



# THE ADVENT OF MAYAN SCRIPT ENCODING

MAPPING THE LAST FRONTIERS OF MAYAN  
HIEROGLYPHIC DECIPHERMENT

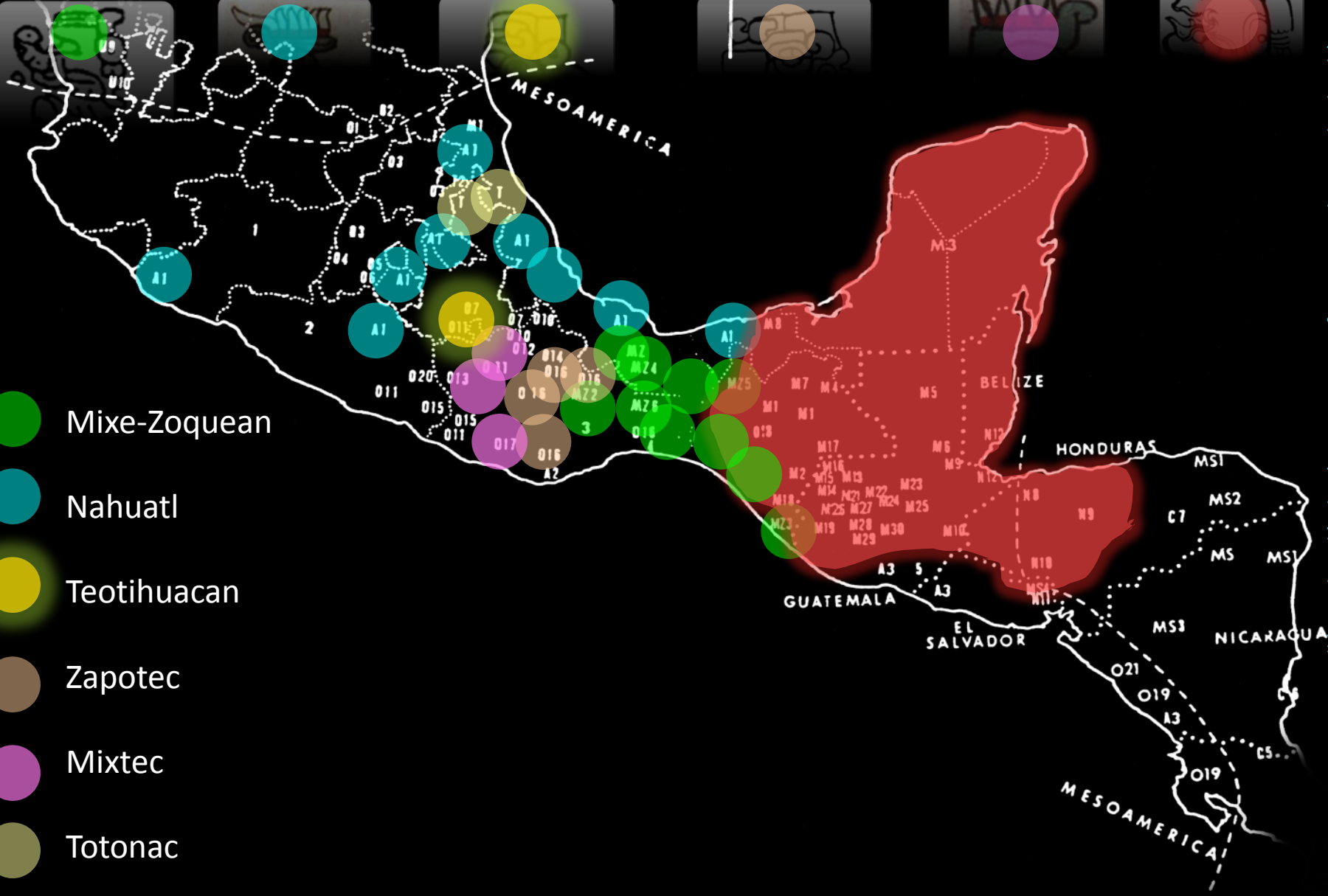
**Carlos Pallán Gayol,**  
NcodeX Project (UC Berkeley / Unicode)  
& University of Bonn (Department of Archaeology  
& Ethnology for the Americas)

