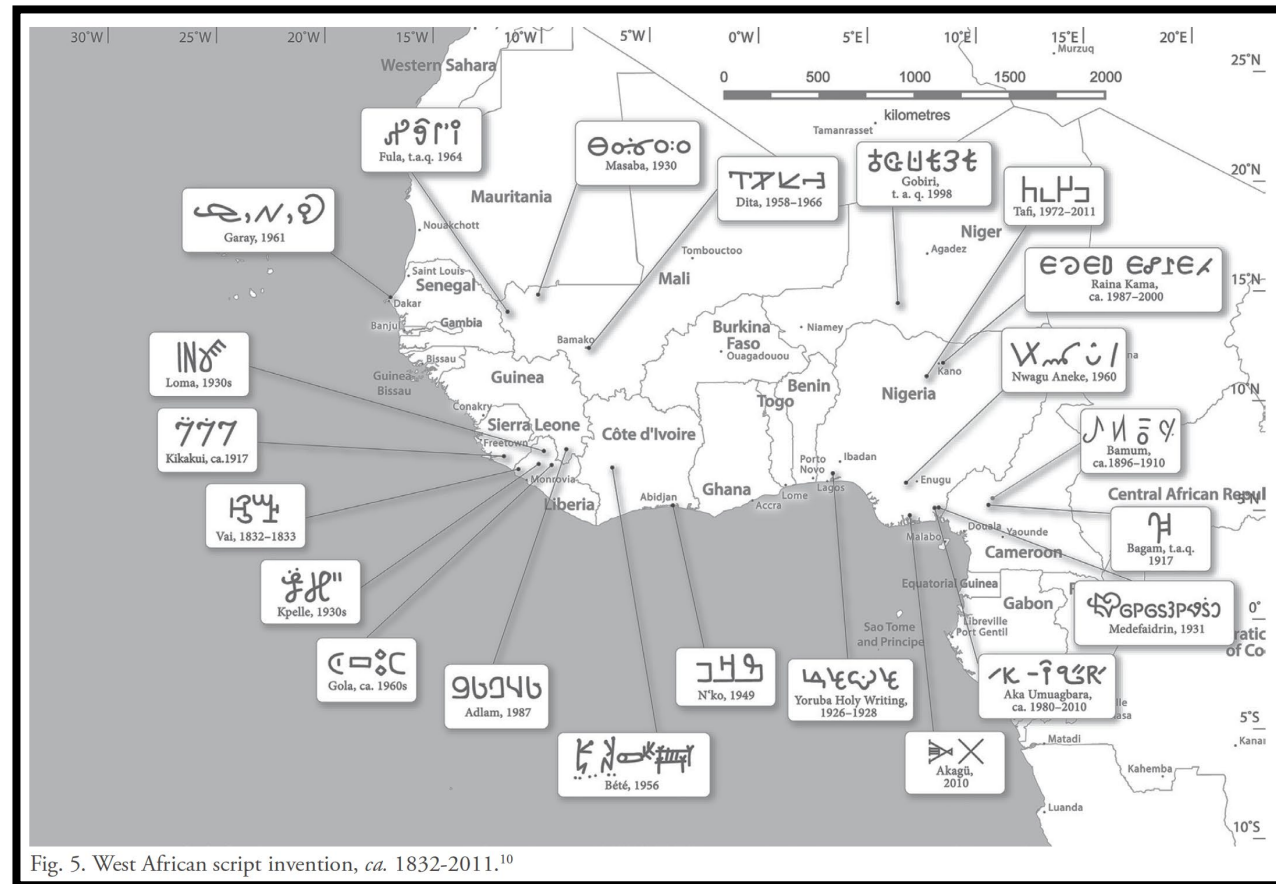
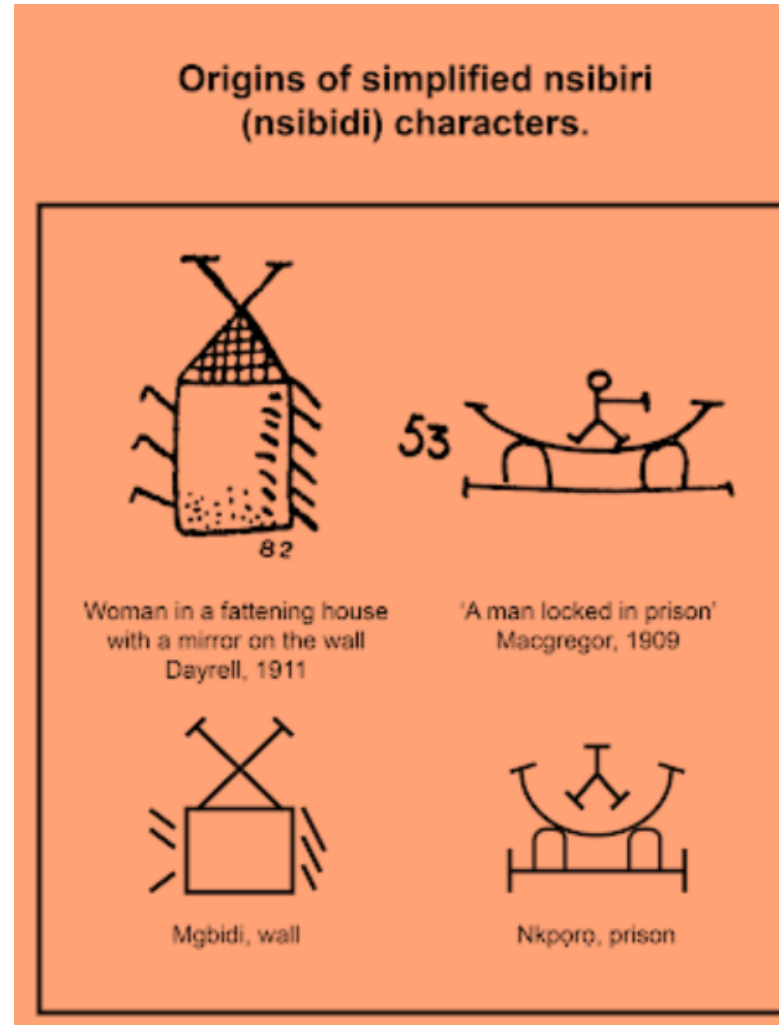


Fig. 5. West African script invention, ca. 1832-2011.<sup>10</sup>



# Features of African neographies

- Some scripts (and their glyphs) may be tied to historical pictographs or objects associated with the group



# Features of African neographies

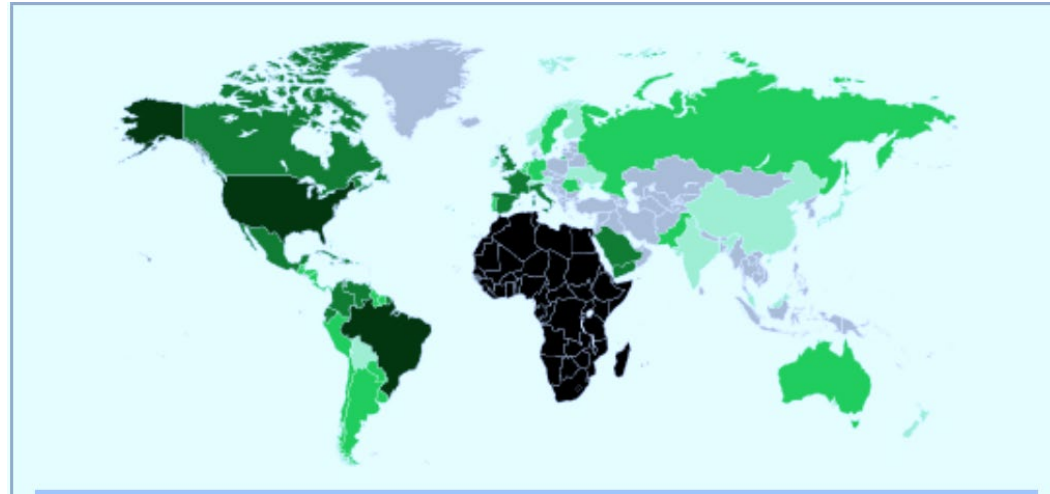
- Several recent neographies have adopted technology:
  - YouTube or Facebook videos online (often with tutorials); Facebook groups



