



# Glyph Dwellers

Report 75

January 2022

## MHD Reference Materials 5: The Historical Development of the Maya Script: Preliminary Results

**Matthew Looper**

Department of Art and Art History, California State University Chico

**Martha J. Macri**

Professor Emerita, Department of Native American Studies

Research Professor in Linguistics

University of California, Davis

### Introduction

Since its inception more than thirty years ago, a primary goal of the Maya Hieroglyphic Database (MHD) project has been to generate the data required to understand more fully the historical development of the Maya script. Upon completing the first volumes of the *New Catalog*, we were able to conduct preliminary statistical surveys of sign use and have periodically redone these studies, often in collaboration with Gabrielle Vail, Jonathan Scholnick, Jessica Munson, and Yuriy Polyukhovych (Looper et al. 2015; Macri 2021; Macri and Looper 2003; Macri and Vail 2009). This report summarizes some of the conclusions we have reached based the current MHD dataset, including discussion of the relationship of the three hieroglyphic books or codices from the Late Postclassic period to the texts in other media, such as monuments and pottery, which date to earlier periods.<sup>1</sup> The sample used in this

---

<sup>1</sup> Dates for time periods mentioned in this note are: Late Preclassic 300 BCE-250 CE; Early Classic 250-600 CE; Late Classic (600-820 CE); Terminal Classic (820-950 CE); Early Postclassic (950-1200 CE); Late Postclassic (1200-1521 CE). This study includes data from the Dresden, Madrid, and Paris codices, but



study includes all signs that can be identified from datable texts, either through precise dates recorded on them or estimates to the nearest 20-year period, regardless of whether or not the signs have been deciphered. In this study, we define a *grapheme* as the fundamental graphic unit of text representing a specific linguistic value or function, a *k'atun* as a 20-year period of time in the Maya calendar, and *innovation* as the first appearance of a new grapheme in the hieroglyphic record.

**The Dataset**

The numbers of graphemes utilized in this study are:

- 1060 from non-codical texts, 796 of which are unique to the non-codical texts
- 414 from the codices, 150 of which are unique to the codices
- 264 graphemes shared among codical and non-codical texts

The majority of graphemes that appear in the codices are shared with non-codical texts, with only a little over a third of codical graphemes unique to the hieroglyphic books. This contrasts with the graphemes in non-codical texts, of which the majority (about three quarters) are unique, and only about a quarter are shared with the codices. Later in this note, we discuss the historical relationship of codical to non-codical texts, based on grapheme populations. The text dataset is summarized in **Table 1**.

**Table 1.** Text dataset.

source	total texts	glyph blocks*	graphemes coded	graphemes coded, not incl. bar-and-dot numerals
codices	3 books	15332	26985	22393
other texts (all)	4887	69670	136132	125819

\*includes eroded/questioned

The dramatic expansion of the MHD dataset over the last 14 years is illustrated by comparing the non-codical glyph block total in **Table 1** (n=69670) to the total of 36416 blocks reported in an earlier progress report (Macri 2008). These data show that the grapheme total for codices is somewhat higher than expected based on its overall proportion of the corpus. It is possible that several of the codical graphemes that are attested only once or twice are misidentified.

In the analyses below, bar-and-dot numerals are not included. In addition, calculations are based on datable non-codical texts, predominantly monuments, but also some other media, primarily ceramics, dated to an approximate k'atun. Rounding to the nearest k'atun helps to smooth out the data given the large number of monuments that were dedicated on k'atun-endings. The scale of this dataset is shown in **Table 2**.

not the Grolier Codex, which contains only tzolk'in notations and "ring numbers," none of which are expressed with graphemes unique to this manuscript.



Table 2. Datable texts dataset (non-codical).

	texts	glyph blocks	graphemes coded, not incl. bar-and-dot numerals
totals	2235	50772	96327

Figure 1 shows the number of coded graphemes that appear in datable texts plotted over time.