**Keynote**

42 Unicode Conference, September 11, Santa Clara, CA



-  Mixe-Zoquean
-  Nahuatl
-  Teotihuacan
-  Zapotec
-  Mixtec
-  Totonac





Lintels



GULF OF MEXICO



Photograph: © Ana Lucia Gastelum  
Monumental sculpture

Codices



Thompson1972 facsimile



Polychrome  
ceramics



K1440 © Justin Kerr

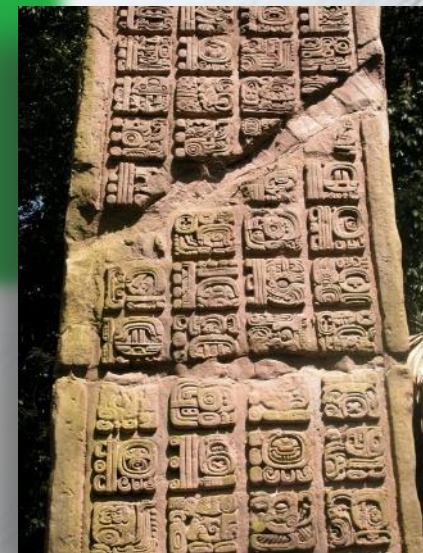


Photo: David Stuart

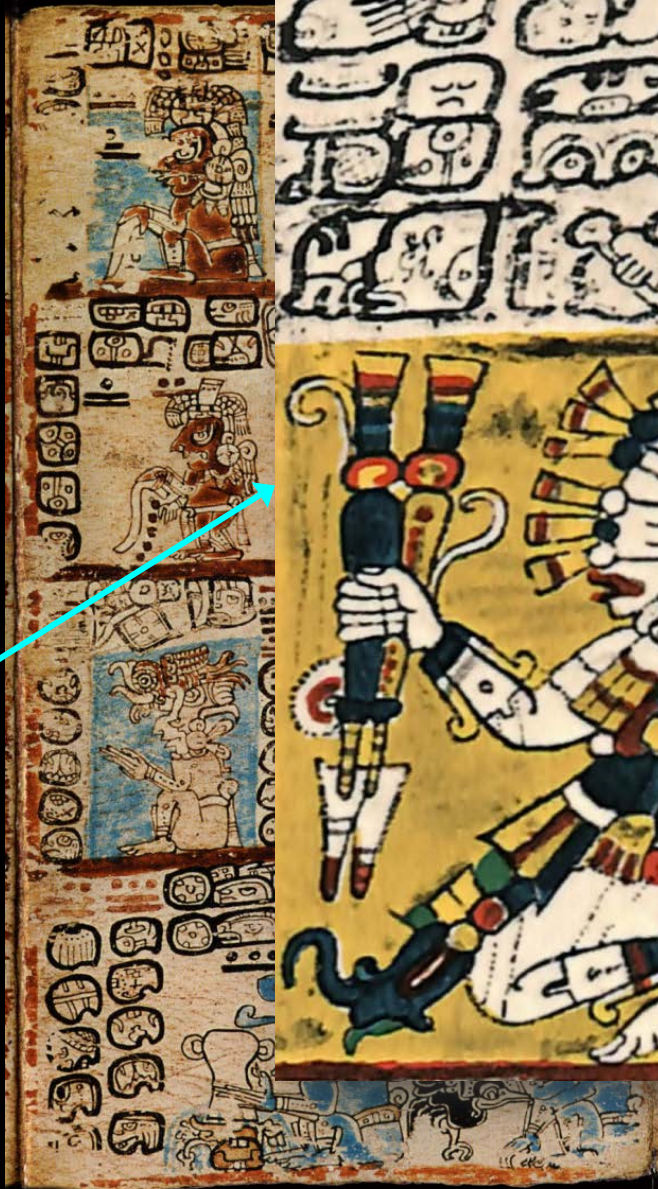
# Mayan Codices



# Mayan Codices



Codex Dresden, p. 50



Codex Madrid, p. 26



Codex Paris, p. 5

# 18<sup>th</sup> Century Expeditions into Palenque, Chiapas



Palenque, Palace & Temple of Inscriptions

Lithography by F. Catherwood, ca. 1840



Palenque, Temple of the Sun

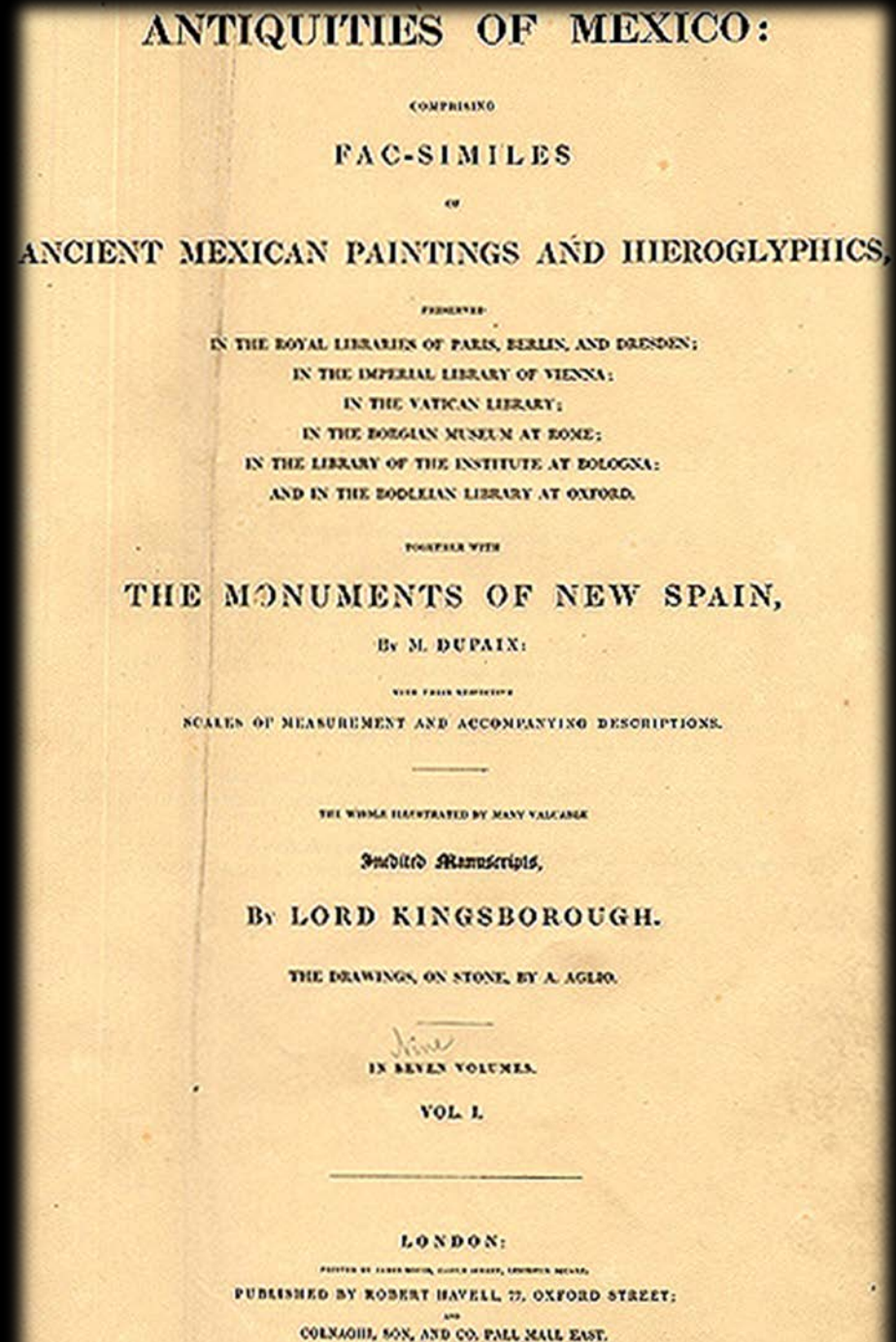
Drawing by Armendáriz, 1787

# Pioneers in Maya Studies



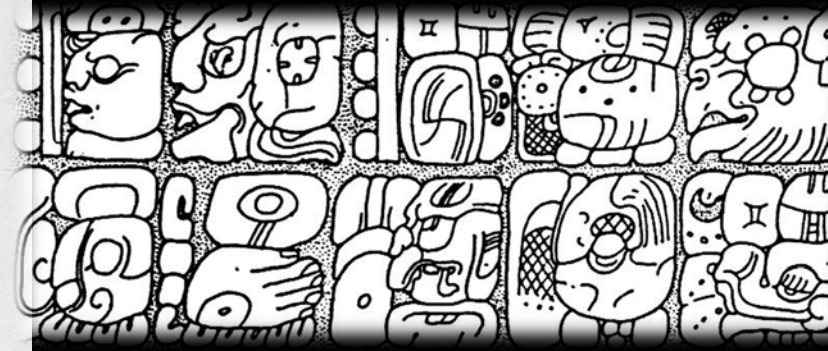
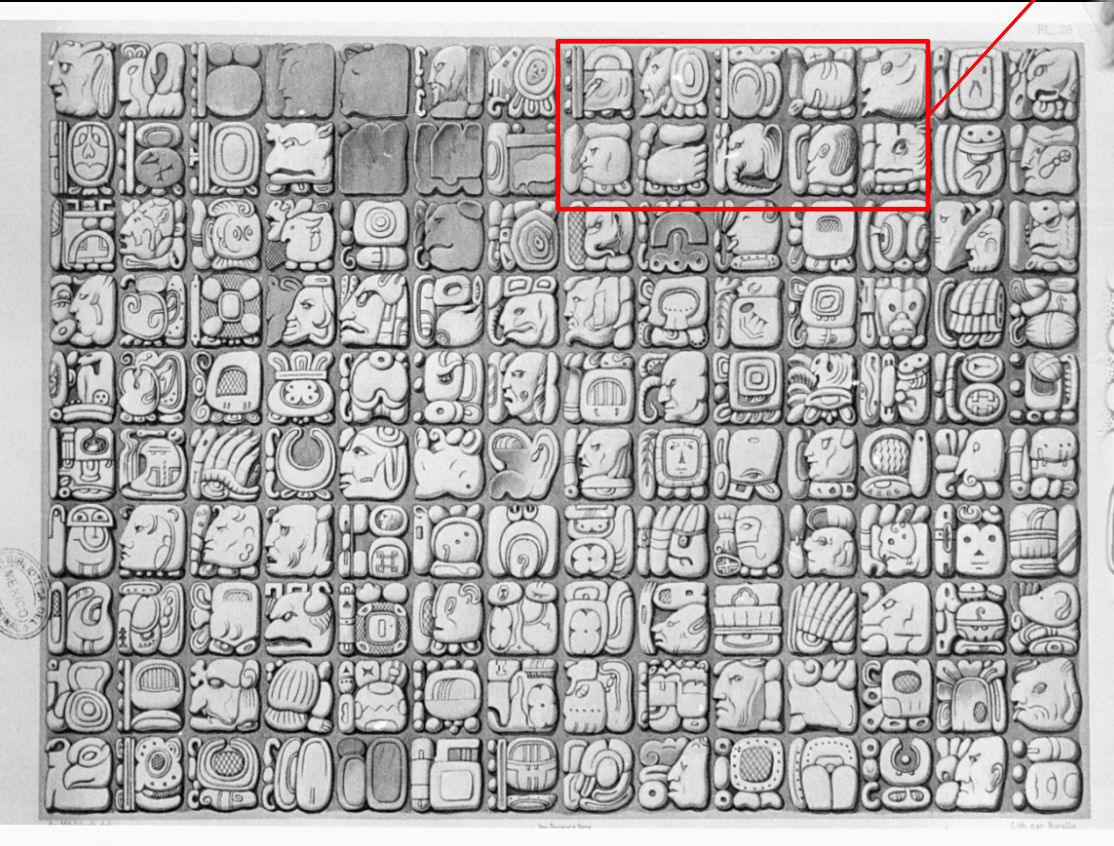
Jean Frédéric Maximilien from  
Waldeck  
(1766(?)-1875)

*Antiquities of Mexico*, Vol. I  
by Lord Kingsborough, London 1831



# Pioneers in Maya Studies:

Original drawing by Waldeck



Modern drawing by Linda Schele,  
ca. 1980s

Waldeck's rendition of Palenque, Temple of Inscriptions, ca. 1830



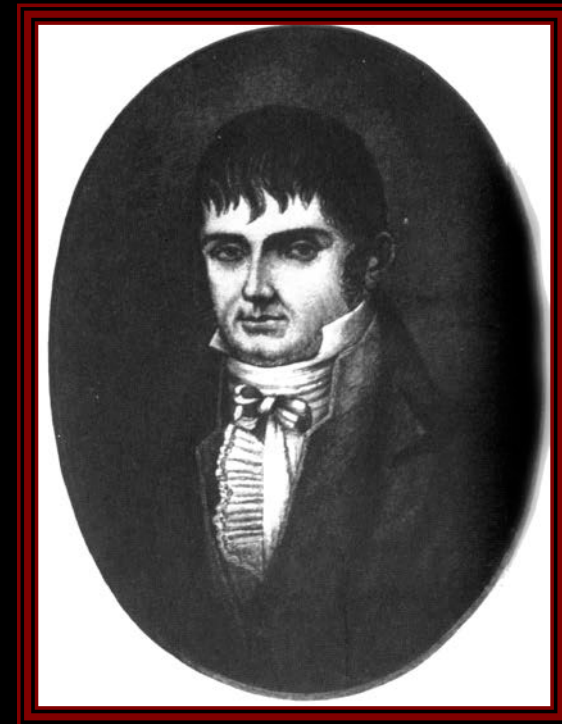
# Mid-19th Century: The Age of Travelers and Explorers

## The Stephens & Catherwood expeditions



*John Lloyd Stephens*

(1783-1840)



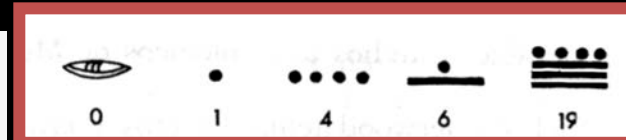
*Constantine Samuel Rafinesque*

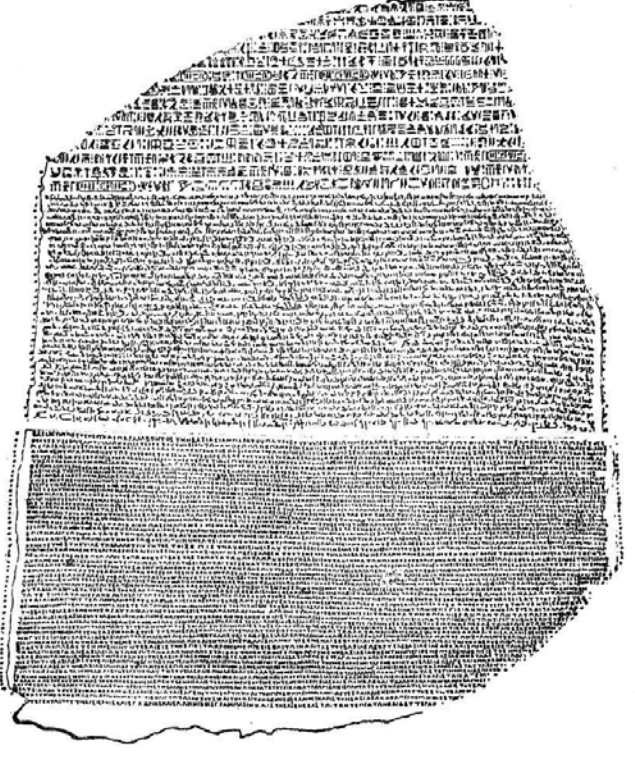
(1783-1840)