Nsibidi for Igbo: Introduction  
<https://www.youtube.com/watch?v=83DQD-GbDWo>

# Features of African neographies

- Online presence shows:
  - knowledge of Unicode (based on ADLaM and N'Ko?)
  - understanding of the importance of digitization for survival of the script (and its role in preserving the language)
  - interest in these scripts amongst those in the diaspora



Map of African diaspora (Image: Wikipedia)



# African Neogeographies: Three Case Studies



# I. Case Studies – Ditema Tsa Dinoko (=Isibheqe Sohlamvu)

Key traits:

- Created in 2010s by linguists, designers, software programmers
- Used for southern Bantu languages in South Africa and neighboring countries
- Inspired by Sesotho mural art / Zulu beadwork
- No Unicode proposal



<https://isibheqe.org.za/>

# I. Case Studies – Ditema Tsa Dinoko (=Isibheqe Sohlamvu)

7 | WN

DIKAKANYO/TSELE LE TSELE

April 2016

## ISIGODLO SEZIFUNDISWA: LUGCINWA ULWAZI EMAGCEKENI AKWANGQONDONKULU

**THEKWANE NDABA**  
*isiZulu*

asilelele isexwayiso salo. KwaNgqondonkulu emagcekeni onke thina abangaphakathi senziwa ukuba sikholwe ukuthi umsebenzi wethu mihle. Iqiniso lisababa nxa kubonwe isimo sezifundiswa ezahlukeniswa nemindeni yazo. Kule ndaba kuzobekwa phansi izindlela ezintathu ulwazi lakwaNgqondonkulu lwebiwa ngazo. Okokuqala kungukusetshenziswa kwesiNgisi nakwesiBhunu emagcekeni emfundo ephakeme. Okwesibili kuyizinhlobo ezimbini zocwaningo eziceliwe ngabaphathi bezigodlo zifundiswa. Okwesithathu kuyizinhlobo zokucindezela zalo cwaningo ezifihliwe.

Ukwesithathu kuyizinhlobo zokucindezela abanengi. Lokhu kusenzeka emagcekeni ethu esimanje. Ikakhulu abantu ezweni lethu bakhuluma izilimi zabo njengesibhaca, isiLala, Sepedi, Xitsonga, isiXhosa, Setswana, Sesotho, Tshivenda, isiBhunu, isiCamtho, SiSwati, isiNdebele, isiZulu nezinye eziningi, kodwa ke sisaduma isiNgisi nesiBhunu emnithweni, komabonwakude, ekuqashweni nangaphakathi kwizombusazwe. Isilungu sesithathile zonke zezikhundla zokuzwakala kuleli zwe.

Yingakho ke kusondela kithi inkinga la. Njengabafundi asifuni neze ukuzilondolozela lo lwazi esilutholayo khona. Cha, kufuze silufake ezandleni zabakithi ekhaya balunwampule. Kodwa pho kunganikezwa njani lo lwazi nxa lusesilungwini? Kuzobachazelwa njani nxa abazali nabomndeni wethu bangakazwile neze la magama athize enhlobo yesiNgisi esibondayo? "Intersectionality" kungathiwa kanjani ngesiZulu? Ne- "Decolonial project",

"Transgender" ne- "Modernism" futhi? Akungukuthi ingekho leyo miqondo ezilimini zethu, okuyinkinga kodwa kungukuswela kwemibhalo ezilimini zethu ezingasimela njengamaseko omilo wengqondo. Ngakho ke kuzomane sicwaninge sicabange ngezilimi zethu ukuze ulwazi lwethu lungandiza emiphemulweni kuze lufike emiphakathini yonke.

Ziningi izinhlobo zalo lwazi olukwazi ukufuna ukuphapha 'maphiko. Lo lwazi nalo cwaningo olwenziwa yizifundiswa luyinhlobonhlobo njengempuphu ebigaywa kamnandi. Ezinye zalezo zinhlobo zibalulekile zimi phezu nokho zisakhona ezinye zazo ezingudaka ngenxa yokusetshenziselwa kwazo ngobukoloni. Esigabeni esizayo kuzokhonjiswa isibonelo salokho. Okunye kungucwaningo lolo olungakwazi ukusiza abantu bomhlaba ekukhuphukeni kwabo, njengokuchaza izitho zezilwanyane ezithi "thorax", "abdomen" nezinye njalo. Angithi sidinga ukuluchitha lo lwazi kodwa ngenkathi iyakhula indlela yabakithi pho ngabe sishayane ngenduku ngezitho zezilwanyane? Sinemithwalo yabazali nayokhokho bethu esidinge siyithwale.

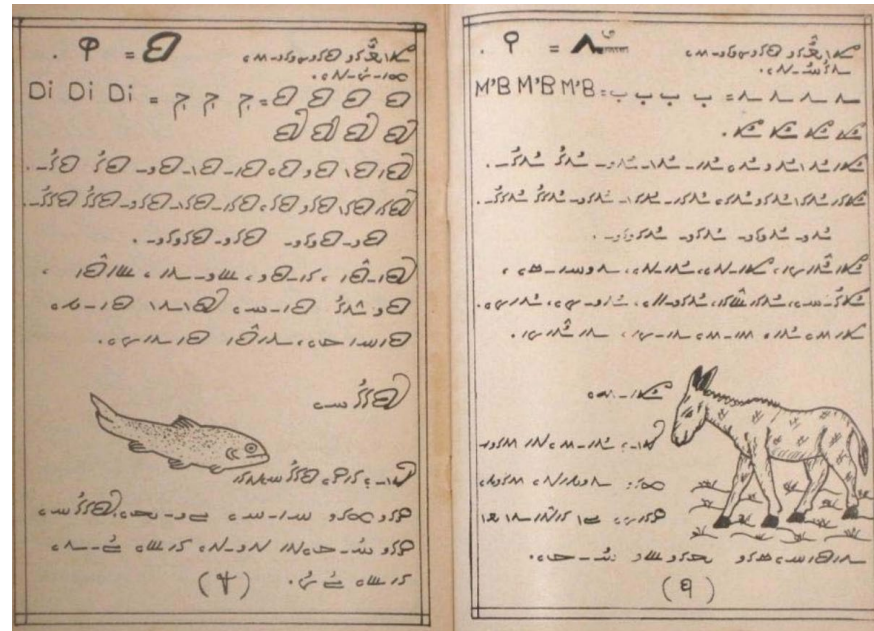
Njengoba ngishilo ngenhla, kade kwaqanjwe ucwaningo

olwalubasiza abelungu bohulumeni wobandlululo ukuba baqhube nokwazi ukubacindezela abantu kanjani. Esinye isibonelo salona silona luka- A.T. Bryant olwalucwaninga ngombuso wesifundazwe saKwaZulu. Lona lwabonwa yizikhulu zaseKoloni yase- 'Natali'. Abelungu bona babefuna ukwazi ukuthi umbuso wakhiwe njani KwaZulu ukuze bethole indlela ethile ukhulasele ngayo abantu bomdabu. Lolu cwaningo lungaphakathi kwenhlobo yolwazi ethiwa ngesilungu i- 'Anthropology', kumbe ngesiZulu ethi ukucwaninga mayelana namasiko, umnotho, ubukhosi, umbuso, inkolo nezinye izinto zonke ezenziwa ngabantu bonke bomhlaba. Luyingozi lo cwaningo ngoba lusaceliwa yizinkampani ezifuna ukuqhamuka ezweni labantu.