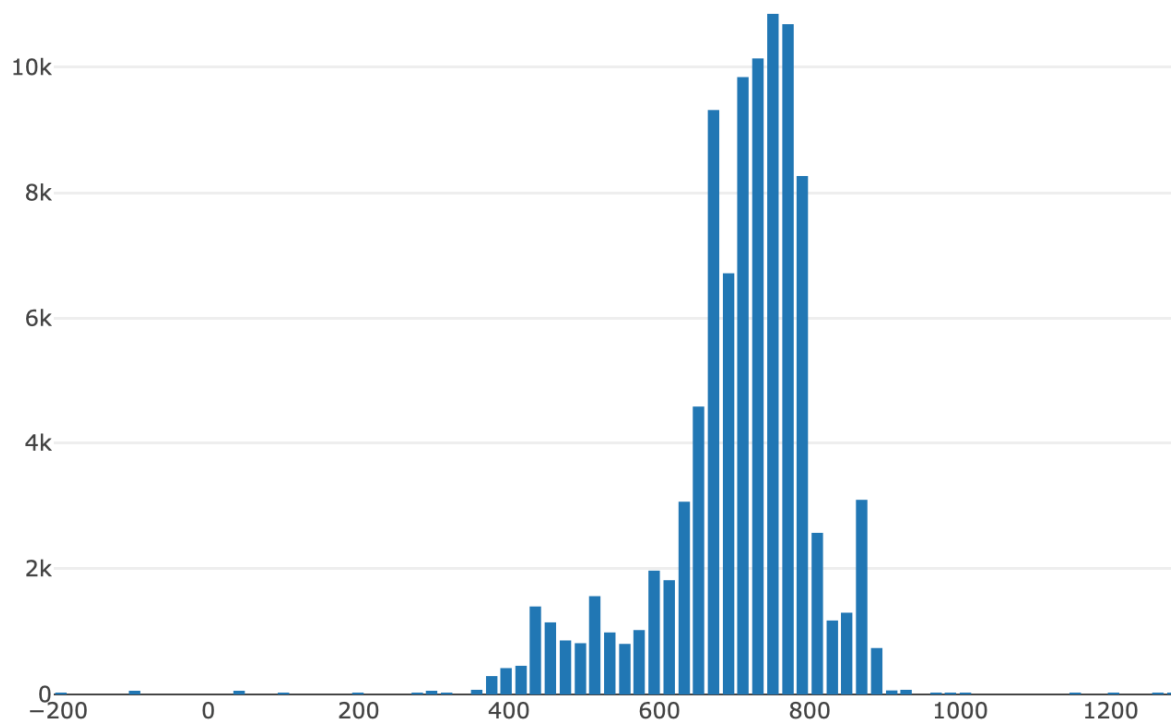


Fig. 1. x=year; y=number of coded graphemes in dated texts.

In sum, a substantial peak of datable text production occurs during the Late Classic period, especially the eighth century (cf. Macri 2008:Fig. 1). This peak would be even more dramatic if currently undated texts, particularly inscribed ceramics, were included.

### Graphemes per K'atun

Since the beginning of modern decipherment of the Maya script, an important question has been the number of unique signs in use at a specific time. This question itself is somewhat vague and indeed impossible to answer, as our dataset does not include complete sign inventories for individual scribes or locales. While irregularities in the sample size and the dating of texts makes it difficult to arrive at a precise number, using the set of dated texts described above, we plotted the number of unique graphemes that appear in these texts during each k'atun (Fig. 2).