*Polygraph naturalist who discovered the functionings of bar-and-dot numerals*

***“ No Champollion has yet thought to them the energies of his inquiring mind. Who shall read them? ”***

Incidents of Travel in Central America, Chiapas, and Yucatan, Vol I





Egyptian

Demotic

Greek



*Friar Diego de Landa (1524-1579)*

**Rosetta Stone**

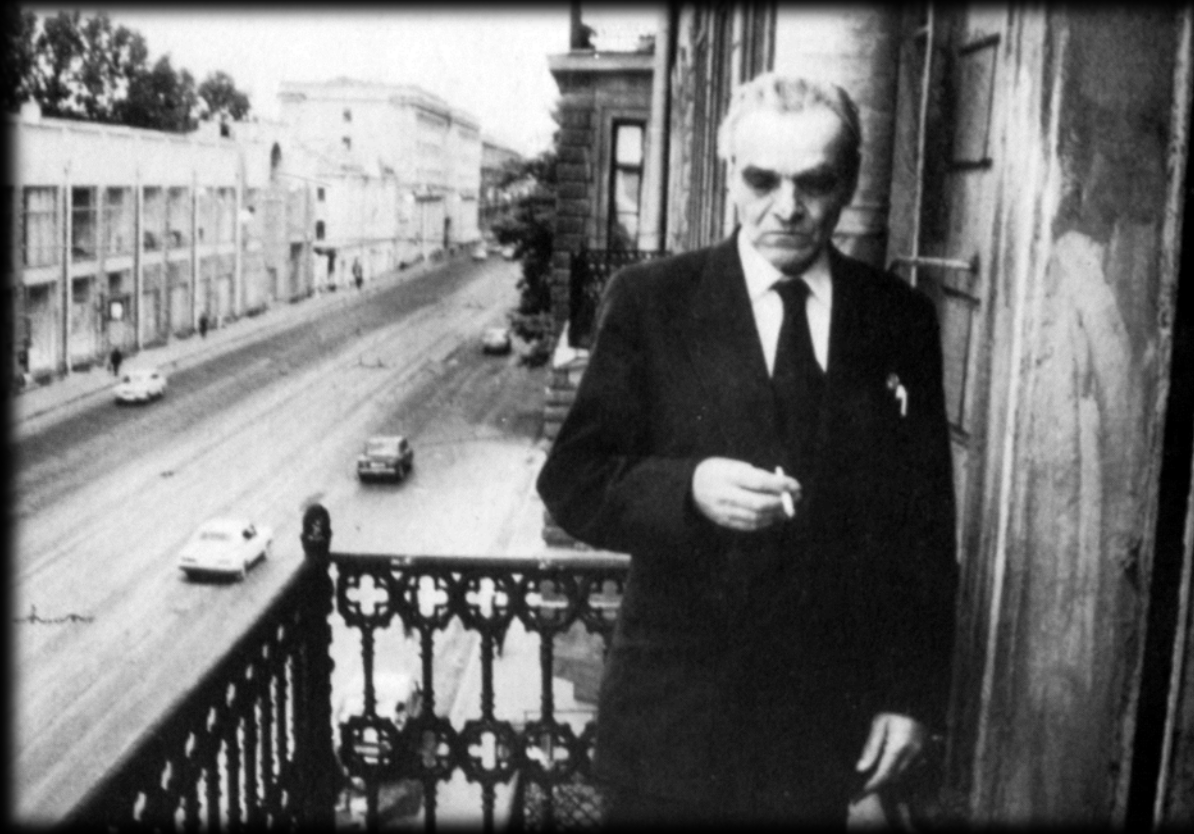
**Landa's "alphabet" (syllabary)**



*Jean François Champollion (1790-1832)*



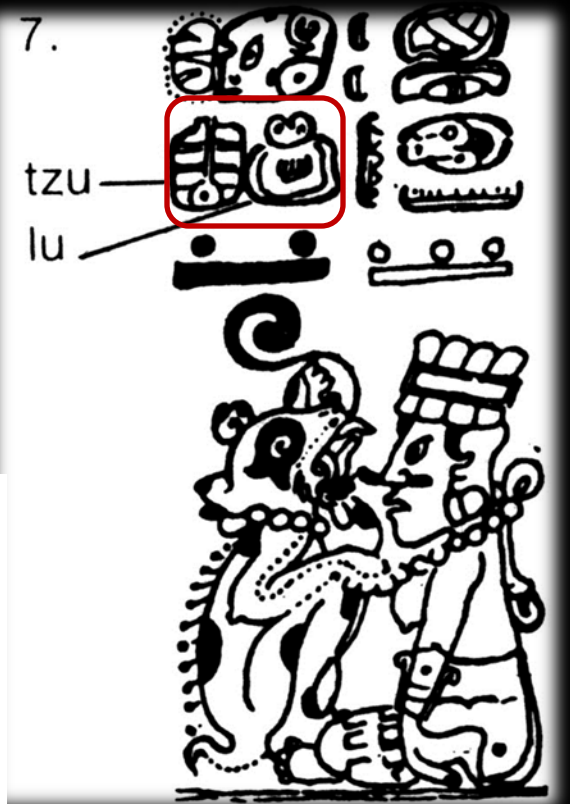
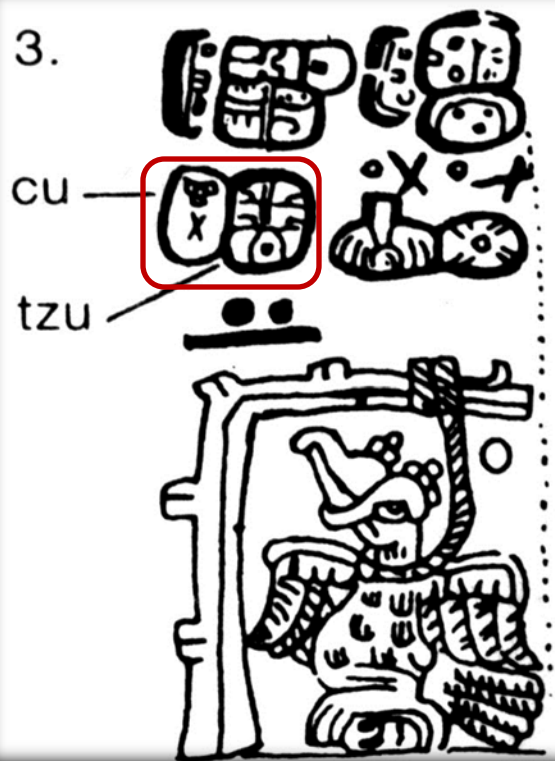
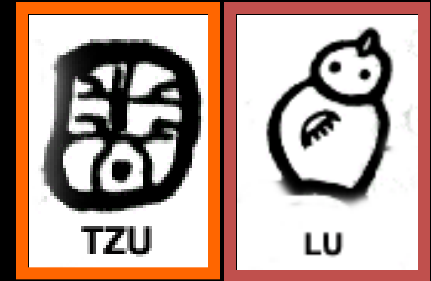
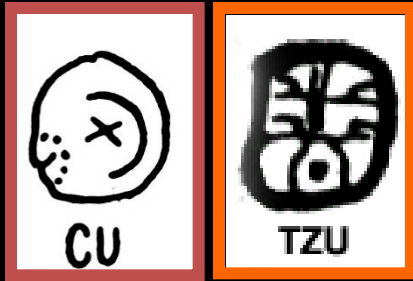
# Knorozov's Phoneticism



**Yuri Valentinovich Knorozov (1922-1999)**

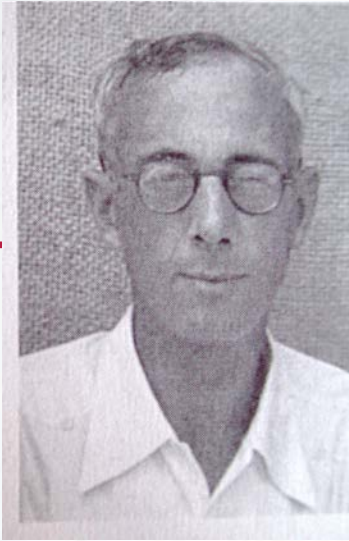
at his Leningrad office, USSR 1989

# Knorozov's Methodology



Decipherments for *kutz* "turkey" & *tzul* "dog"

# DECIPHERMENT: HISTORICAL AGE, CA. 1960

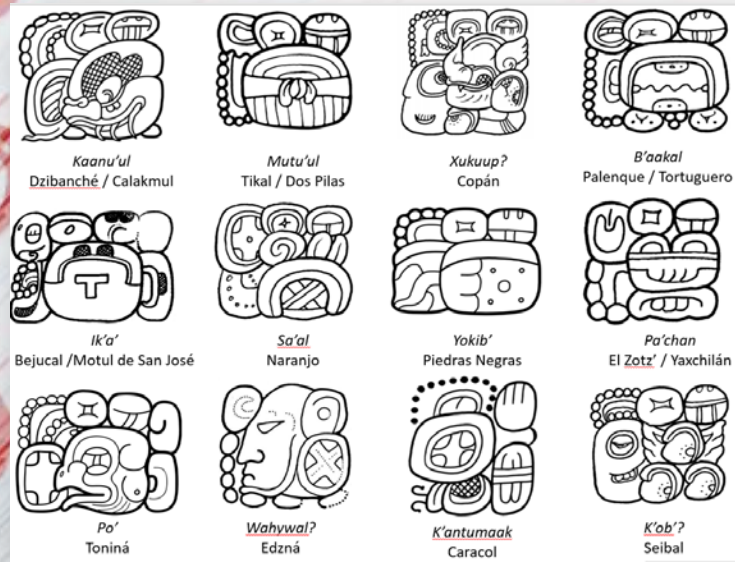


Heinrich Berlin (1915-1988)

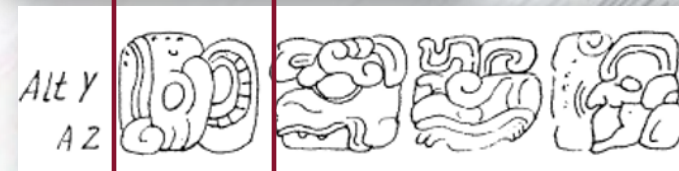
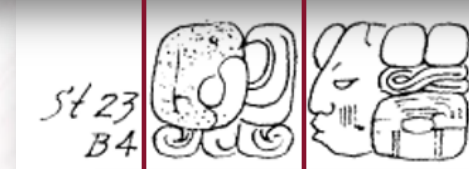


Tatiana Proskouriakoff (1909-1985)

## EMBLEM-GLYPHS:



## HISTORICAL DATES IN THE INSCRIPTIONS:



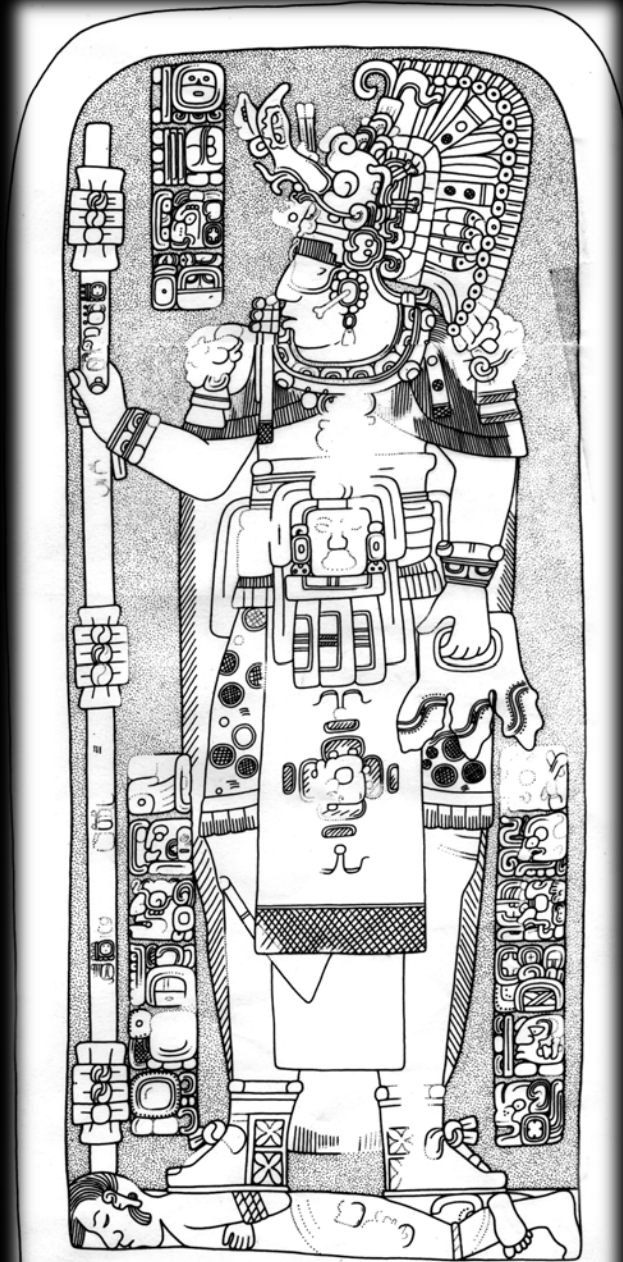
Ian Graham (1923-2017):  
*Corpus of Maya Hieroglyphic Inscriptions (CMHI)*