Ngakho ke kungathiwa ukuthi amagceka akwaNgqondonkulu ayizigodlo zezifundiswa ngoba lo lwazi lwazo lungavunyelwe ukuphuma kuleso sigodlo ukuba luhambe kuze lufike ezandleni zabakithi abangenayo imithuba yokufunda. Ngaphezu kwalokho kwaNgqondonkulu kufuneka ucwaningo oluyokhombisa izindlela ezintsha ukweba umhlaba wabantu ngazo.

Those of us who are within the walls of the university often forget or do not

## II. Case Studies – Garay



### Key traits:

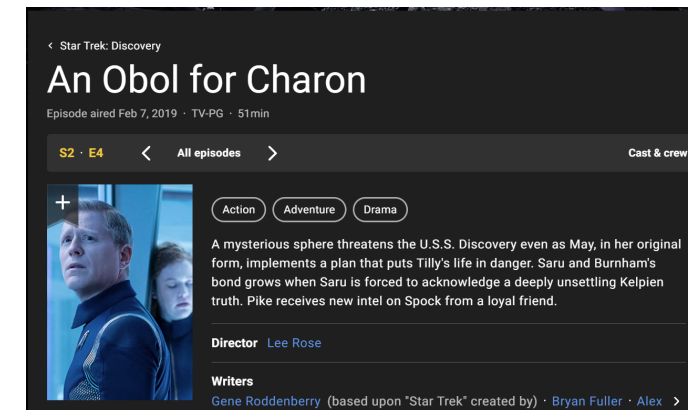
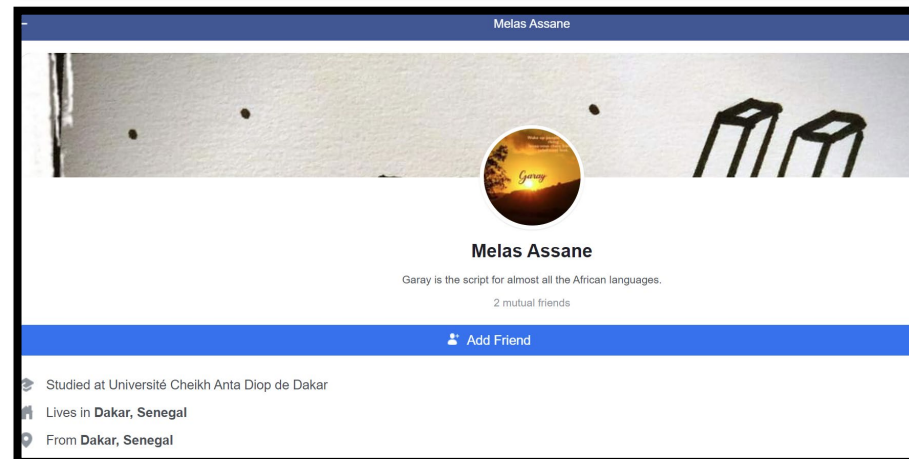
- Created in 1961 by Assane Faye
- Used for Wolof in Senegal and environs
- Some written materials, including Koran
- Taught informally
- Proposal still under review

Photo: <https://www.unicode.org/L2/L2017/17322-garay-progress.pdf>

Text image from proposal: <http://www.unicode.org/L2/L2016/16069-n4709-garay-revision.pdf>



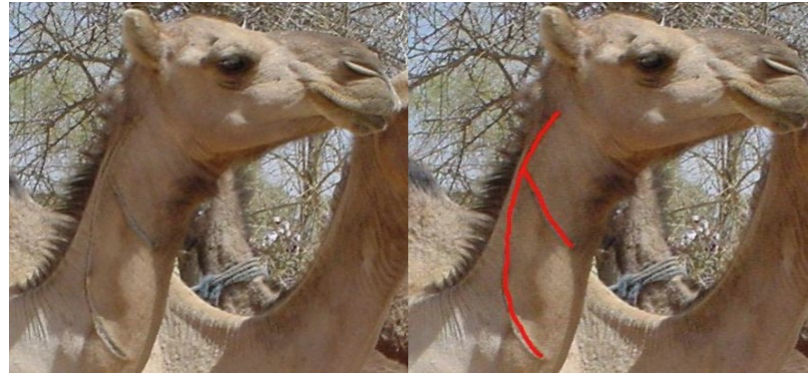
## II. Case Studies – Garay



### III. Case Studies – Beria

Key traits:

- Created in 1950s by Adam Tajir for writing the Zaghawa language of Sudan (Darfur) and Chad; in 2000, Siddick Adam Issa modified the script
- Glyphs derive from camel brands
- Unicode preliminary proposal
- Groups use script differently; need community to reach consensus

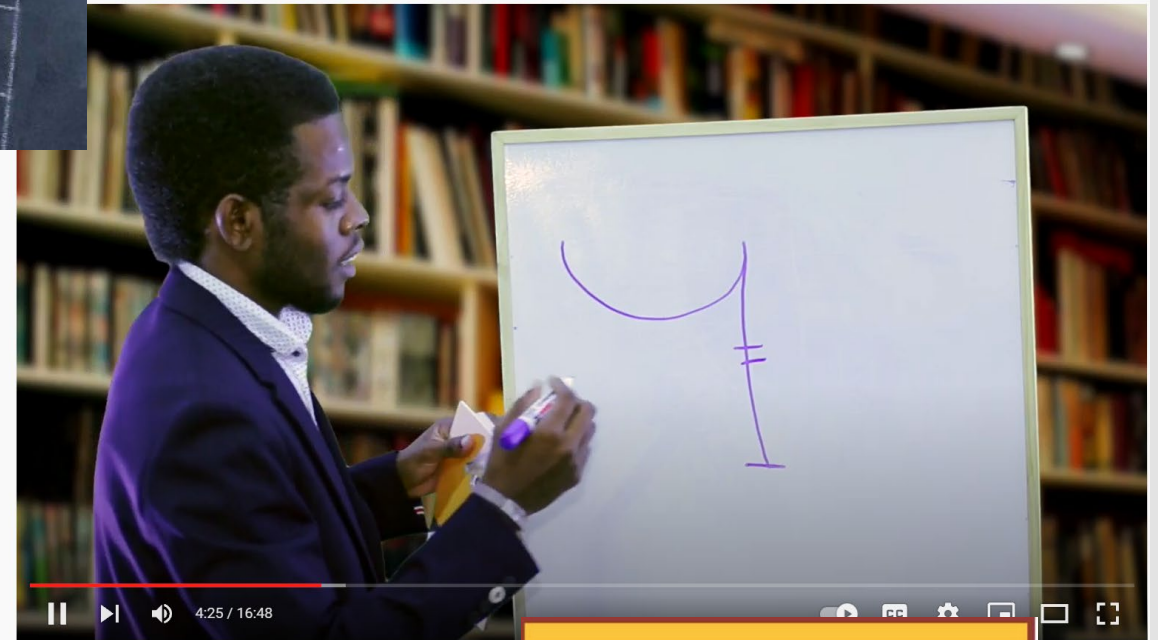




Siddick Adam Issa

### III. Case Studies – Beria

Alsadik Sadik (Israel)