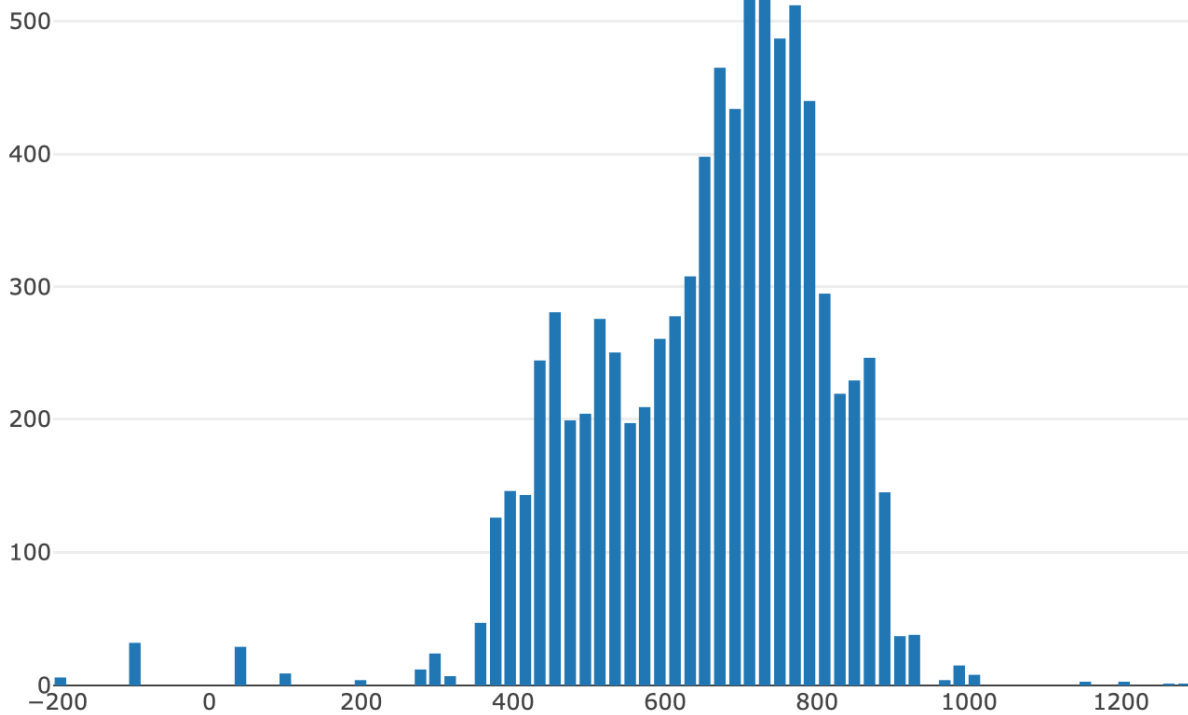


Fig. 2. x=year; y=total unique graphemes in dated texts.

Although we count 1210 graphemes total in the Maya script, only a subset of these is attested as being in use at any given time. The data show that the number of unique graphemes per k'atun peaks at n= about 500 during the eighth century. Adding in undated objects would be expected to increase this total to near 600 graphemes per k'atun. Of course, these figures are likely much larger than the number of unique signs utilized by an individual scribe, as we have considered the Maya area as a whole and grouped grapheme use by k'atun. Indeed, even looking at the data for objects dated precisely to the period ending of 9.17.0.0.0 (771 CE), when the largest number of graphemes are coded, only 321 unique graphemes are attested. It is likely, therefore, that average competency of an individual scribe would be significantly lower than these numbers might suggest.<sup>2</sup>

Presumably, some of the lower grapheme counts for the Early Classic and Terminal Classic can be explained by smaller sample size from these periods. It is likely that the number of graphemes documented for the codices (414 signs) relative to the non-codical texts at their peak is also an artifact of sample size and redundancy of content in these manuscripts.

### Syllabic vs. Logographic Usage

The Maya script is a mixed logosyllabic system, in which logographs represent entire words, while syllabic signs (or syllabographs) represent a consonant + vowel (CV) combination. Many graphemes have known values or are known to be either logographs or syllables. While syllabographs generally only have

<sup>2</sup> Our figures are consistent with Nikolai Grube (1994:178), who estimated the competency of the average scribe to be no more than 250 signs.



a single value, some logographs have multiple values, such as XH3, which is used to represent words for "sky," "snake," "guard", and "four." Such signs are counted as a single logograph in the present study. The following analysis excludes graphemes that are not known to be logographic or syllabic, which explains the differences between Figures 2 and 3. Because of the large number of uncategorized graphemes (about  $\frac{1}{3}$  of the total), estimates about logographic vs. syllabic development should be considered tentative.

Many graphemes in the Maya script have both logographic and syllabic values, depending on context (53 graphemes total; **Table 3**; see also Fox and Justeson 1984). Graphemes that have both logographic and syllabic functions sometimes have completely unrelated logographic and syllabic values, such as ZC1, which can be logographic **TUUN**, the day Kawak, or syllabic /ku/ (**Table 3**, Row 1). Others have a syllabic value that is derived acrophonically from its logographic reading. In these cases, a logographic function is more common and is attested in earlier texts, i.e. a syllabic value is derived at a later date. An example is AP9, which can be logographic **BAH**, **BA'H**, or **BAAH**, as well as syllabic /ba/ (**Table 3**, Row 2). A few signs have multiple logographic values as well as a syllabic value derived from one of these (**Table 3**, Row 3). AA1 ("fish/fin") is a special case, which would normally be considered a syllabograph; however, because it is known to have a value of **CHAY** in some contexts, it cannot simply be considered syllabic /ka/ used to underspell "kay."



Table 3. Examples of graphemes with both logographic and syllabic values. Drawings by Matthew Looer and Martha J. Macri.

class	grapheme
1. signs having unrelated logographic and syllabic values	
2. signs having syllabic values derived acrophonically from the logograph	
3. signs with multiple logographic values, one of which is the source of syllabic value	



Table 4 shows the number of graphemes attested exclusively in non-codical texts, those in the codices, and those shared between the two data sets, distinguishing between logographs, syllables, and signs that have both functions.