*The Road to Ruins*  
IAN GRAHAM



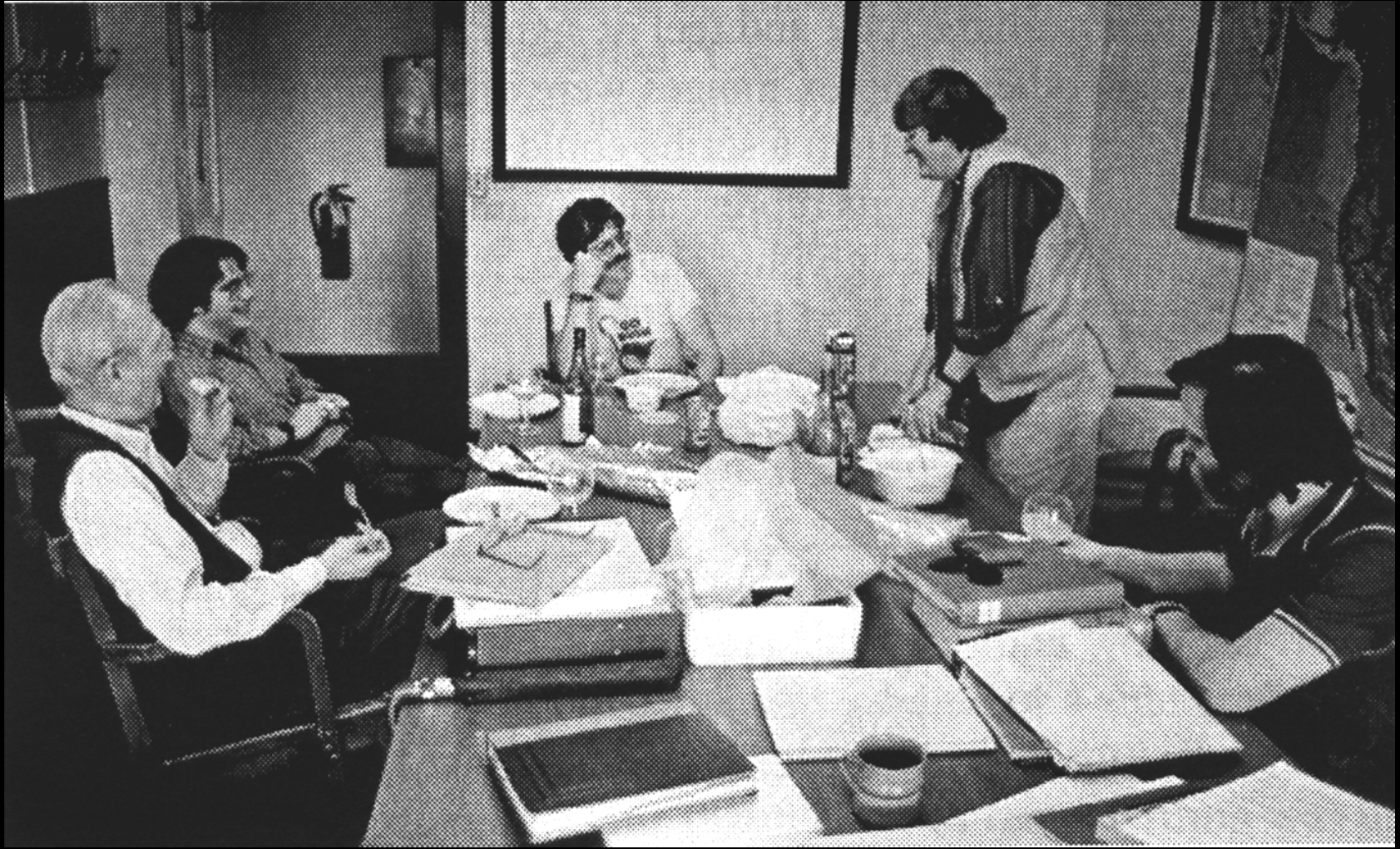
Lintel from Yaxchilán (Mexico)



Stela from Naranjo (Guatemala)

Drawings: Ian Graham

# The Modern Era of Mayan Decipherment



Left to Right: Floyd Lounsbury, David Stuart, Peter Mathews, Linda Schele and John Justeson at Harvard, 1984



# Yaxchilán Expeditions 2007-2008 with Ian Graham collaboration Peabody Museum (Harvard) with INAH, Mexico



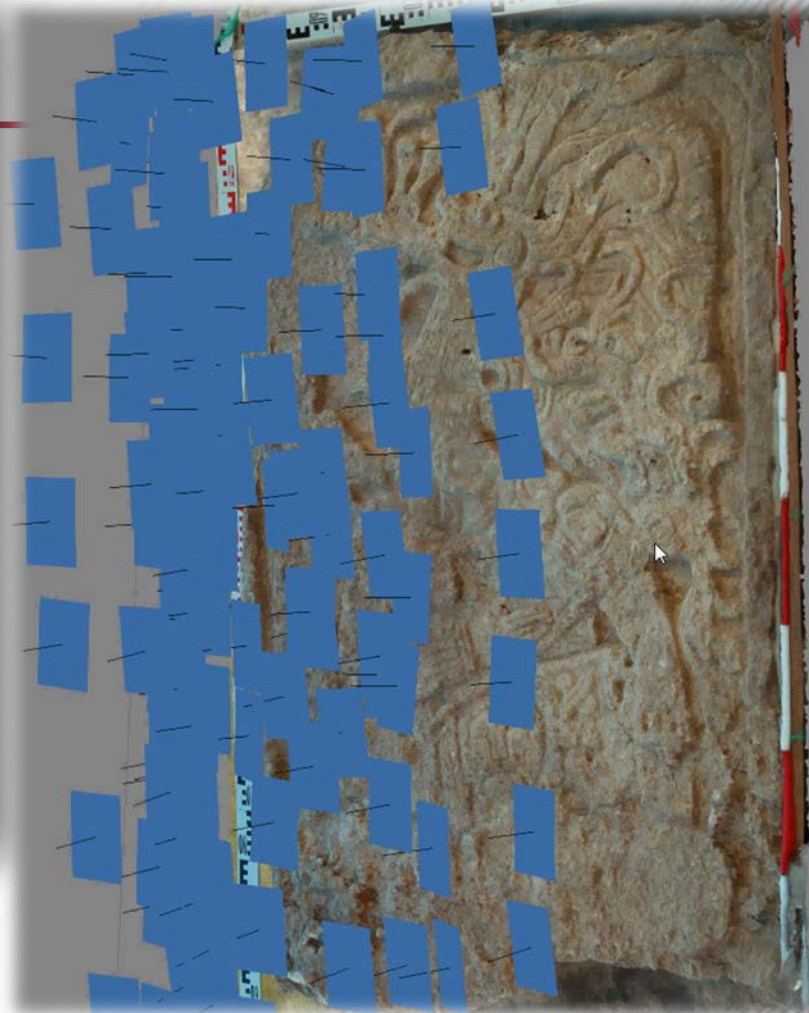
Expedition to Yaxchilan. Ian Graham with Carlos Pallan checking Stela 4.

# Fieldseason 2013: Manantial, Guatemala

## Discovery of Manantial Stela 6

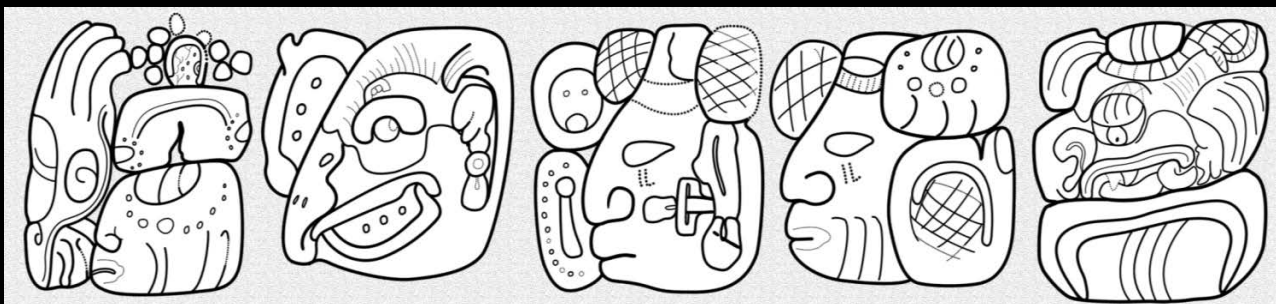
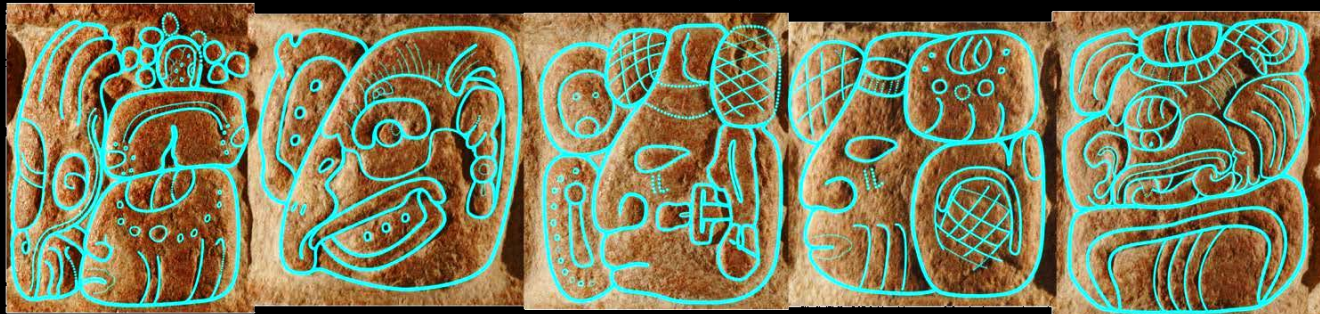


The Uaxactun Archaeological Project  
Dr. Milan Kováč with Carlos Pallán and team members



EL MANANTIAL, PETEN, Stela 6,  
3D Photogrammetry by Carlos Pallán.  
Uaxactún-SAHI Project 2012

# Digital Humanities methods first introduced in Mexico (Edzna, Campeche, 2006)



Photos & Drawing: Carlos Pallán  
AGIMAYA-INAH 2006

EDZNA ALTAR

ca. 657 AD.

