Ancient Mesoamerica (2024), 1–26
doi:10.1017/S0956536123000317

CAMBRIDGE UNIVERSITY PRESS

Research Article

The duplication diacritic: A case study of variation and change in Mayan writing

David F. Mora-Marín
Department of Linguistics, University of North Carolina, 307 Smith Building, CB #3155, Chapel Hill, NC 27599, USA

Abstract

This article studies the duplication diacritic of Epigraphic Mayan (ISO 639-3 emy) during the Classic period (A.D. 200–900). Cataloged as grapheme 22A, it consists of two dots optionally and rarely affixed to another grapheme to command the reader, in the majority of cases, to read a syllabogram twice in sequence. This article reviews prior literature on the diacritic, elaborates a typology of four distinct but ultimately related functions, and employs a data set compiled by means of the Maya Hieroglyphic Database to determine via statistical tests whether scriptal, linguistic, media, geographic, and temporal factors were influential in its distribution, and more narrowly, its various functions. The results indicate that two lexemes, *käkäw* 'cacao' and *k'ahk'* 'fire,' account for several of the scriptal and linguistic traits that show significant relationships with 22A, with the former, *käkäw*, likely serving as a major prototype in the evolution of 22A. It is also pointed out that 22A is absent from the Postclassic (A.D. 900–1521) codices, suggesting that one of the Classic regional subtraditions with lowest frequency of use of 22A may have been a direct ancestor of the subtraditions responsible for the codices.

Resumen

Este artículo estudia el diacrítico de duplicación del maya epigráfico (ISO 639-3 emy) durante el período Clásico (200–900 d.C.). Catalogado como el grafema 22A, consta de dos puntos añadidos opcional e infrecuentemente a otro grafema para comandar, en la mayoría de los casos, que el lector lea un silabograma dos veces seguidas. Este artículo revisa la literatura previa sobre el diacrítico, elabora una tipología de cuatro funciones distintas pero en última instancia relacionadas, y emplea un conjunto de datos recopilados por medio de la Maya Hieroglyphic Database (MHD) o Base de Datos de Jeroglíficos Mayas por Looper y Macri (1991–2023), para determinar, por medio de pruebas estadísticas, si los factores escriturarios, lingüísticos, mediáticos, geográficos y temporales influyeron en su distribución y, más específicamente, en sus diversas funciones. Tales funciones incluyen la duplicación secuencial (Stuart 2014), la duplicación no-secuencial (Stuart 2014; Zender 1999), la marcación de logogramas verbales con raíces C_1VC_1 (Kettunen and Helmke 2020; Prager 2020) y la abreviación colocacional (supra-grafémica) (Mora-Marín 2022b). Los resultados indican que dos lexemas, *käkäw* 'cacao' y *k'ahk'* 'fuego', explican varios de los rasgos lingüísticos y escriturarios que muestran relaciones significativas con 22A, y que el primero, *käkäw*, probablemente sirvió como un prototipo importante en la evolución de 22A. También se señala que 22A está ausente de los códices del Posclásico (900–1521 d.C.), lo que sugiere que por lo menos una de las subtradiciones regionales del Clásico con menor frecuencia de uso de 22A (norte, sur, Pasión) puede haber sido un antepasado directo de las subtradiciones responsables de los códices.

Keywords: Mayan epigraphy; Mayan linguistics; Variationist approach; Duplication diacritic

This article pertains to the duplication diacritic, or "doubler," of Epigraphic Mayan (ISO 639-3 emy), first identified by Stuart (2014) (originally circulated in 1990). Epigraphic Mayan—the written language of Lowland Mayan society of Mexico, Guatemala, Belize, Honduras, and El Salvador—was in use between circa 300 B.C. and the late seventeenth century. It represented varieties of the Ch'olan and Yucatecan subgroups of the Mayan language family, with Ch'olan serving as the basis of the script itself (e.g., Bricker 1986; Gronemeyer 2014; Houston et al. 2000; Justeson and Fox 1989; Justeson et al. 1985; Josserand and Hopkins 2002; Lacadena and Wichmann 2002; Mora-Marín 2009). The script was characterized by three major types of distinctive signs or graphemes: logograms, which represent lexemes (e.g., Mora-Marín 2005, 2010); syllabograms, which represent CV or CVC sequences (C = consonant, V = vowel); and diacritics, a term first applied by Zender (1999) to the case of Mayan writing, referring to a grapheme that is juxtaposed or graphically affixed to another grapheme to indicate a deviation from its unmarked value or unmarked application.