Table 4. Graphemes in non-codical texts, codices, and shared.

	logographic	syllabic	both (overall)
signs non-codical only	409	86	15
codical only	36	3	2
shared between codices and non-codical texts	146	69	36
total	591	158	53

Excluding graphemes that function both as logographs and syllables, 78.2% of the Classic inventory is logographic vs. syllabic, which is close to the codical proportion of logographs to syllables (71.7%). Adjusting the tally for the codices based on attested usage of bivalent (logographic/syllabic) graphemes yields 75.3% logographic signs. The similarity of these raw numbers is partly the result of the large number of shared signs between the codices and non-codical texts, relative to graphemes that are unique to the codices. This indicates a strong structural continuity between the two datasets, a topic explored further below.

Figure 3 shows the proportion of logographs vs. syllables used in non-codical texts over time, adjusting for the use of graphemes that can be both logographic and syllabic.



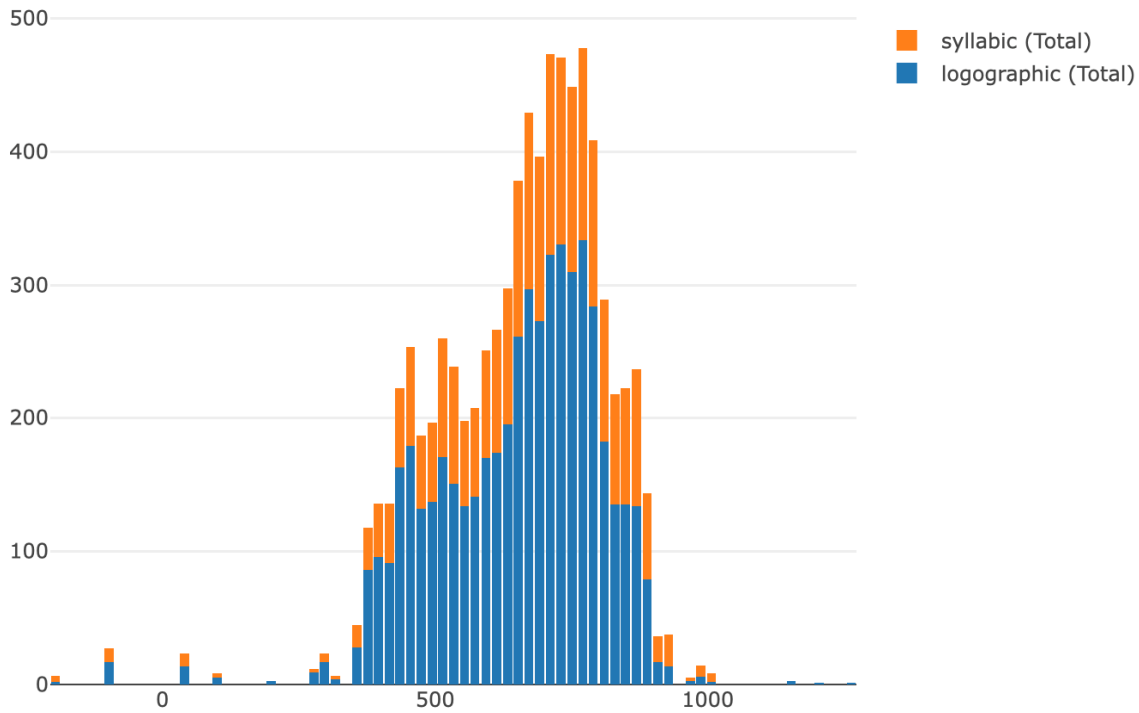


Fig. 3. x=year; y=number of unique logographs and syllables used, including polyvalent signs.

Figure 4 shows the percentage of logographs in non-codical texts over time, removing very early and late examples (burn-in and -out).