# New Syllabary for Codical signs (NcodeX Project 2017-2018)



PROJECT NcodeX: Updated syllabary of Mayan codical hieroglyphic signs (v3)  
by Carlos Pallán Gayol (January 17, 2018)



	VOWELS	b'	ch	ch'	h	j	k	k'	l	m
a										
e										
i										
o										
u										

n	p	s	t	t'	tz	tz'	w	x	y
					see /se/ (related to tze)				

# SCRIPT ENCODING INITIATIVE AT UC BERKELEY

## SEI

**Script Encoding Initiative**  
Department of Linguistics  
University of California, Berkeley

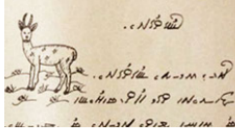
- Home
- Scripts to Encode
- Progress Overview
- UTC Reports
- News & Presentations
- Press
- About Us
- How to Donate
- Donors

### 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

The goal of the SEI project is to fund the preparation of proposals that will be successfully approved by the Technical Committee and WG2 (ISO/IEC 10646) via extensive revision or involvement of the committee.



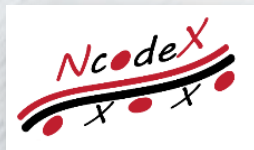
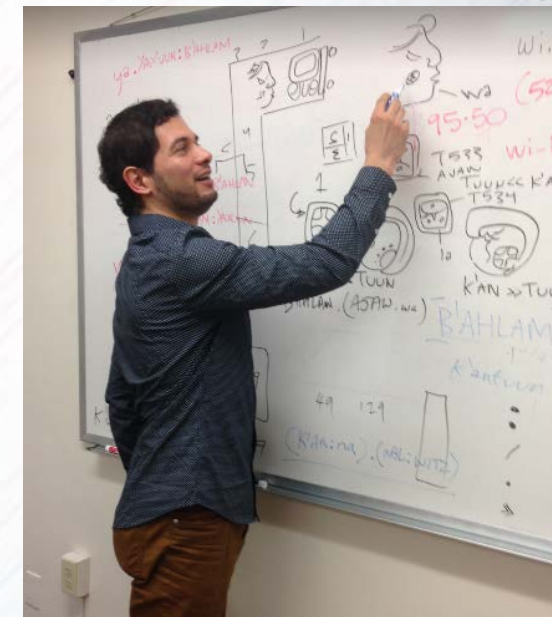
A secondary goal of the SEI project is the creation of a Unicode-compatible font. This will help with the widespread adoption and implementation of the standard.



Dr. Deborah Anderson

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

- 2015 – 2017, set up meetings at UC Berkeley between Carlos Pallán and Unicode experts
- Discussed
  - Requirements to generate a proposal for Mayan
  - Ways to handle complex Mayan features in Unicode
  - **Inception of the NcodeX Project:**
- Funding received from **Adopt-a-Character program of the Unicode Consortium (2017)**
- Project supervision by Lisa Moore / Deborah Anderson



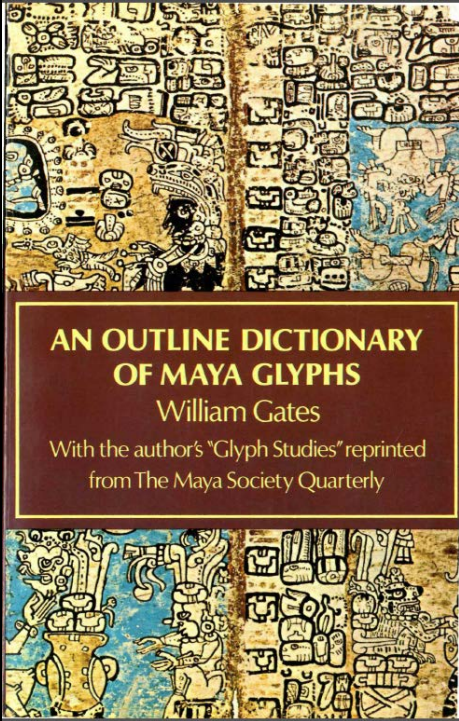
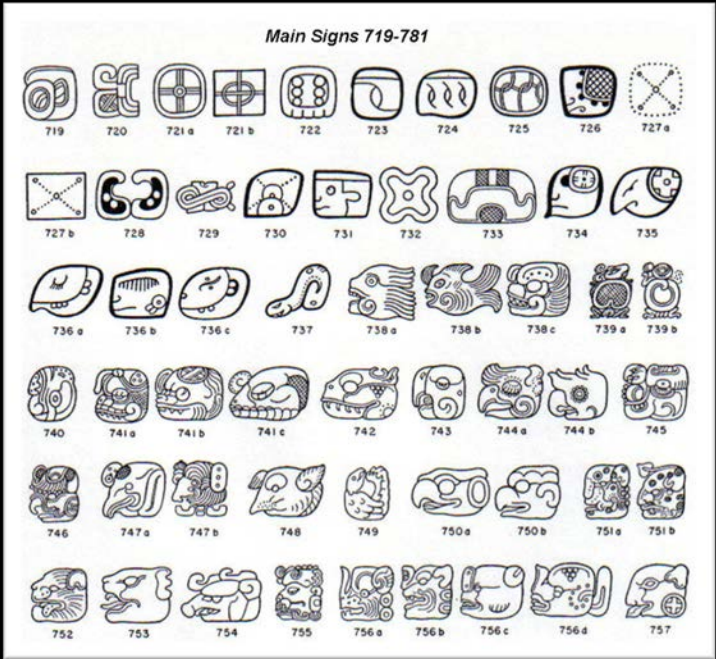
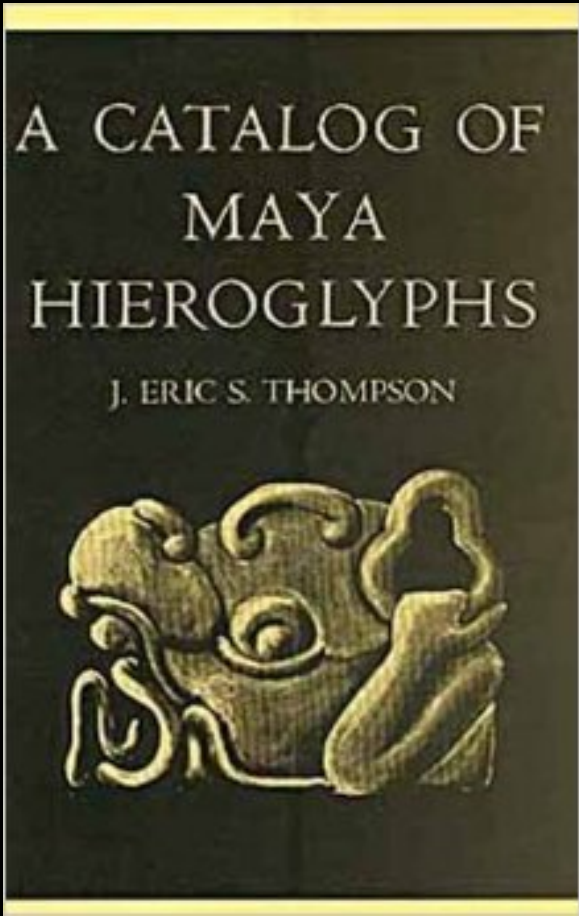
# WHY UNICODE?



- Stable format - international standard
- Allows copying and pasting text
- Search capability both within documents and across Internet
- Provides long-term access
- Aids in discoverability



# Previous Mayan Glypharies & Catalogs



Gates 1931  
 "An Outline Dictionary of  
 Maya Glyphs"

Thompson, J. E. S.  
 (1964). *A Catalog of Maya Hieroglyphs*. University of Oklahoma Press.

# NEW GLYPHARY TOOL. Mapping of signs & contexts

CODICAL  
CONTEXTS

CLASSIC-PERIOD  
CONTEXTS

The screenshot displays the Glyphary software interface. At the top, there is a menu bar with options like 'Layout: V2a\_GLYPHARY\_CONTEXT', 'View As:', and 'Preview'. Below this is a header area with fields for 'id', 'id\_SIGNS', and 'ID\_PhotoCatalog'. The main workspace is divided into several sections:

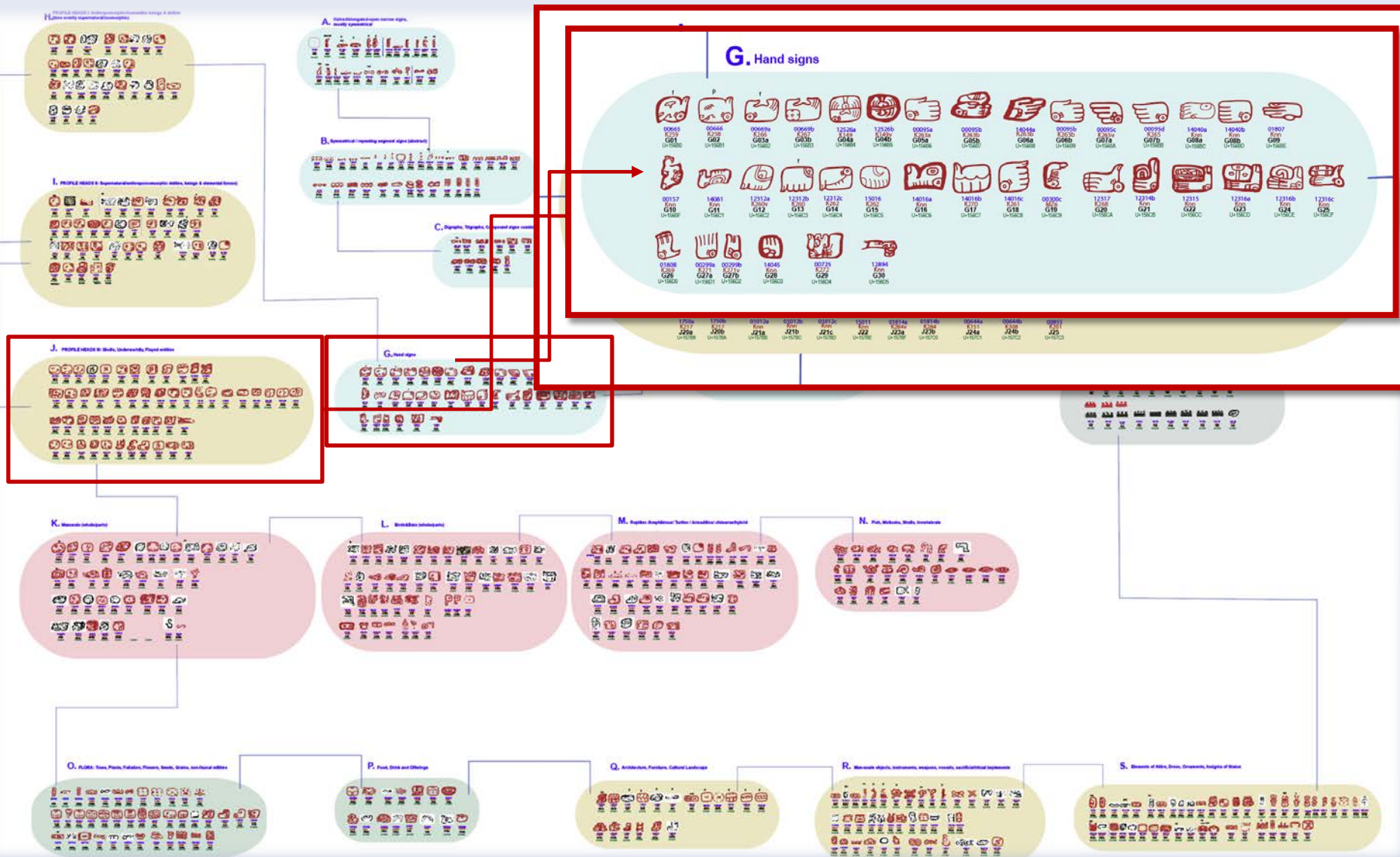
- Left Panel:** A vertical list of signs, each with a small thumbnail and a color-coded label (e.g., 1384, 12310, 1715, 3390, 3971, 4587, 5569, 6393, 6662, 6955, 7604, 7729, 7810).
- Central Workspace:** A large area showing a grid of signs and their contexts. The grid has columns for different contexts and rows for different signs. A sign with ID 12310 is highlighted. Below the grid, there are detailed views for specific signs, including 'Tokoviline 2017' and 'KNORO SOV 1983'. These views show the sign's appearance in different contexts and provide additional information like 'Name AT: CH'EE' and 'Name YK: 155'.
- Right Panel:** A vertical list of signs, similar to the left panel, but with a different set of thumbnails and labels (e.g., 1384, 12310, 1715, 3390, 3971, 4587, 5569, 6393, 6662, 6955, 7604, 7729, 7810).