Corresponding author: David F. Mora-Marín, Email: davidmm@unc.edu
Cite this article: Mora-Marín, David F. (2024) The duplication diacritic: A case study of variation and change in Mayan writing. *Ancient Mesoamerica* 1–26. https://doi.org/10.1017/S0956536123000317

The grapheme in question, cataloged as 22A (Looper et al. 2022), was a diacritic in the sense just defined. Its most frequent function is to tell the reader to read a syllabogram twice (Stuart 2014) and therefore constitutes a deviation from its unmarked application (a single reading). Zender (1999) conducted a thorough review of the literature on 22A up until the late 1990s and systematically evaluated the relevant claims and hypotheses, especially those outlined by Stuart and Houston (1994), which suggested that 22A could also be applied to logograms. Zender's (1999) conclusion was that 22A applied only to syllabograms, not to logograms, and that its function was to duplicate the reading of the syllabogram to which it was affixed. Nevertheless, recent work by Kettunen and Helmke (2020:20) and Prager (2020:3) has called attention to instances in which 22A is affixed to logograms with a C_1VC_1 shape. Even earlier, Stuart and Houston (1994) had made reference to a possible collocation-distinguishing function, in which 22A would be used to distinguish between similar collocations of signs spelling different terms (e.g., *chik'in* 'west' from *k'ihnich* 'heated [Sun God]' both based on *k'in 'sun', and both potentially employing the logogram **K'IN** and the syllabogram **chi**—albeit usually in different reading orders) or different values of a polyvalent logogram (e.g., the grapheme ZBB for **B'UTS'** for *b'utz'* 'smoke' and **K'AK'** for *k'ahk'* 'fire', but see disclaimer below about the former value). And more recently, Mora-Marín (2022b) has argued that in a few instances, 22A functioned as a punctuation marker, indicating the abbreviation of a common collocation.

With this background in mind, one of the goals of this article is to utilize the Maya Hieroglyphic Database (MHD) by Looper and Macri (1991–2023), a comprehensive digital corpus of 85,565 records consisting of close to 5,000 individual texts and spanning the entire history of the script, which became accessible online in early 2022, to conduct a comprehensive examination of all the evidence for the proposed functions of the 22A diacritic. The second goal of this article, made possible by the MHD, is to conduct a quantitative study of 22A aimed at understanding its historical development and the factors—scriptal, linguistic, temporal, geographic—that may have influenced such development.

I begin with a review of the literature to evaluate the validity of the proposed functions of 22A and to elaborate a classification of such functions. Then, I introduce the methods applied in this study, including the definition of the relevant scriptal and linguistic categories, the definition of the relevant geographic and temporal categories, the procedures for preparing the data set, and the types of statistical tests used to analyze the data set. I continue with a presentation of the quantitative results in three parts: (1) the results relevant to time as an independent variable and its possible influence on the uses and functions of 22A; (2) the results relevant to the possible influence of scriptal, linguistic, and geographic factors on the use of 22A; and (3) an examination of the lexical distribution of 22A in the corpus. The quantitative study, serving the dual role of a hypothesis-forming and a hypothesis-testing tool, will allow for the formulation of a model outlining the key factors influencing the evolution of 22A during the Classic period. The article then summarizes the major findings and offers a set of conclusions and desiderata for further research.

More specifically, the article proposes that 22A started on portable media by the beginning of the Early Classic period and that its early evolution (i.e., its set of preferred graphemic and phonological targets) was strongly influenced by its use in the spellings of proto-Ch'olan *käkäw 'cacao', which served as a scriptal and linguistic prototype for scribes. This does not mean that 22A was originally applied to the spellings of 'cacao'. It only means that this root quickly became its primary lexical target and influenced the scribes' application of 22A from then on. Once accepted on more formal monumental media, which became its preferred medium during the Late Classic period, 22A was frequently applied to the lexeme *k'ahk' 'fire', which accounts on its own for some of the associations that characterize 22A during that time. Interestingly, 22A is completely absent from the Postclassic codices. Two factors are posited to account for this: (1) perhaps the Postclassic codex tradition was a direct descendant of one of the geographic subtraditions (Northern, Southern, Pasion) where 22A was rarer than normal; and (2) perhaps the dramatic decline in text production at the end of the Classic period led to a kind of bottleneck effect, through which only common scribal practices survived, and rare practices such as the use of 22A simply did not.