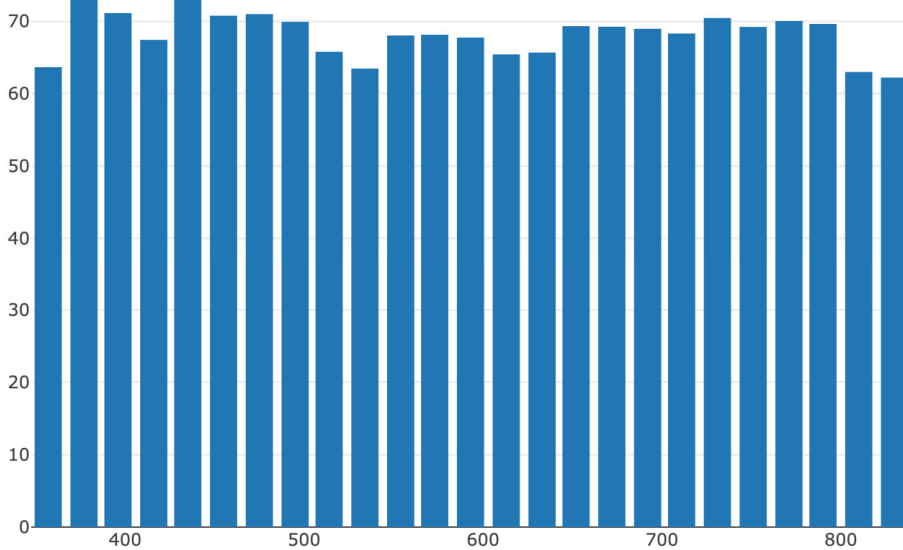


Fig. 4. x=year; y=percentage unique logographic vs. syllabic signs used during the Classic period, including polyvalent signs.





The proportion of logographic vs. syllabic graphemes used in dated texts stays relatively constant at around 70.9% logographic between ca. 300-810 or 68.6% if syllabic and logographic usages of polyvalent graphemes are included.<sup>3</sup> As others have observed, there is an apparent increase in syllabic usage during the ninth century, which is likely related to an increase in the relative proportion of texts from the northern lowlands during this time, which tend to employ more syllabic signs (see Wichmann and Davletshin 2006). Another important point is that the average proportion of logographic signs during the Classic period is also close to the value of 75.3% logographic graphemes in the codices.<sup>4</sup> This seems to be another example of the overall formal consistency in the script over time. However, these conclusions are based on known values for only about two thirds of the signs used in the script and should therefore be considered tentative.

### Sign Innovation and Obsolescence

Figure 5 shows number of new logographic and syllabic signs innovated over time during the Classic period.

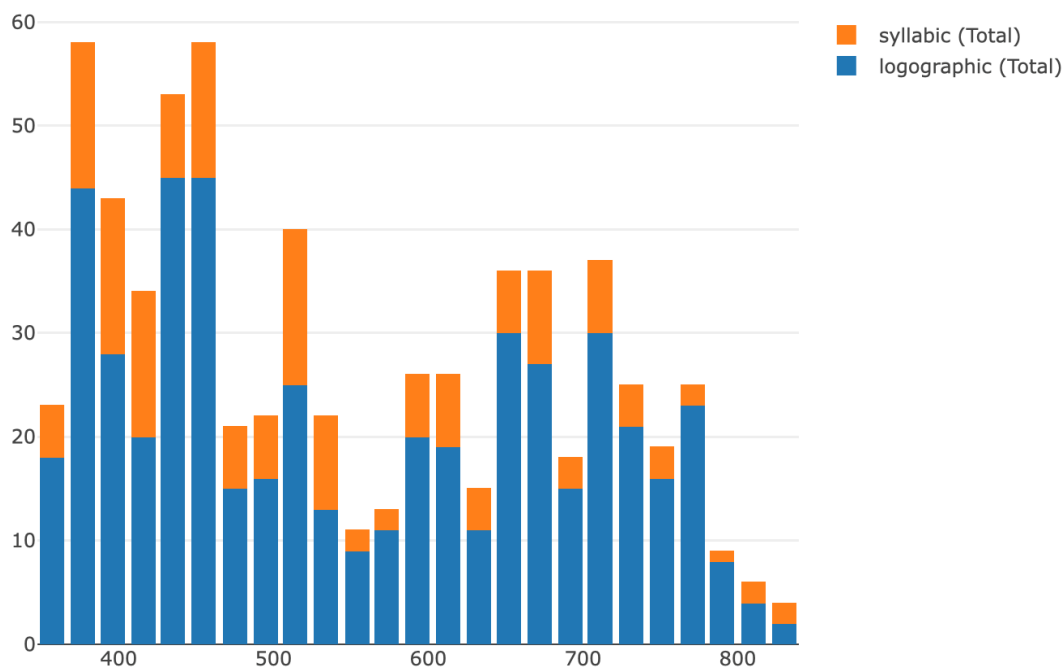
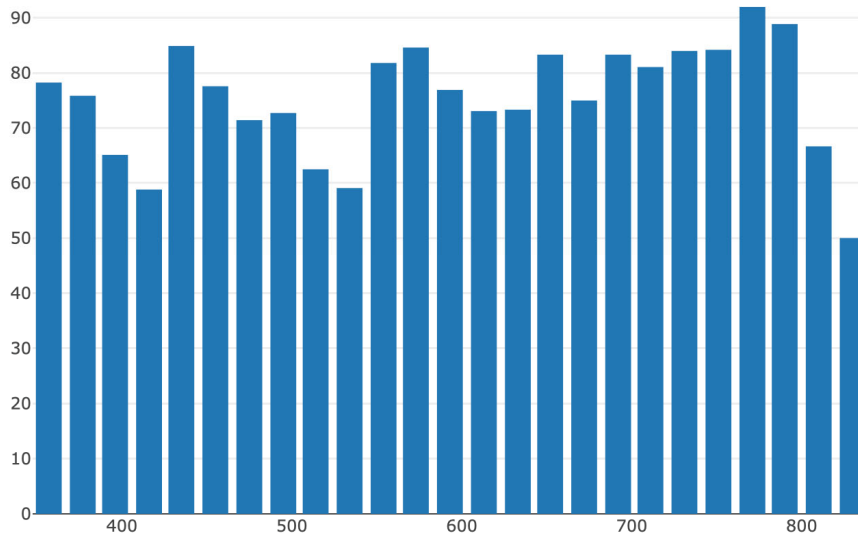