Red arrows point from the labels 'CODICAL CONTEXTS' and 'CLASSIC-PERIOD CONTEXTS' to specific areas of the interface, indicating the mapping of signs and contexts.



# • GLYPHARY tool

## • NEW TAXONOMY FOR CODICAL SIGNS. EXPAND TO CLASSIC DATASETS













































# SIGN-REPERTOIRE LIST. UNICODE CODE-POINT CHART

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
1550																
1551																
1552																
1553																
1554																
1555																
1556																
1557																
1558																
1559																
155A																
155B																
155C																
155D																
155E																
155F																
1560																
1561																
1562																
1563																
1564																
1565																



# OTHER HIEROGLYPHIC SCRIPTS IN UNICODE

- **Egyptian hieroglyphs** (1,071 chars. published in Unicode 5.2, 2009)

	1300	1301	1302	1303	1304	1305	1306	1307	1308	1309	130A	130B	130C	130D
0	 13000	 13010	 13020	 13030	 13040	 13050	 13060	 13070	 13080	 13090	 130A0	 130B0	 130C0	 130D0
1	 13001	 13011	 13021	 13031	 13041	 13051	 13061	 13071	 13081	 13091	 130A1	 130B1	 130C1	 130D1
2	 13002	 13012	 13022	 13032	 13042	 13052	 13062	 13072	 13082	 13092	 130A2	 130B2	 130C2	 130D2





























<http://unicode.org/charts/PDF/U13000.pdf>

- **Meroitic hieroglyphs** (published in Unicode 6.1, 2012)



<https://fonts2u.com/meroitic---hieroglyphics.font>































- **Anatolian hieroglyphs** (published in Unicode 8.0, 2015)

	1440	1441	1442	1443	1444	1445	1446	1447	1448	1449	144A	144B	144C	144D
0	 14400	 14410	 14420	 14430	 14440	 14450	 14460	 14470	 14480	 14490	 144A0	 144B0	 144C0	 144D0
1	 14401	 14411	 14421	 14431	 14441	 14451	 14461	 14471	 14481	 14491	 144A1	 144B1	 144C1	 144D1

<https://www.unicode.org/charts/PDF/U14400.pdf>

# Unicode approaches relevant to Mayan: Chinese-Japanese-Korean (CJK) descriptors

Pattern

 U+2FF1	 U+2FF0	 U+2FF8	 2FF4	 2FF3	 2FF2	not available	not available	not available	not available	not available	not available
											
关	母	匄	囿	咎	衙	not attested?	not attested?	not attested?	not attested?	not attested?	not attested?
											

CJK

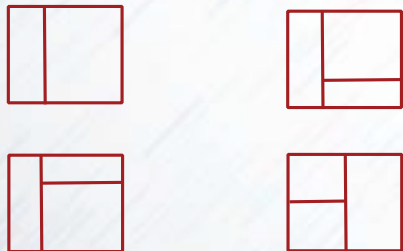
Mayan

# LIST OF *QUADRATS*

New, more advanced *Quadrats*  
 (being created by Carlos Pallan, in  
 collaboration with Stephen White and  
 Andrew Glass)  
 Manually created bounding-boxes:



Auto-generated *Quadrat*  
 abstractions



A. CLASS	B. RANKING	C. QUADRAT DIAGRAM	D. EXAMPLE FROM CODICES	E. NUMERIC GENERAL DESCR.	F. LETTER GENERALDESCR	G. LETTER SPECIFIC DESCR.	H. NUMERIC SPECIFIC DESCR.
1.01	340			1	a	A	11
2.01	668			1H2	a.b	a.B	1H22
2.02	248			1H2	a.b	A.b	11H2
2.03	105			1H2	a.b	A.B	11H22
2.04	112			1V2	a:b	A:B	11V22
2.05	97			1V2	a:b	a.B	1V22
2.06	68			1V2	a:b	A:b	11V2
2.07	11			1H2	a.b	(A°).B	(11R)H22
2.08	16			1L2	a[b	AαB	11II22
2.09	2			1L2	a[b	A`B	11SI22
2.10	2			1L2	a[b	A`B	11SI22
2.11	26			1L2	a[b	A`B	11S22
2.12	6			1L2	a[b	Aαb	11II2
2.13	2			1L2	a[b	A`b	11SI2
2.14	1			1L2	a[b	a»B	1I22
2.15	3			1H2	a.b	(A).((B))	(11(H))22))
2.16	2			1L2	a[b	a»B	1I22
2.17	2			1H2	a.b	a.∞B	1HDI22
2.18	2			1V2	a:b	A:(B°)	11V(22R)
2.19	0			1V2	a:b	(A°):(B)	(11R)V(22)



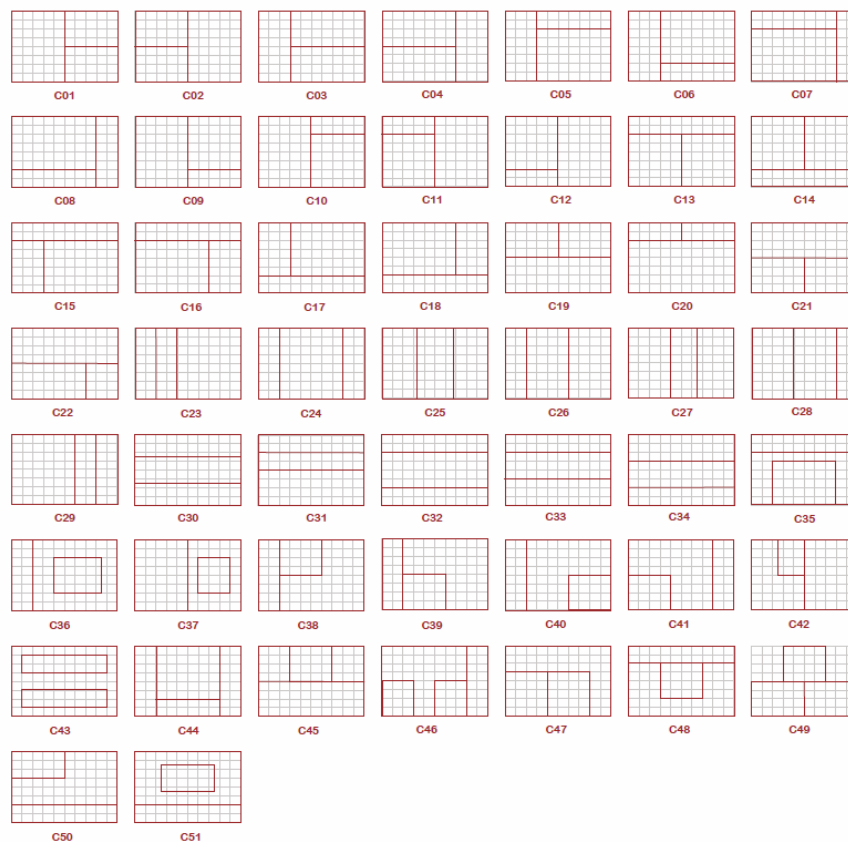
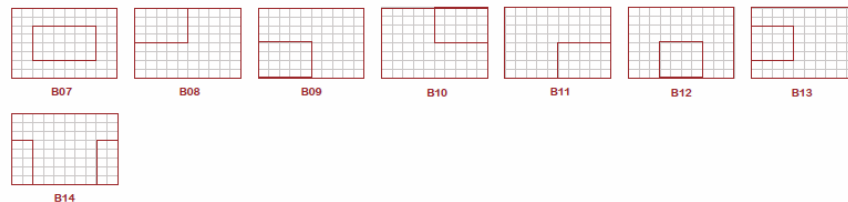
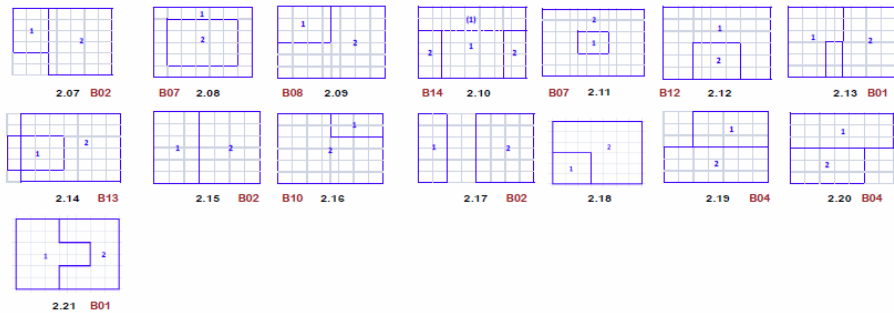
SEI

Script Encoding Initiative  
Department of Linguistics  
University of California, Berkeley