Background and classification

Basic graphemes in Mayan writing

I am following the definition of grapheme as "the minimally contrastive unit in a writing system" (Henderson 1985:135), and I occasionally use it interchangeably with the term sign in the Peircean sense, although I use these terms to refer not just to a unit of form and meaning (e.g., a logogram) but more generally to a unit of form and value (e.g., logograms or syllabograms) or of form and function (e.g., diacritics). In this sense, Mayan writing employs three major types of graphemes: logograms, syllabograms, and diacritics. Following Fox and Justeson (1984), I render logograms in uppercase bold letters, syllabograms in lowercase bold letters, and following more recent conventions, the duplication diacritic as a bold superscript <2>. The two most important types are logograms and syllabograms, which could be utilized in three major types of spellings. These can be illustrated with the term *ʔunen* 'baby; child', or its possessed form *y-unen* 'his/her/its baby': (1) logographic, when only a logogram was used, as in **ʔUNEN** (Figure 1a); (2) syllabographic (or simply syllabic), when only syllabograms were used, as in **yu-ne** (Figure 1b) or **ʔu-ne** (Figure 1c); and (3) logosyllabic, when logograms and syllabograms were combined, as in **ʔu-ʔUNEN-ne** (Figure 1d) or **ʔUNEN-ne** (Figure 1e).

I am following Zender (1999:101–102) in his extension of the term *diacritic* to apply to nonalphabetic scripts, and in the case of Mayan, to both "semantic determinatives"

https://doi.org/10.1017/S0956536123000317 Published online by Cambridge University Press