Fig. 5. x=year; y=number of new graphemes appearing during each k'atun during the Classic period, including polyvalent signs used logographically vs. syllabically.

<sup>3</sup> This contrasts with the conclusion of Wichmann and Davletshin (2006), which seems to suggest an overall increasing trend in "phoneticism" (i.e. syllabic vs. logographic usage) over the course of the Classic period. However, one observes that in their study (Wichmann and Davletshin 2006:Fig. 1), the Late Classic period is characterized by relatively consistent percentage of CV signs, which is what we find as well.

<sup>4</sup> Here, we agree with the findings of Wichmann (2006:290) regarding proportions of CV graphemes in the codices compared with non-codical texts.

**Figure 6** shows percentage of new logographs appearing during each k'atun, removing very early and late examples (burn-in and -out):



**Fig. 6.** x=year; y=percentage of logographic graphemes innovated during the Classic period, including polyvalent graphemes used logographically.

**Figure 5** makes it seem as if grapheme innovation peaked twice, once in the Early Classic and again in the Late Classic. However, in a previous analysis, we demonstrated that the Early Classic "peak" of innovation is probably an artifact of small sample size, and the Late Classic innovation shows a linear relationship to sample size (Looper et al. 2015). **Figure 6** shows that the proportion of new logographs to syllables remains fairly constant over time. Again, we see an increase in syllabic innovation after 800, likely owing to a greater proportion of texts from the northern lowlands. However, we also caution that sample size is dramatically smaller during this period.

Not including signs shared between codical and non-codical texts, we calculate the median duration of grapheme usage at only 60 years. This number is lower than expected because more than 45% of non-codical graphemes were used for one k'atun or less. Of the non-codical logographic graphemes, the median duration of usage is 119 years (6 k'atuns). Of the non-codical syllabic graphemes, the median usage is slightly longer, at 193 years (9.8 k'atuns) These figures support the notion of the relative stability of syllabic versus logographic signs (see Grube 1994:179).

### Graphemes in the Codices Compared to Non-codical Texts

**Figure 7** shows the cumulative totals of new graphemes unique to non-codical texts (orange) compared to new graphemes that are shared between codical and non-codical texts (blue):