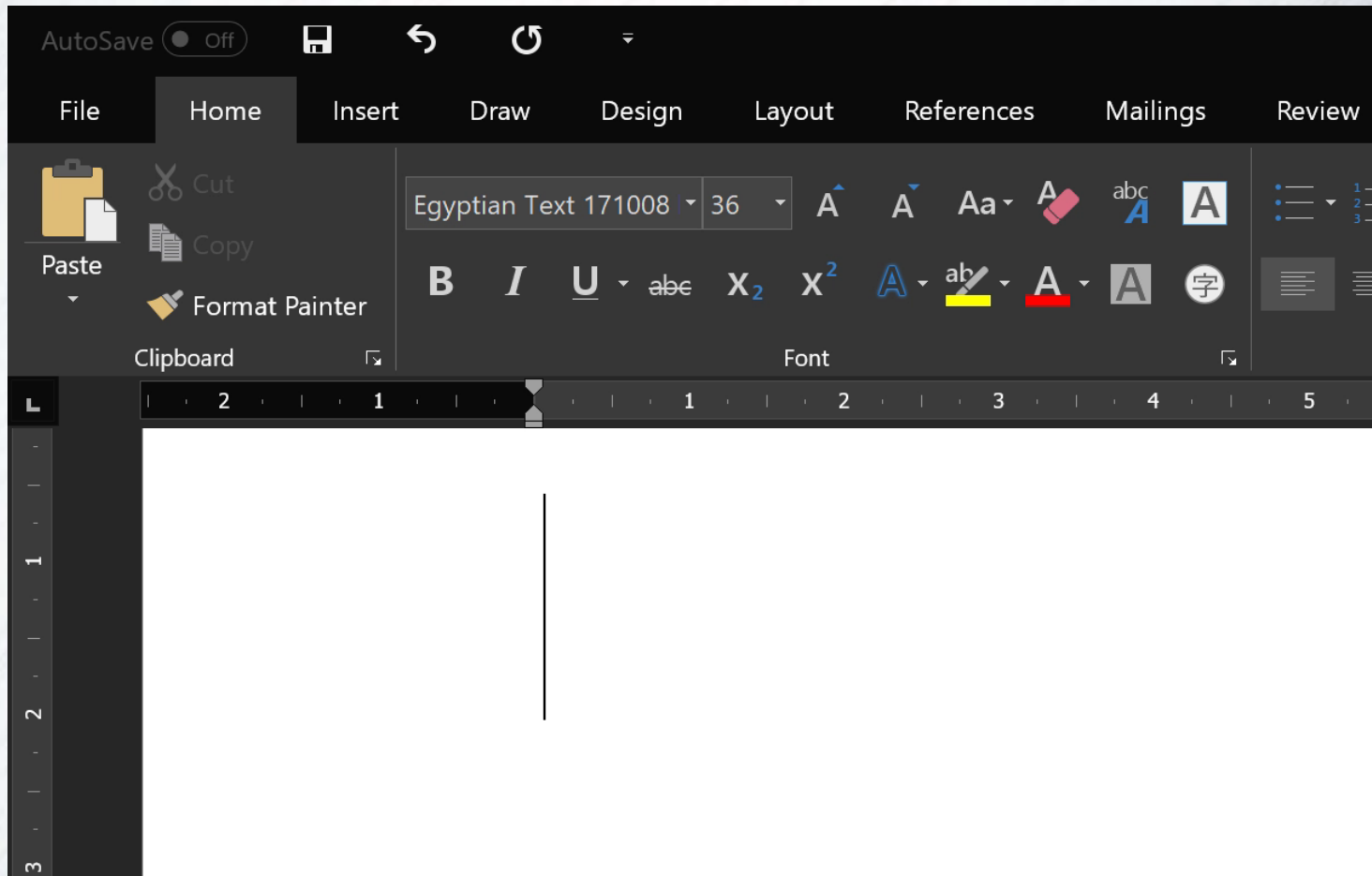
# LIST OF NEXT-GENERATION *QUADRATS* (ARRANGED BY COMPLEXITY)

Previous List of Quadrats (2017). Total count: 164

New & simplified List of Quadrats (2018). Total count: 142

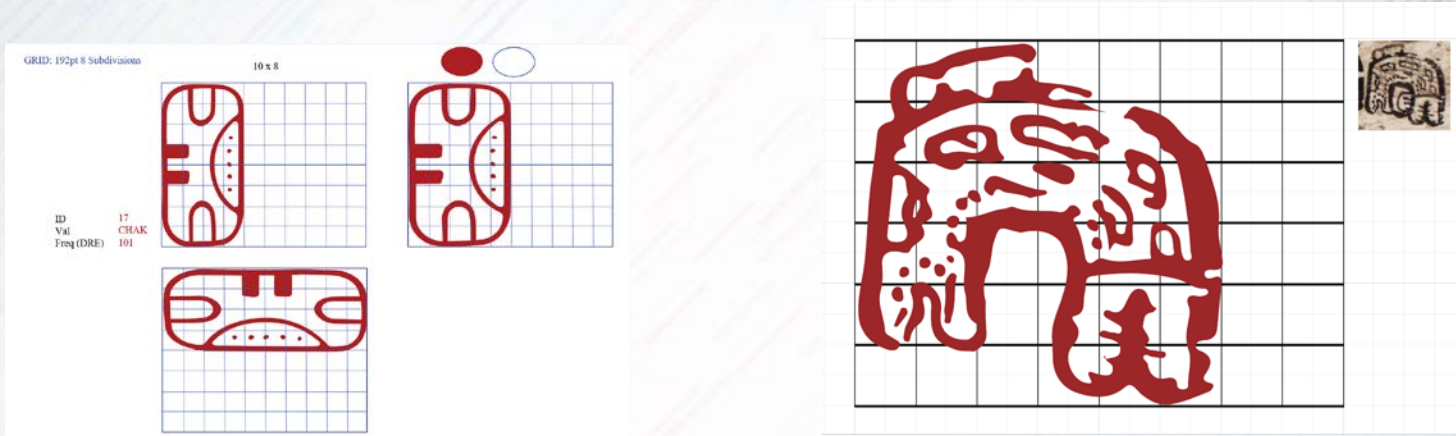


# Unicode approaches relevant to Mayan: Egyptian hieroglyphs clustering (by Andrew Glass, Microsoft)



- User types sequences and presses a conversion key to convert to the associated Hieroglyph
- Structuring is dynamic based on internal font logic for the signs and joining controls
- Prototype font is fully Unicode conformant, anticipating quadrat controls that will become official in Unicode 12

# MAYAN FONT DEVELOPMENT



Visualizations for Open-Type Mayan font  
 (by Carlos Pallan, in collaboration with Andrew Glass, Microsoft)

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
1.55E																
1.55F																
1.560																
1.561																
1.562																
1.563																
1.564																



Special thanks to Adobe Inc. & Matthews Rechs for supporting our Project with complimentary Creative Cloud licenses



# ENCODING MAYAN TEXT: NON-LINEAR RENDERINGS OF GLYPH-BLOCKS



a	b
MV/HE6	MV/2S6
55	50
u	K'AK'
3221	3216

a	b	c
MV/3M2	MV/XH3	MV/1G1
129	524	1130
ti	CHAN/KAN	na
3293	3678	4279

a	b	c
MV/004	MV/2S1	MV/BP4
1722	12276	11296
CHAN/KAN	NAL?	MO'
4485	3317	4647

a	b	c
MV/XQ3	MV/ZC1	MV/XH2
12299	12297	510
K'IN	ku/TU:	HA:B'
3660	3645	3664

A01.B12



N02.B04:A03



N04.J06:G05



B21:B07.N22




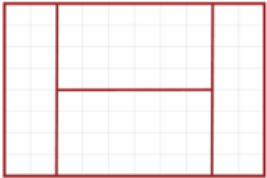




Result: vectorial font correctly rendered non-linearly with the above *quadrats*:

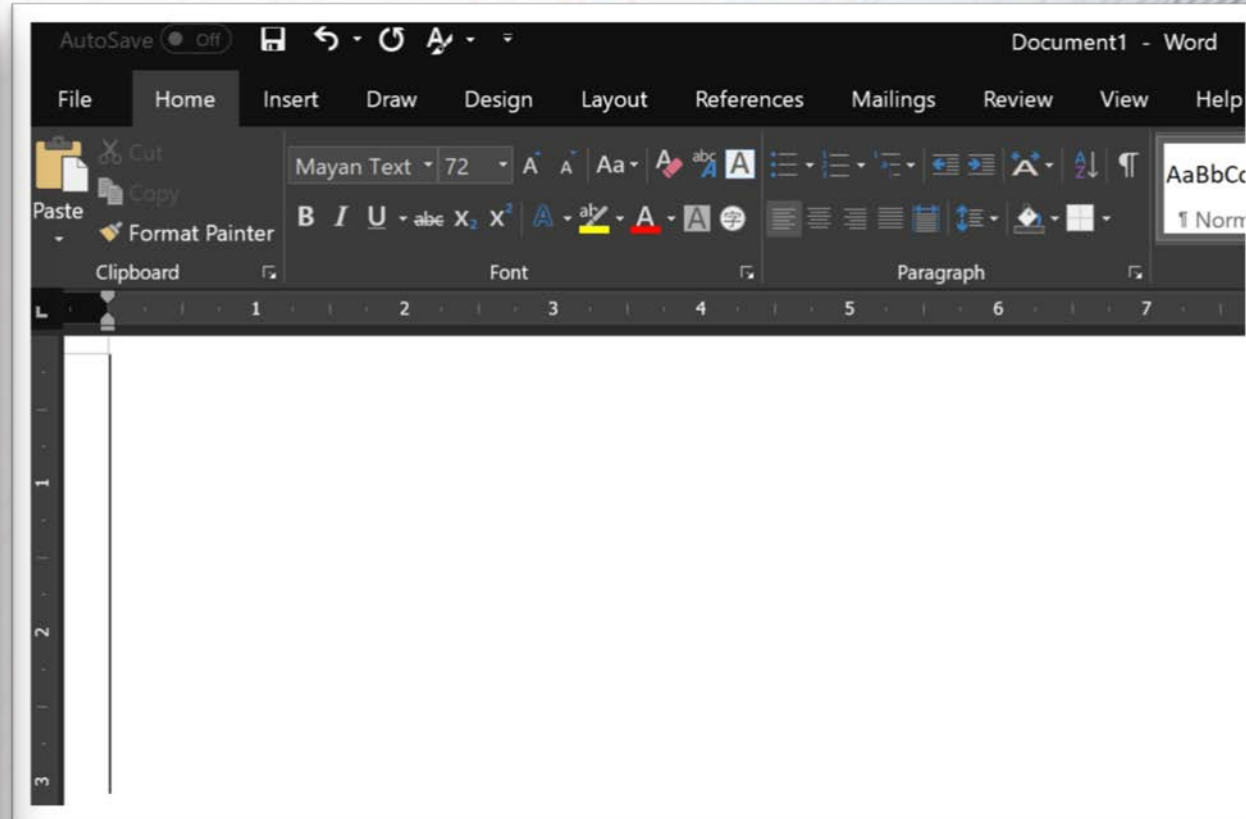


# New Mayan Font in Practice

Foundational data  
(by Carlos Pallan)