الدرس الثاني في لغة البرياء The second lesson in beria language

(Turkey)





# South Asian Neographies

# South Asian Neographies in Unicode

## In Unicode

- Ol Chiki
- Warang Citi
- Sora Sompeng
- Mro
- Pau Cin Hau alphabet
- Masaram Gondi
- Toto
- Wancho
- Hanifi Rohingya

## Approved for inclusion

- Tangsa
- Nag Mundari

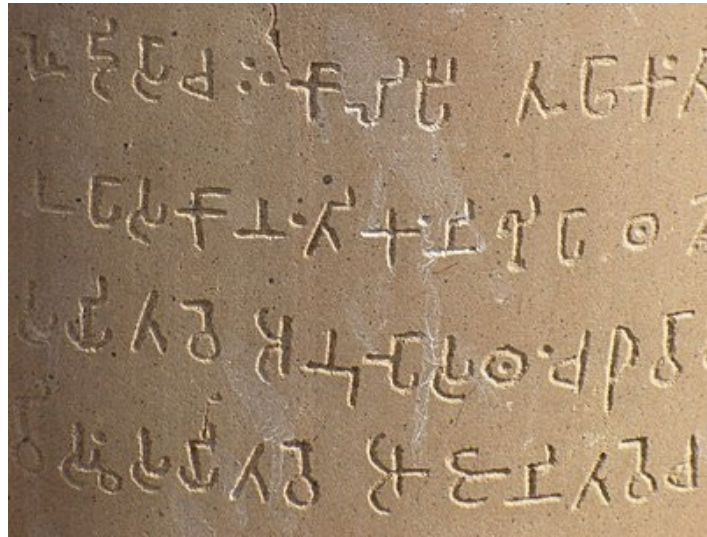
## Pending

- Sunuwar
- Tolong Siki
- Gurung Khema



# Neographies in South Asia - Historical context

- Script development in South Asia
  - continuous phenomenon since at least 5<sup>th</sup> century BCE (... or perhaps 24<sup>th</sup> century BCE ?)



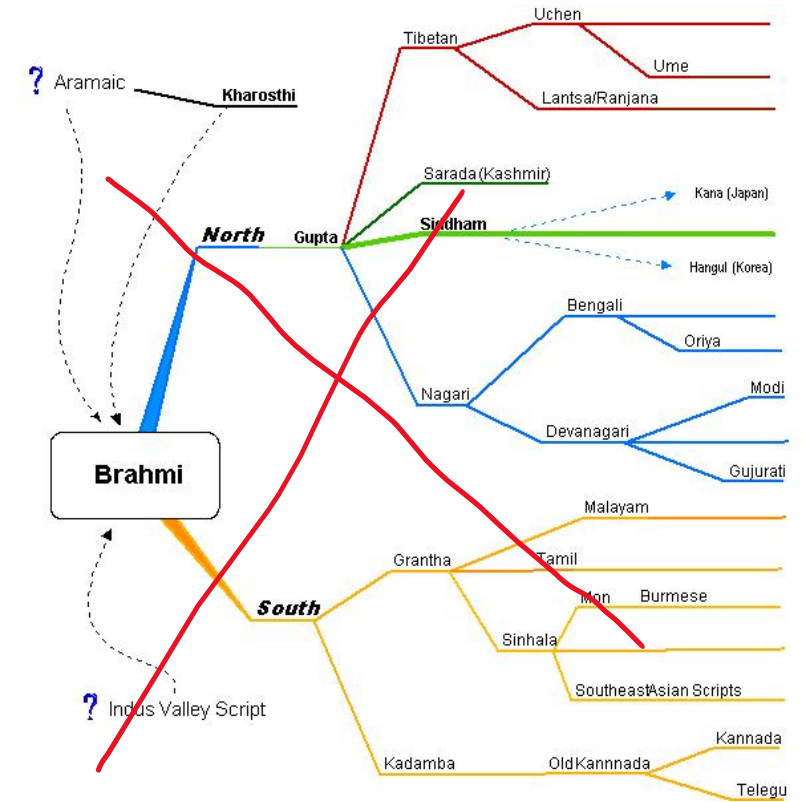
Excerpt of edict of Ashoka in  
Brahmi on a pillar at Girnar,  
2<sup>nd</sup> century BCE



'Unicorn' seal from Harappa,  
24<sup>th</sup> to 2250 BCE

# Neographies in South Asia - Definitions

- South Asian 'neography':
  - Genetically unrelated to other scripts
    - Not organically derived from Brahmi, Latin, etc.
  - Indigenous creation

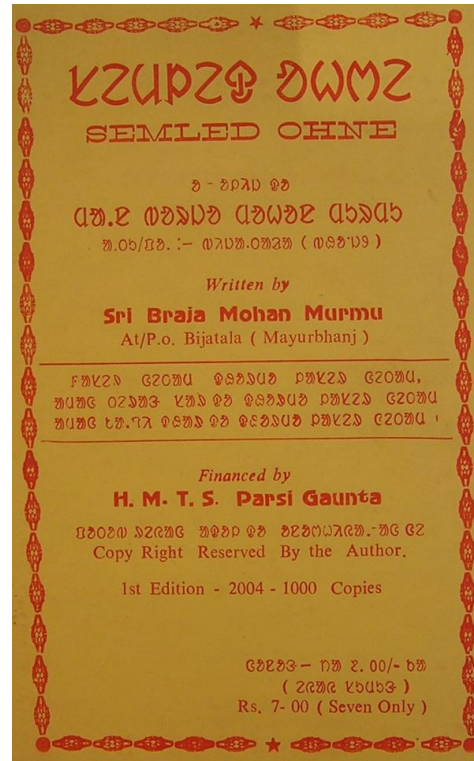




## Motivations for script creation

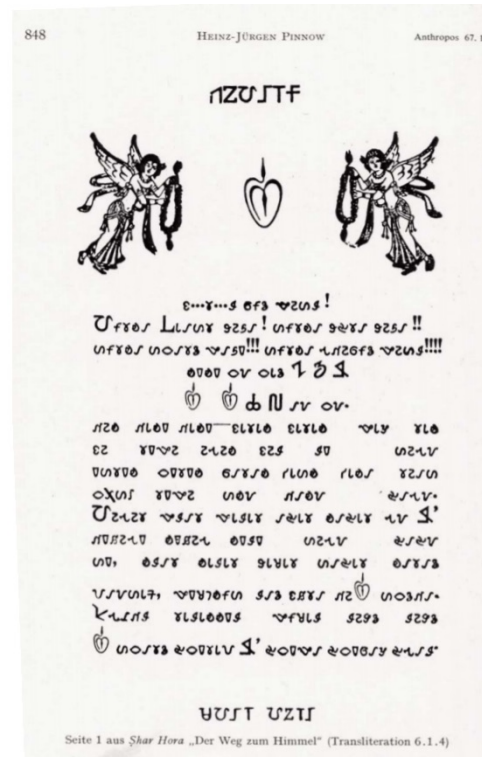
- Distinctive visual marker for language and community
- Parity with other written languages
- Separation from other languages
- Development of an orthography that is more suited to phonology of language