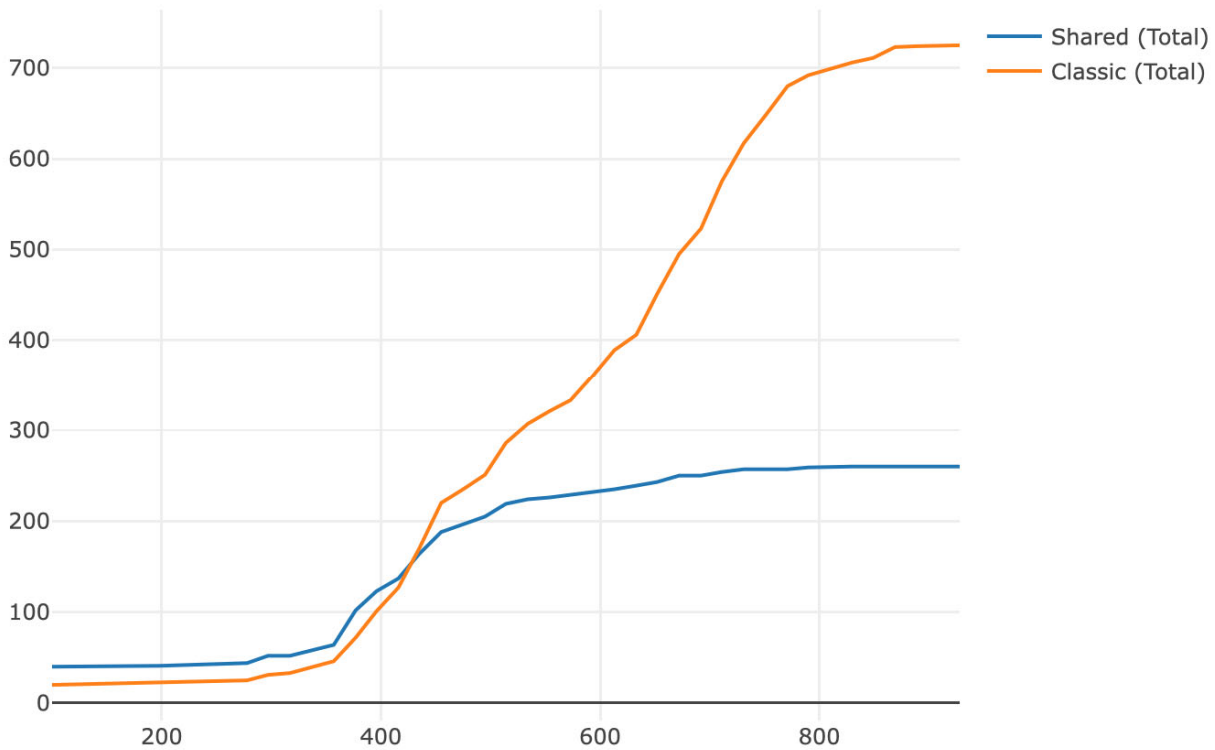


Fig. 7. x=year; y=cumulative total of new non-codical grapheme vs. new graphemes shared between non-codical texts and codices

As can be seen from this graph, texts at the beginning of the Classic period, have more graphemes that are shared with the codices. They far outnumber the graphemes that are unique to non-codical texts. During the fifth century, when more texts are documented, the number of new shared signs starts to decrease. This trend accelerates during the Late Classic period, when very few new signs also appear in the codices. In sum, **Figure 7** shows that most of the signs that appear in both codices and non-codical texts were introduced relatively early.<sup>5</sup> Because data for early times (i.e. Late Preclassic—Early Classic) are relatively sparse, it would be expected that a larger data set would show an even stronger correlation between early script and codices.

Looking at the data more closely, of the 264 graphemes shared between Classic-period texts and the codices, the vast majority are used continuously from their initial date of attestation, typically before the Late Classic, until the decline of the textual record at the end of the Classic period, after around 771 CE. Of the graphemes that are shared between codical and non-codical texts, 94 are used syllabically, while 170 are used logographically. Of the syllabic graphemes, only 6 shared signs were introduced in the Late Classic (6.4%), which contrasts with logographs, of which 20 shared signs appear in the Late Classic (11.8%).

Many of the shared signs that appeared for the first time in the Late Classic period are rare, suggesting that, once again, sample size plays a role. In two significant examples, rare signs appear in an unusual

<sup>5</sup> This confirms the observation of Grube (1994:179) that many signs that appear in the earliest phases of the script remained in use in the codices.

context in the Late Classic period as well as in the codices, possibly signaling a specific continuity in the genre of numerical and calendrical notation. One is 1GC, which appears in the Classic only on the Lunar Table mural and numerical array on the north wall of Xultun Structure 10K2 (731 CE), but is common in the codices (**Fig. 8a**) (Saturno et al. 2012). Because it is used in both contexts as 'zero', the Classic heritage of the codical grapheme seems secure. The unusual context in wall paintings that represent mathematical tables suggests a special association with painted books. An analogous argument may be made in the case of the "ring" grapheme ZB4, which is used to indicate counts of days both in the Xultun Structure 10K2 mural and the codices (**Fig. 8b**).



**Fig. 8.** Comparison of non-codical and codical signs used in numerical/calendrical notation: **a.** 1GC, **b.** ZB4. Drawings by Matthew Looper and Martha J. Macri.