Mayan hieroglyphs clustering  
(by Andrew Glass, Microsoft)

	Transcription: u-te-k'a-ja Context: DRE_08e_D1
 D12	Associated quadrat: D12
	<b>F09</b> ID: 1110 U+1566F Phonetic value: u
	<b>F12</b> ID: 1500 U+15675 Phonetic value: te
	<b>G02</b> ID: 666 U+15681 Phonetic value: k'a
	<b>E01</b> ID: 260 U+155E0 Phonetic value: ja



- User types sequences and presses a conversion key to convert to the associated Hieroglyph
- Structuring is dynamic based on internal font logic for the signs and joining controls
- Prototype font is fully Unicode conformant, the required quadrat controls not yet official, but will be similar to those already being implemented for Unicode 12

# TEXT REPOSITORY: CODICAL TEXTS SELECTED FOR ENCODING

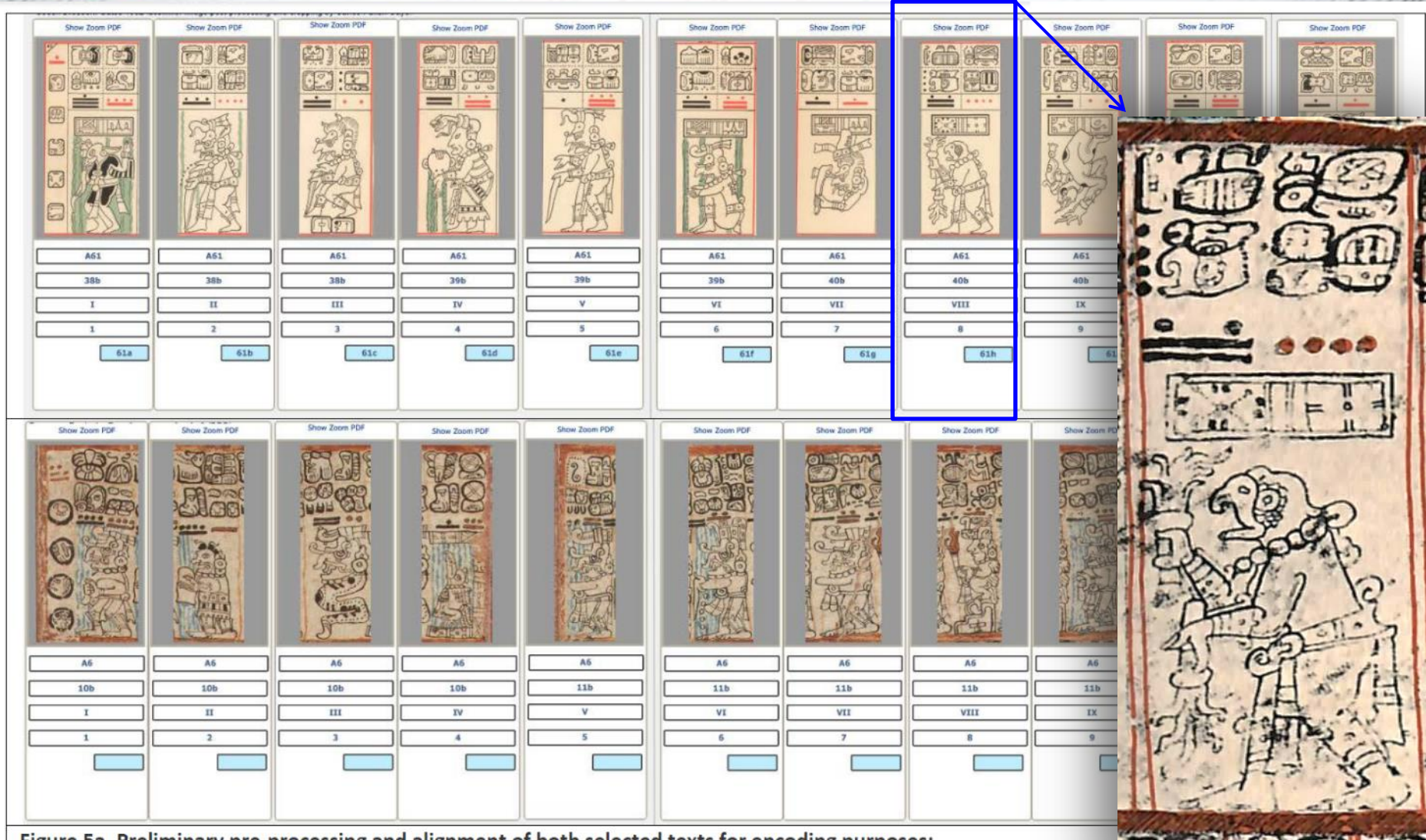


Figure 5a. Preliminary pre-processing and alignment of both selected texts for encoding purposes:

TEXT 1 (top): Dresden Codex (Almanac 61, pages 38b-41b, 11 frames)

TEXT 2 (bottom): Madrid Codex (Almanac 6, pages 10b-11b, 9 frames)

- CLOUD-COLLABORATIVE RESEARCH & DISSEMINATION OF PROJECT OUTCOME
- MAYAN-READ (Research Environment for Ancient Documents)

- Work towards implementing online Mayan-READ with Andrew Glass & Stephen White since 2017
- Planned to continue into 2018-2020

• Auto-generated Palaeography charts:

The screenshot shows the Mayan-READ interface. On the left, there's a sidebar with 'CATALOG', 'EDIT', 'VIEW', and 'LAYOUT' sections. The main area displays a document image with green rectangular annotations. Below the image is a 'Paleography for BC04' chart.

	a	i	u	e	o
vowel					
k	ka	ki	ku	ke	ko
ks		ksi			
kh	kha			khe	
g		gi		ge	
gr	gra				
gh					gho
c			cu	ce	co
i	ia	ii	iu	ie	

• Advanced text-image annotation & markup

<https://github.com/readsoftware/read/wiki>

The screenshot shows the Mayan-READ interface with a document image and a detailed text view below. The text view shows a list of lines with their corresponding glyphs and annotations.

BC04 [lepidici](#)

41 777 + 777777 + ..... karata vñiñadi bhavidi. d. ....

42 na haksati - se apa ñegena margbhavane haksati dukho + 777 + ..... (\*tredhduade viratasa viragaatusa)

43 (\*e) otane haksadi chidasa viñadi priñi haksati viñadi ñoa na haksati trae ca (\*durgadi na haksati trae ca sugadi (\*haksati) trae ca saparala meksa haksati trae sadriñia

44 77 trae sadriñia suha haksati kaisandegia trae ca ñuho na haksati trae ca ñuho (\*haksati) trae añuho na haksati trae ca kugala haksati trae akugala na haksati

45 (\*trae sapurupa)darñano haksati budhapatreana (\*trae durgana na haksati : 7777 mapurvagama (\*apapurupa)) (\*navadma na) haksati sarvasapati (\*haksati) + ..... (\*haksati

ti - vado nidana ca akusala pavena karat\*mena kajrane kusale purfaksae na haksati

46 ..... karata vñiñadi bhavidi. d. ....

47 ..... ki haksati? 777 + .....

48 ..... (\*vñiñadi kujala ca karana vñiñadi [to do c] - [u h] - r. - patoña - na haksati apañes(\*e)ña (\*m)gathah-

49 (\*vane haksati dukho) + ..... (\*e)dhduade viratasa viragaatuse c tane haksati chidasa viñadi priñi

410 haksati viñadi ñoa na haksati trae ca durgadi na haksati trae ca sut\*gedi haksati trae? + + 7 ñ sadriñia ca trae suha haksati trae dukha na haksati kalace-

411 ñegia trae ca ñuho haksati trae añuho na haksati trae kusala haksati trae (\*kujala na haksati trae sapurupadarñano haksati budhapat(\*e)ña trae ñuho na (\*ña) haksati

412 ya mama ca sakñetere sarvasapata na haksati sarvasapata haksati mokasaha ca hal\*kañadi (me anuñada haksati sarvasaha paricata = 1 gaganadivalegamañaga-

413 (\*ñhu) tarania chisamerthaj ñaki uadane gaganadivalegamañadhadu(\*ñhu)ñe vñeja ñaki (\*ñadana ko vardi pa 7 pe gaganadivaleja maladñadu tañra)ñeja ñaki

414 paritcana gagañeñiñazamaladñaduñe vñeja ñaki paritcana (ga) ko vardi ta a ro paricata na uadi na uadane te sarve ñe anuvadane te sarve sapati mokasapati ca - 2

415 ..... + 7 tribosae sa añañhedi kareñi gahañrañadi + ..... 77 nisamarho va ñuho va sudñamo ca gajabati (ca) se maha ñemi udi odhamañuade

<https://github.com/readsoftware/read/wiki>

## EXPANDING INTO THE CLASSIC-PERIOD CORPUS

Potential team members and collaborators contemplated for 2018-2020 activities:

- Dr. Gabrielle Vail
- Holly Neville
- Céline Tamignaux
  
- Mayan-READ:
  - Andrew Glass
  - Stephen White
  - Ian McCrabb
  
- INAH & Mexican DH experts:
  - Dr. Diego Jiménez Badillo
  - Dr. Edgar-Román Rangel



Vessel K1440. Photograph: © Justin Kerr. Courtesy: Dumbarton Oaks Research Library and Collections, Washington DC



Copán, Honduras, Stela A  
Photograph: © Ana Lucía Gastelum

## Concluding Remarks:

Bringing our vast Ancient American Heritage in the Information Age



Illustration from © National Geographic Magazine by Terry Rutledge



## Many Thanks!

- Lisa Moore, UNICODE, CFO & Chair of the UTC
- Alolita Sharma, UNICODE, Board Director
- Dr. Deborah Anderson, SEI / UC Berkeley (Dept. of Linguistics)
- Andrew Glass, Microsoft, READ
- Stephen White, READ
- Ken Whistler, Unicode
- Instituto Nacional de Antropología e Historia (INAH, México)
- Acervo Jeroglífico e Iconográfico Maya (Ajimaya-INAH)
- Roozbeh Pournader, WhatsApp
- Project "Codices"
- Project "MAAYA"
- Research Environment for Ancient Documents (READ)
- University of California at Berkeley, *Department of Linguistics*

## Acknowledgements:

- This project was supported through funding from the Unicode Consortium's Adopt-a-Character program
- Adobe Inc.
- Universal Scripts Project (Script Encoding Initiative) housed at the University of California at Berkeley (Department of Linguistics).
- National Endowment for the Humanities:  
NEH Grant PR-253360-17



Seibal, Guatemala (Carlos Pallán 2012)