# 'Classic' South Asian Neographies



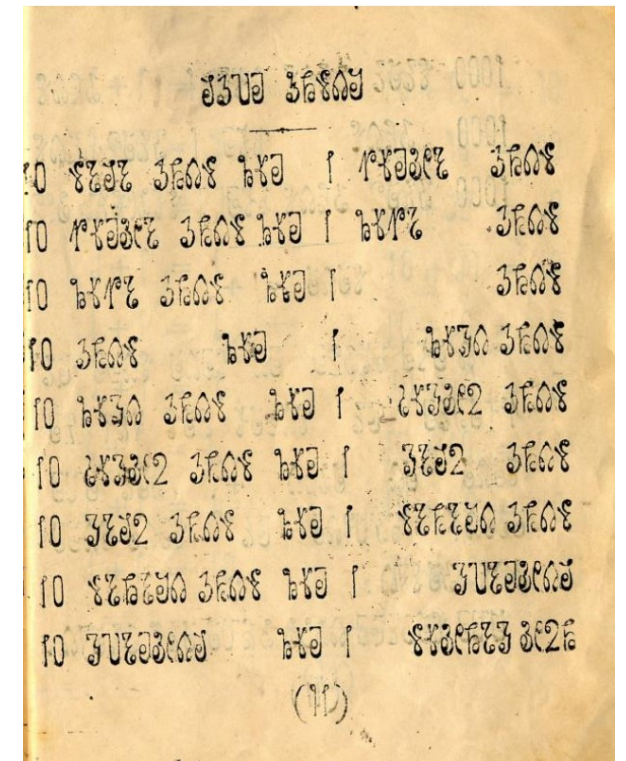
Ol Chiki

ଓଡ଼ ଗ୍ରନ୍ଥ



Warang Citi

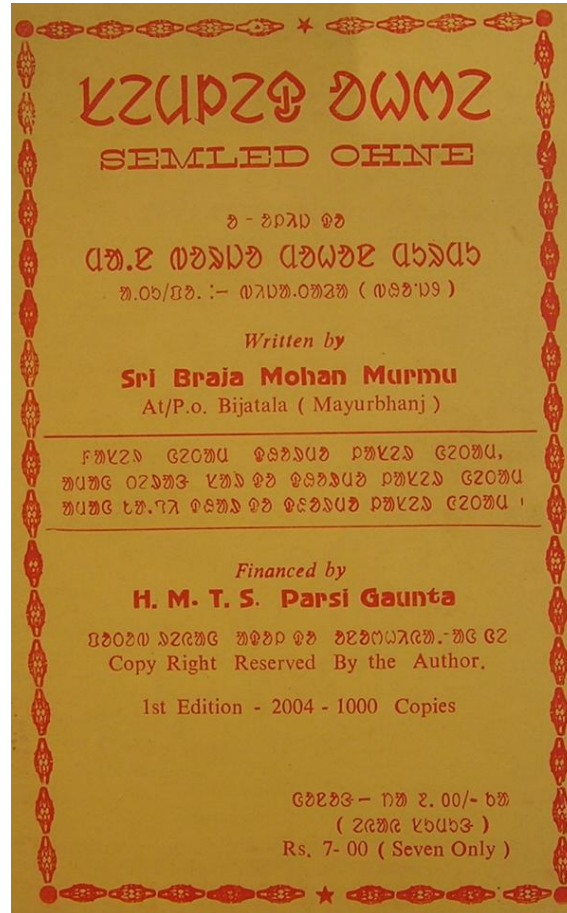
□□□□□  
□□□□□



Sora Sompeng

ꠘꠞꠟꠞ ꠘꠞꠟꠞꠞꠞ

# Ol Chiki

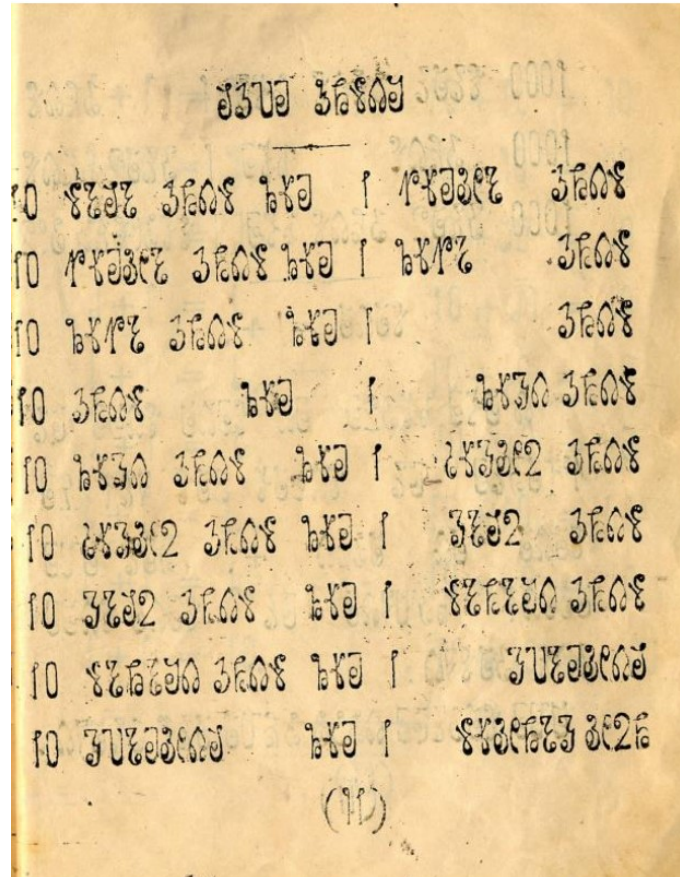


- Developed by Raghunath Murmu in 1925 for Santali, a Munda language
- Provide distinctive script that is tied to Santali relationship to nature





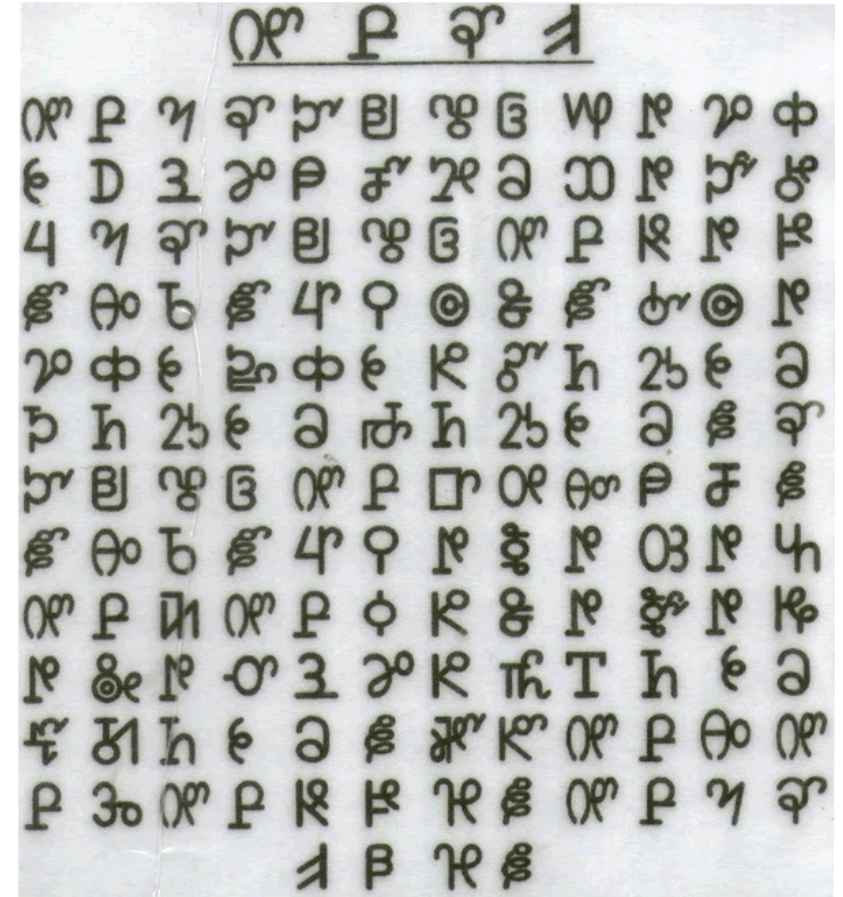
# Sora Sompeng



- Developed by Mangei Gomango for writing , a Munda language, in 1936
- Liturgical script for Sora religious practices