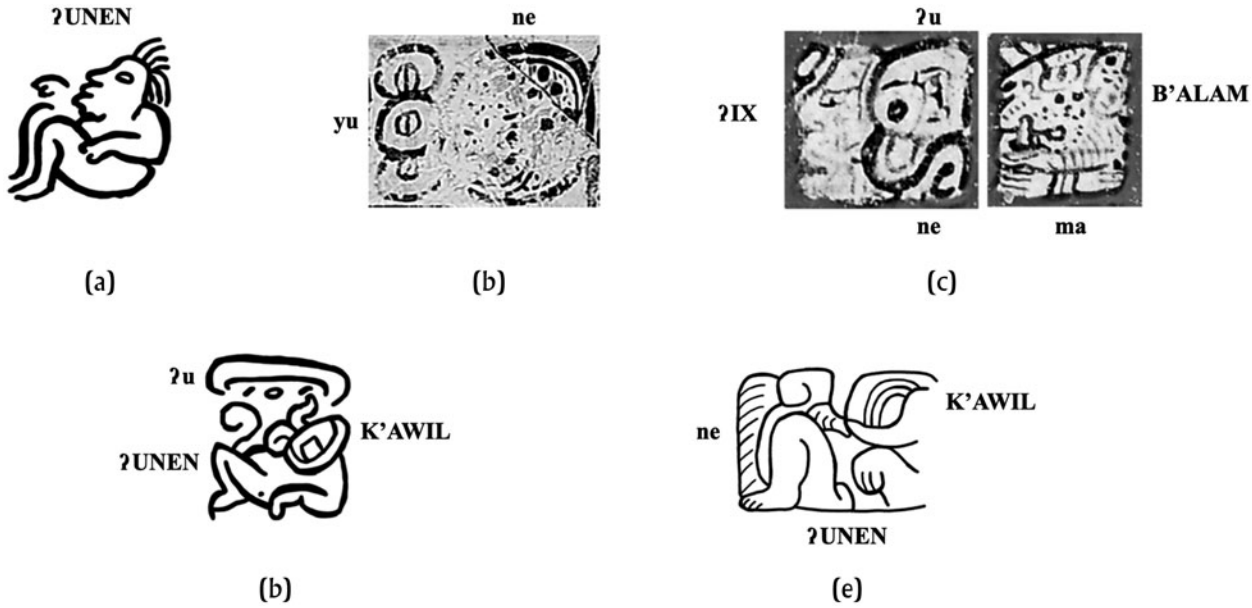

Figure 1. Types of spellings in Epigraphic Mayan. (a) Logographic spelling **ʔUNEN** for *ʔunen* 'baby; child', collocation U01 on Uaxactun Structure B13 Mural (UAXB13Mu). Drawing by author (after drawing by Antonio Tejeda in Smith 1950:Figure 47). (b) Syllabographic spelling **yu-ne** for *y-unen* 'his child', collocation Q on Vase 1986.1080 at the Art Institute of Chicago (COLK0635). Photograph by Justin Kerr from Maya Vase Database at http://research.mayavase.com/kerrmaya.html. (c) Syllabographic spelling **ʔIX-ʔu-ne B'ALAM** for *ʔix-ʔunen b'ahläm* 'Lady Baby Jaguar', collocation B'01 on Vase of the Eleven Gods at the Kimbell Art Museum (COLK7750). Photograph by Justin Kerr from Maya Vase Database at http://research.mayavase.com/kerrmaya.html. (d) Logosyllabic spelling **ʔu-ʔUNEN:K'AWIL-ne** for *ʔunen k'awil* 'baby k'awil', collocation B05 on Comalcalco Urn 26 Pendant 08 (CMLU26Sp08). Drawing by author (after drawing by Marc Zender in Martin et al. 2002:II-51). (e) Logosyllabic spelling **ʔUNEN:K'AWIL-ne** for *ʔunen k'awil* 'baby k'awil', collocation C02c on Palenque House C Hieroglyphic Stairway (PALHCHS). Drawing by author (after drawing by Yuriy Polyukhovych in Sánchez Gamboa et al. 2020:7–11, Figure 5).

(referred to as lexical determinatives by Mora-Marín [2023]) and the "doubler." Here, I refine Zender's definition, extending its usage from one in which a diacritic points to a departure of a grapheme's unmarked phonetic or lexical value (e.g., syllabographic, logographic) to one in which it may *also* point to a departure of its unmarked scope of application (e.g., a single reading versus two readings within a spelling). Determinatives do not point to general lexical or semantic domains, but instead, to specific lexemes (Mora-Marín 2023). Some scripts (e.g., Sumerian, Egyptian, Chinese), in addition to lexical determinatives, also employ semantic classifiers, which point to general lexical or semantic domains. Mayan does not have a class of semantic classifiers proper. The closest to such a category are the classifiers identified by Hopkins (1994) and Hopkins and Josserand (1999), also discussed by Mora-Marín (2008), but which the present author now considers to be iconographic classifiers that function within the artistic subsystem of the script rather than the graphematic subsystem that interfaces with the spoken language.

The third type, diacritics, were graphemes that were "affixed" to a logogram or syllabogram to indicate a deviation from its default or unmarked value, function, or scope of application within a spelling. One such sign has been known for quite some time: the so-called semantic determinative (Hopkins 1994; Hopkins and Josserand 1999; Justeson 1986:447–449; Kelley 1976:206–211; Mora-Marín 2008; Schele 1983:19–21); more recently, Mora-Marín (2022a, 2023) has redefined such signs as *lexical determinatives*, for they point to specific lexemes rather than more general semantic classes. These are equivalent to Gelb's (1963:105) notion of "specific determinatives," rather than his notion of "determinatives" or "semantic indicators" (1963:252) (also known as "classifiers"), which point to a general semantic class. Although some writing systems bear both types (e.g., Sumerian, Egyptian, Chinese), others exhibit only one type. Mayan may bear only the first type, lexical determinatives: what Hopkins (1994), Hopkins and Josserand (1999), and Mora-Marín (2008) have called "semantic classifiers" in the past, Mora-Marín (2022a, 2023) has more recently redefined as "iconographic classifiers" and suggested that they bear no relevance to the script/language interface.

Figure 2 illustrates examples of lexical determinatives. First, Figure 2a illustrates the use of T533 **ʔAJAW** for *ʔajaw* 'lord, ruler' used as a lexical determinative in conjunction with T670 to determine the value **CH'AM** for *ch'äm* 'to hold/receive', whereas Figure 2b shows that when combined with the SPIRAL diacritic, T670 bore the value **YAL** (or **ʔAL**) for *y-al* 'her child' instead. Figures 2c–d illustrate a different example, T713, which, unmarked, conveys the value **K'AB'** for *k'ab'* 'hand, arm' (Figure 2c), but when combined with T617 as a lexical determinative, it bears the value **K'AL** for *k'al* 'to wrap, close' (Figure 2d). In these cases, the signs T533, SPIRAL (possibly same as ZRJ), and T617 function as diacritics called lexical determinatives because, when combined with another sign, they determine that other sign's value.