## Conclusions

In summary, we offer the following general observations about grapheme distribution and innovation in the Maya script:

1. *Sample size* is essential in evaluating grapheme use and innovation. Conclusions are most reliable for the eighth century, when data are most abundant. In future studies, it will be important to add more texts to the sample, particularly the large corpus of inscribed ceramics and other portable objects that could be assigned estimated dates. Adding spatial controls to the analysis of grapheme development and usage may also reveal significant patterns.
2. At the end of the Late Classic period, the *number of unique graphemes* in use is just over 500 per k'atun, though the actual number of signs being known to a hypothetical scribe at this time would likely be considerably less. In the codices, which are the work of several different scribes, 414 unique signs are documented.
3. The *proportion of logographs to syllabographs* in the script remains remarkably stable over time, even though signs of both classes are being constantly cycled in and out of the script. Similar proportions of logographs to syllabographs appear in the codices compared to the non-codical texts. These conclusions are tentative, since undeciphered graphemes are not included in this survey.
4. For the Late Classic, the *grapheme innovation* rate correlates to sample size. The proportion of logographs to syllabographs innovated over time is stable, with an apparent increase in syllabic signs created after 800 CE. Syllabic signs show slightly greater permanence overall compared to logographs.
5. Many *codical graphemes* first appear during the early phase of the Maya script, prior to the Late Classic. Of the shared signs between codical and non-codical texts, a greater proportion of syllabic signs appears before the Late Classic period compared to logographs. Two specific graphemes are known only



from a Late Classic astronomical mural and from the codices, suggesting that in some cases, genre or content of a text may have been significant in determining sign usage.

## References

Fox, James A., and John S. Justeson

1984 Polyvalence in Mayan Hieroglyphic Writing. In *Phoneticism in Mayan Hieroglyphic Writing*. John S. Justeson and Lyle Campbell, eds. Pp. 17–76. Albany: Institute for Mesoamerican Studies, University of New York at Albany.

Grube, Nikolai

1994 Observations on the History of Maya Hieroglyphic Writing. In *Seventh Palenque Round Table, 1989*, Palenque Round Table Series. Virginia M. Fields, ed. Pp. 177–186. San Francisco: Pre-Columbian Art Research Institute.

Looper, Matthew, Jonathan Scholnick, Jessica Munson, Yuriy Polyukhovych, and Martha J. Macri

2015 Measuring Grapheme Innovation in Classic Maya Writing. Paper presented at the Society for American Archaeology 80th Annual Meeting, San Francisco, Calif.

Macri, Martha J.

2008 Dated Monuments in the Hieroglyphic Database. *Glyph Dwellers*, Report 26. <http://glyphdwellers.com/pdf/R26.pdf>.

2021 Prepositions *Ti* and *Ta* in Classic Maya Monument Texts. *Glyph Dwellers*, Report 69. <http://glyphdwellers.com/pdf/R69.pdf>.

Macri, Martha J., and Matthew G. Looper

2003 *The New Catalog of Maya Hieroglyphs, Volume One: The Classic Period Inscriptions*. Norman: University of Oklahoma Press.

Macri, Martha J., and Gabrielle Vail

2009 *The New Catalog of Maya Hieroglyphs, Volume 2: The Codical Texts*. Norman: University of Oklahoma Press.

Saturno, William A., David Stuart, Anthony Aveni, and Franco Rossi

2012 Ancient Maya Astronomy from Xultun, Guatemala. *Science* 336: 714–717.

Wichmann, Søren

2006 Mayan Historical Linguistics and Epigraphy: A New Synthesis. *Annual Review of Anthropology* 35:279-294.

Wichmann, Søren and Albert Davletshin

2006 Writing with an Accent: Phonology as a Marker of Ethnic Identity. In *Maya Ethnicity: The Construction of Ethnic Identity from the Preclassic to Modern Times, 9th European Maya Conference, Bonn, December 2004*, Frauke Sachse, ed. Pp. 99-106. Markt Schwaben: Verlag Anton Saurwein.



Glyph Dwellers is an occasional publication of the Maya Hieroglyphic Database Project at California State University, Chico, California. Its purpose is to make available recent discoveries about ancient Maya culture, history, iconography, and Mayan historical linguistics deriving from the project.

Funding for the Maya Hieroglyphic Database Project is provided by the National Endowment for the Humanities, grants #RT21365-92, RT21608-94, PA22844-96, the National Science Foundation, grants #SBR9710961 and IBSS1328928, the Department of Native American Studies, University of California, Davis, and the Department of Art and Art History, California State University, Chico.

© 2022 Matthew G. Loper. All rights reserved. Written material and artwork appearing in these reports may not be republished or duplicated for profit. Citation of more than one paragraph requires written permission of the publisher. No copies of this work may be distributed electronically, in whole or in part, without express written permission from the publisher.

ISSN 1097-3737