## Other Neographies:

### Laipian lai – Pau Cin Hau

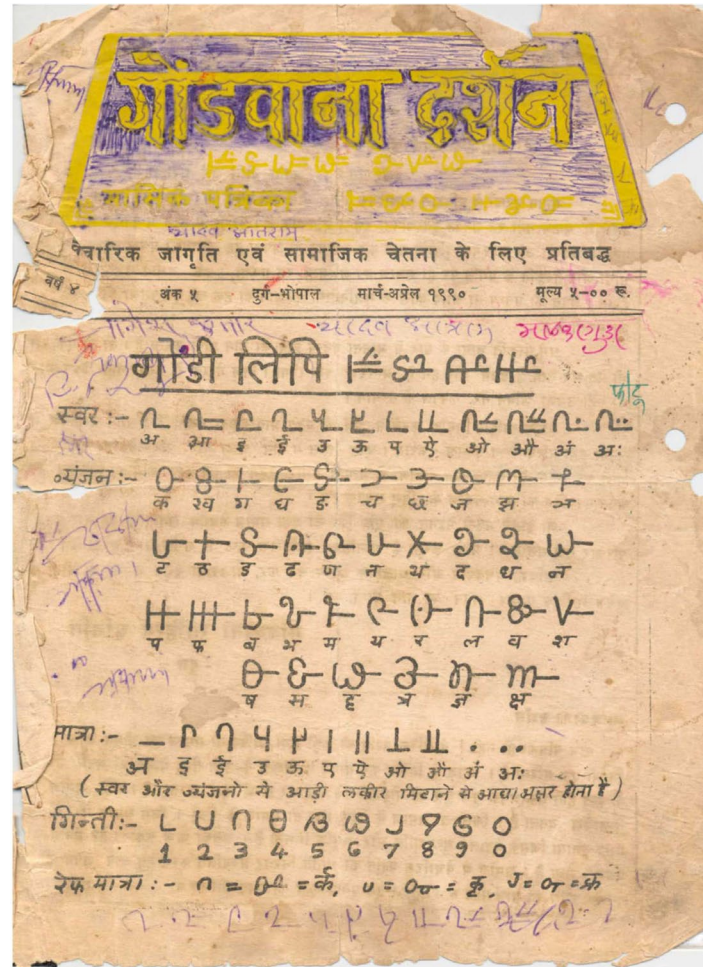











- Developed by Pau Cin Hau in 1902
- Syllabary for writing Tedim, Tibeto-Burman language
- Representation of tones





# Other Neogeographies: Gondi



	<h1 style="text-align: center;">जय परमा पैत</h1>	
<p>अ</p> 	<p>॥ ५ ०</p> 	<p>॥ ५ ० = अड़का ॥ ५ = अना</p>
<p>आ</p> 	<p>॥ ० २</p> 	<p>॥ ० २ आकी ॥ १ ५ आमोट</p>
<p>इ</p> 	<p>॥ १ ॥</p> 	<p>॥ १ ॥ इमला ॥ १ इमा</p>
<p>ई</p> 	<p>॥ १ २</p> 	<p>॥ १ २ ईली ॥ १ ५ ईमाल</p>

- Developed by Mangal Singh Masaram for Gondi, a Dravidian language, in 1928
- Modelled on Brahmi, with unique view towards the inherent vowel



# External Neographies

## Gondi

### Vowels

a	aa	e	ee	u	uu	ae
aaae	i	o	oo	ou	aum	

### Consonants

ka	gha	gna
cha	ja	
ta	da	
tha	dha	na
pa	bha	ma
ya	ra	la
sa	ha	va

### Independent vowels

a	aa	e	ee	u	uu	ae
aaae	i	o	oo	ou	aum	

### Long consonants

ka	gha	gna	cha	ja	ta	da
tha	dha	na	pa	bha	ma	ya
ra	la	va	sa	ha		

## Savara

### VOWELS

A	Aa	E	Ee	U	Uu
Ae	Aae	i	o	ou	Aum

### CONSONANTS

Ka	Ga	Gna
Ja	Eni	
Da	Tha	Na
Pa	Ba	Ma
Ya	Ra	La
Va	Sya	

### CONJUNCT VOWEL SOUNDS

Aa	E	Ee	U	Uu
Ae	Aae	i	o	ou

### LONG LENGTH CONSONANTS

Ka	Ga	ana	Ja	eni	Da
Tha	Na	Pa	Ba		
Ma	Ya	Ra	La	Va	Syah

Prasanna Sree has designed scripts for at least 19 indigenous languages of India

# External Neographies



THE NEW  
**INDIAN EXPRESS**

[NATION](#) [WORLD](#) **[STATES](#)** [CITIES](#) [BUSINESS](#) [SPORT](#) [GOOD NEWS](#) [MOVIES](#) [PHOTOS](#)

**SUBSCRIBE TO E-PAPER**

STOCK MARKET	BSE	61305.95	▲	568.90(0.94%)	NSE	18338.55
--------------	-----	----------	---	---------------	-----	----------

Home > States > Andhra Pradesh

## Andhra University professor enters World Atlas of Endangered Alphabets

*S Prasanna Sree's name, as the first Asian and Indian, has been incorporated in the World Atlas of Endangered Alphabets, the US, 2019, which was inaugurated recently.*

Published: 29th January 2019 09:05 AM | Last Updated: 29th January 2019 09:09 AM  | **A+ A A-**

Unclear motivations?...



# Competing Neographies

प्रभात खबर 17 दिसम्बर 2012 (सोमवार)

## लिपि के लिए तोलोंग सिकि व कुडुख बन्ना की दावेदारी

► प्रथम अंतरराष्ट्रीय कुडुख सम्मेलन संपन्न, कार्यकारिणी की आवश्यक बैठक आज

संवाददाता ■ रांची

प्रथम अंतरराष्ट्रीय कुडुख सम्मेलन का समापन रविवार को हो गया। अंतिम दिन कुडुख भाषा की लिपि के लिए तोलोंग सिकि व कुडुख बन्ना ने अपनी सशक्त दावेदारी पेश की। डॉ नारायण उरांव 'सैदा' रचित तोलोंग सिकि लिपि को झारखंड सरकार ने मान्यता दे दी है। वहीं,

रांची विवि के आर्यभट्ट सभागार में आयोजित सम्मेलन को संबोधित करते छतीसगढ़ से आये सीपी बाखला।

चक्रधरपुर के बासुदेव राम खलखो का दावा है कि उनके द्वारा विकसित कुडुख बन्ना अधिक सरल, सुंदर है व तेजी से लिखी जा सकती है। कहा : कुडुख कई राज्यों में बोली जाती है और कोई एक राज्य इसकी लिपि पर अंतिम निर्णय नहीं ले सकता। यह केंद्र सरकार का मसला है। सोसाइटी के प्रदेश अध्यक्ष प्रो महेश भगत ने कहा कि इस मुद्दे पर कार्यकारिणी समिति सोमवार को निर्णय लेगी।

**आठवीं अनुसूची के लिए लिपि जरूरी :** किसी भाषा को आठवीं अनुसूची में शामिल करने के लिए उस भाषा की अपनी लिपि जरूरी है। राज्य में झारखंड अधिविद्य परिषद की अनुमति से सिसई व डुंवारी के कुछ स्कूल के विद्यार्थी तोलोंग सिकि लिपि से मैट्रिक की परीक्षा दे रहे हैं। डॉ नारायण उरांव ने 1993 में तोलोंग सिकि लिपि तैयार की, जिसे वर्ष 2009 में जैक की मान्यता मिल

गयी। उधर, बासुदेव राम खलखो के अनुसार उन्होंने वर्ष 1991 में ही अपनी लिपि विकसित कर ली थी। कुडुख के विद्वानों को दोनों लिपियों की निष्पक्ष समीक्षा करनी चाहिए। सम्मेलन के आयोजन में प्रो हरि उरांव, प्रो महेश भगत, अशोक कुमार बाखला, डॉ सोसिल खाखा, संजय टोप्पो, सुखराम मिंज, मनीलाल उरांव, विरेन्द्र उरांव व शशि विनय भगत ने महत्वपूर्ण योगदान दिया।

**भाषा किसी भी समाज की पूंजी**

रांची विवि के आर्यभट्ट सभागार में आयोजित दो दिनी सम्मेलन के समापन सत्र में मुख्य अतिथि डीआइजी शीतल उरांव ने कहा कि भाषा किसी भी समाज की पूंजी है। सभी कुडुख भाषियों को अपनी मातृभाषा का प्रयोग करना चाहिए। अपनी भाषा, संस्कृति के प्रति जागरूकता जरूरी है। नेपाल के सांसद शिवनारायण उरांव ने प्रस्ताव दिया कि अगले अंतरराष्ट्रीय कुडुख सम्मेलन का आयोजन नेपाल में किया जाये। फादर आलबर्टुस मिंज ने कहा कि कुडुख को शिक्षा का माध्यम बनाने व रोजगार से जोड़ने की आवश्यकता है। विमल टोप्पो ने कहा कि दुनिया भर के कुडुख भाषी एक दूसरे के साथ लगातार संपर्क में रहे। प्रो चौधरी उरांव ने कहा कि हर दस कदम पर बोली बदलती है, पर साहित्यिक दृष्टिकोण से लेखन ने एकरूपता होनी चाहिए।

**संयोजन, खूंटी**  
खूंटी - 835210

अधिकार  
अध्ययन  
वर्ग

Phone & Fax No. : 06528-221709

**के संचालक हेतु**

Two different scripts for writing Kurukh

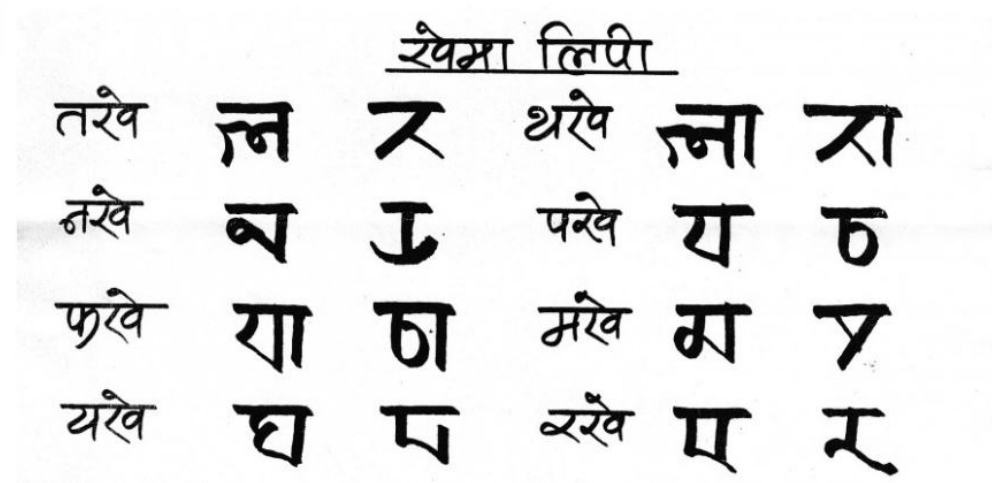
# Competing Neographies

THE KATHMANDU POST

CULTURE & ARTS

## The Gurung language once had no written script. Now it has five

In an effort to preserve its oral language, the Gurung community came up with its own scripts, but it is now divided over which to use.



Five different scripts for writing Gurung



# Features of South Asian Neographies

- Typology
  - Alphasyllabic, on the pattern of Brahmi
  - Alphabetic, LTR, including cursive-joining RTL
  - Syllabic
- Origins
  - Internal development
- Motivations
  - Dream or vision: Pau Cin Hau, Warang Citi
  - First written form for language: Ol Chiki
  - Visible and distinctive identity: all
- Status
  - Script creation is active and scripts are proliferating

# Navigating Neographies:

## Some metrics

- Standardized usage?
  - Is there a standard form of the script
  - Is the script the primary writing system?
- Institutional support?
  - Is there an officially-recognized standard?
  - Is the script routinely used in education?
- Publication?
  - Is the script used in publications? Digital media?
- Literary tradition?
  - Is literary religious material written in the script?
- Community sentiment?
  - Does the target community view the script positively?
  - Are there competing scripts?
- Technological resources?
  - Is the user community developing fonts, keyboards
- Requester?
  - Who is requesting the inclusion of the script in Unicode?



# Motivations for encoding neographies

## Case 1:

- I am an inventor and I have designed a brand new writing system for my community or for a specific linguistic community
- I would like for it to be in Unicode so that it will be standardized.
- Once it is in the Standard I will have something to point to, which might help me persuade my people to use it.

## Case 2:

- We have developed a new writing system for our people.
- We are using this new writing system to publish books and periodicals.
- Our new writing system is being taught in our schools and our people have embraced the writing system wholeheartedly."

# Case Study: Sunuwar

ISO/IEC JTC1/SC2/WG2 N4028  
L2/11-218  
2011-05-31

## Proposal to Encode the Jenticha Script in ISO/IEC 10646

Anshuman Pandey  
University of Michigan  
Department of History  
Ann Arbor, Michigan, U.S.A.  
pandey@umich.edu

May 31, 2011

### 1 Introduction

This is a preliminary proposal to encode the Jenticha script in the Universal Character Set (ISO/IEC 10646). It replaces the preliminary proposal submitted for encoding the script (N3962 L2/10-466R). Changes made to the proposal include the use of English names for signs, the addition of a new combining sign, and more general detail about the proposed characters. Code points have been revised to reflect changes to the allocation for the script in the Unicode Roadmap.

### 2 Background

#### 2.1 Overview

Jenticha is a script used for writing Sunuwar (ISO 639-3: suz), a Kiranti language of the Tibeto-Burman family that is spoken in Sikkim, India and eastern Nepal. Sunuwar is spoken by 40,000 persons in Nepal. The number of speakers of the language in Sikkim is unknown because data for Sunuwar is absent from the 2001 Census of India. However, Sunuwar was recognized as an official language of Sikkim in 1996. Jenticha is a recognized script in Sikkim. Devanagari is also used for writing Sunuwar and is the most commonly used script for the language.

Sunuwar (सुनुवार *sunuvāra*), alternately spelled ‘Sunwar’, is also known by the name ‘Mukhiya’ (मुखिया *mukhiyā*). Both are Nepali exonyms for the language and its speakers. The autonym is कौइच *kōica*, which is normalized as ‘Kōits’. The language is called कौइच लो *kōica lo* ‘Kōits Lo’ or किराँती-कौइच *kirāṭī-kōica* ‘Kiranti-Kōits’. The name ‘Kōits’ is also romanized as ‘Koinch’, ‘Koincha’, ‘Koints’, etc.

The script was developed by Krishna Bahadur Jenticha (1926–1991) in 1942. It is known indigenously as जेंतिच ब्रेसे *jētica brese*, ‘Jenticha Brese’. The script is also called कौइच ब्रेसे *kōica brese* ‘Kōits Brese’, or the ‘Kōits script’; ‘brese’, or script, is also romanized as ‘brehs’ or ‘blehs’, ie. ‘Koinch Brehs’. The script is referred to here as ‘Jenticha’. It has no genetic relationship to other writing systems. However, some of its letters resemble those of other scripts, such as Limbu and Latin.

L2/21-157  
2021-09-10

## Proposal to encode the Sunuwar script in Unicode

Anshuman Pandey

pandey@umich.edu  
pandey.github.io/unicode

September 10, 2021

### 1 Introduction

The last proposal to encode the Sunuwar script in Unicode was submitted over a decade ago (L2/11-218). Since then, I have intermittently monitored the script and have seen evidence of its ongoing usage by Sunuwar communities in both Nepal and Sikkim. However, the encoding effort has been on hold due to inconsistent contact with the user community. A renewal of the project began in June 2020, when Jeannette Stewart of Translation Commons initiated a collaboration between Dev Kumar Sunuwar (Indigenous Media Foundation, Nepal), Deborah Anderson (Script Encoding Initiative), Craig Cornelius, Craig Cummings, Gerry Leonidas, Tex Texin, myself, and others. This group has met at least once a month through video calls to discuss the usage and development of the script. Dev Kumar Sunuwar has facilitated discussions and exchange of documents between the group and the Sunuwar Welfare Society. Lal-Shyakarelu Rapaca, who provided me with information on the script in 2010, is involved in these discussions as part of the Society’s team. This collaboration has confirmed that the script has been in continuous usage since its creation by K. B. Jentich in the 1940s. It has been enhanced by users to better represent the Sunuwar language and is actively used, taught, and promoted today. Given the stability and usage of the script, it is suitable for inclusion in Unicode.

The current proposal supersedes the following documents:

- L2/10-466R: “Preliminary Proposal to Encode the Jenticha Script”
- L2/11-218: “Proposal to Encode the Jenticha Script”

The major changes between it and L2/11-218 are as follows:

- Change of the Unicode designation for the script from ‘Jenticha’ to ‘Sunuwar’
- Usage of Sunuwar names for the letters instead of Latin transliterations
- Ordering of letters according to Jentich’s arrangement
- Addition of letters to complete the repertoire
- Recommendation to use existing combining signs for various diacritics instead of script-specific signs
- Expanded description of nasalization sign, vowel length mark, tone marks, and legacy diacritics
- Addition of an auspicious sign representing a mystical sound in Sunuwar culture

A request has been submitted to register the code “sunu” for the script in ISO 15924.

# Sunuwar



३४४१५३४ (प्रकाशकीय)

90 वृषजघ्न (हामी आयौं)

[illegible]

**हाम्रो** मार्फत कोईच सुनुवार सामू हामी एउटा सदस्यका रुपमा जोडिएका छौं। नाम कमाउने अभिलासा होइन, मनग्रे धन कमाउने लोभ पनि होइन। कोईच समक्ष केही सिकाउन र सिक्ने जोशमा जोडिएका हौं। कोईच समुदायको आवाजलाई बुलन्द गर्न थोरै भएपनि **हाम्रो**ले योगदान अवश्य पुर्याउने छ। हाम्रो प्रयासमा तपाईंहरू पनि जोडिए आभार्यक छ।

अरूले कसो गर्छन् भन्नेतिर हैन, हामीचाहिँ कोइँचसामु शब्द र अक्षरका प्रतिबद्धता लिएर आएका हौं । यो मैदानमा आफैँले उठाएका प्रत्येक मुद्दा, विषय र त्यसका पुष्टिकरणमा केन्द्रित रहनेछौं । यो आफैँसँगको निरन्तरको स्पर्धा हनेछ ।

अवस्था र परिवेश जोखिमपूर्ण छ । यो जोखिमलाई नजरअन्दाज नगरी थोरै भए पनि जानकारी दिन सकिन्छ भन्ने धेयले नै हामी आएका हौं । हामी **हाम्रो**माफ्ट कोइँचको संस्कार, संस्कृति र परम्परालाई जगेर्ना गर्ने उद्देश्यले मैदानमा औपचारिक रुपमा आएका छौं । हामी कामका माध्यमबाट असली 'बढाई'को लोभमा छौं, कोइँचबापी पाठकको मायाका लागि 'लालची' छौं ।

*for*

प्रकाशक

## Koĩts Script (कोइँच लिपीको वर्णमाला)

(ways of pronunciation with alphabetical order)

Alphabet	Name	pronouns
<b>A. Segmental Alphabet</b>		
<b><i>a. Primary</i></b>		
ᮊ	Dewi	/d/
ᮋ	Tasla	/t/
ᮌ	Eko	/e/
ᮍ	Imar	/i/
ᮎ	Reu	/r/
ᮏ	Utthi	/u/
<b><i>b. Secondary</i></b>		
ᮓ	Kiki	/k/
ᮔ	Mama	/m/
ᮕ	Appho	/a/
ᮖ	Pipi	/p/
ᮗ	Gil/Gas	/g/
ᮘ	Hamsō	/h/
ᮙ	Charmi	/tʃ/
ᮚ	Na	/n/
ᮛ	Bur	/b/
ᮜ	Jurmi/juljal	/j/
ᮝ	loacha	/l/
ᮞ	Otthi	/o/
ᮟ	Secha/selel	/s/
ᮠ	Warchi	/w/
ᮡ	Yat	/y/
ᮢ	Awa	/b/
<b>c. Temshyo Apchi</b>		
ᮣ	Aal	/a/
ᮤ	Donga	/d/

Alphabet	Name	pronouns
ṯ	Thari	/tʰ/
Ḷ	Phar	/f/
Ṽ	Ngar	/ŋ/
Ṽ	Kha	/kʰ/
Ṽ	shyer	/ʃ/
Ṽ	Chhelap	/cʰ/
Ṽ	Tentu	/t/
Ṽ	Thele	/tʰ/
Ṽ	Kloko	/ʔ/
Ṽ	Pwopwo	/p/
<b><i>B. Non-Segmental Alphabet</i></b>		
˘	Nesante	nasal
˙	SotteṼsi	stress
ˆ	Ta RenmareṼsi	high tone
˜	Di RenmareṼsi	low tone
:	Laissi	length
<b><i>C. Digit</i></b>		
Ḷ	Sum	/sum/
Ḷ	Kaa	/ka/
Ḷ	Niashi	/niʔsi/
Ḷ	San	/sã/
Ḷ	Le	/le/
Ḷ	Nga	/ŋa/
Ḷ	Raku	/raku/
Ḷ	Chani	/ʔani/
Ḷ	Sasi	/sasi/
Ḷ	Yan	/vã/



## Discussion

What factors should be taken into consideration when evaluating proposals for encoding neographies?

Thank you

Email: [pandey@umich.edu](mailto:pandey@umich.edu)

Website: <https://pandey.github.io/unicode/>

Email: [dwanders@berkeley.edu](mailto:dwanders@berkeley.edu)

Website: <http://linguistics.berkeley.edu/sei>

Support from NEH grant PR-268710-20 and donations



**NATIONAL  
ENDOWMENT  
FOR THE  
HUMANITIES**