https://doi.org/10.1017/S0956536123000317 Published online by Cambridge University Press

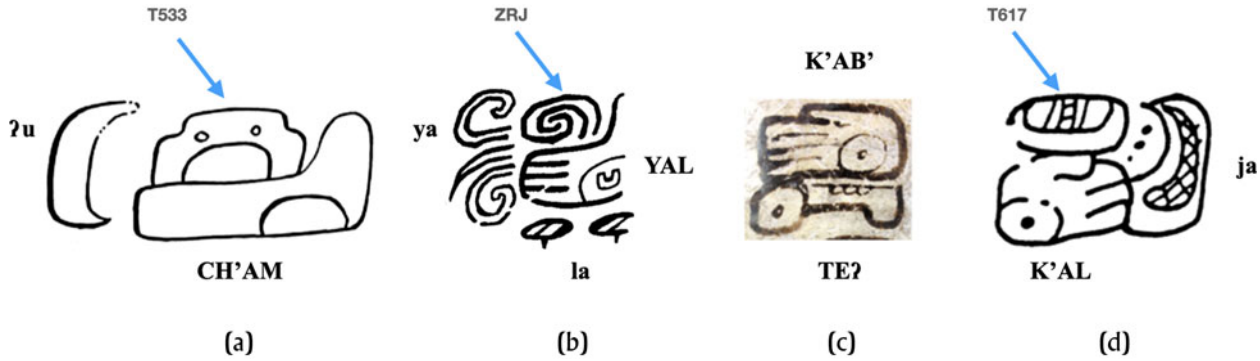

Figure 2. Examples of lexical determinatives. (a) Example of T670 with T533/ZA1a in the **ʔu-** ^{T533}**CH'AM** collocation for *u-ch'äm(-aw)-Ø* 's/he received/took it' on incised conch shell trumpet. Drawing by author. (b) Example of T670 with SPIRAL in the **ya-**^{SPIRAL}**(Y)AL-la** collocation for *y-al* 'her child' on jade belt plaque at Museo del Jade, San José, Costa Rica. Drawing by author. (c) Example of T713 as **K'AB'** in **K'AB'-TEʔ** expression for *(u-)k'ab' teʔ* 'tree branch (hand/arm)' on Vessel K7149 (COLK7149). Photograph by Justin Kerr from Maya Vase Database at http://research.mayavase.com/kerrmaya.html. (d) ^{T617}**K'AL-ja** collocation for *k'a[h]l-aj-Ø-Ø* 'it was wrapped/closed' on inscribed bone (COLDMA129). Drawing #7320 from the Linda Schele Drawings Collection at http://research.famsi.org/schele.html.

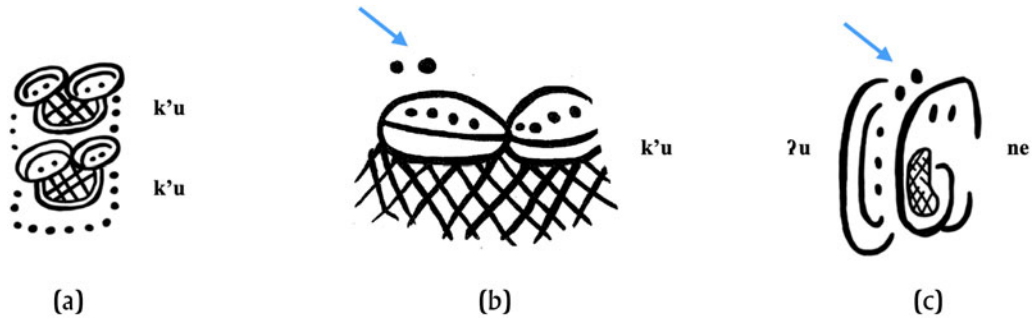

Figure 3. Examples of use of 22A. (a) Syllabographic spelling **k'u-k'u** for *k'uk'* 'quetzal' (NARSt32, Z04). Drawing by author after detail of drawing by Ian Graham (1978:86). (b) Syllabographic spelling ^{2}**k'u** for *k'uk'* 'quetzal' (COLK1874, A04). Drawing by the author (after photograph provided by Justin Kerr). (c) Syllabographic spelling **ʔu-**^{2}**ne** for *ʔunen* 'baby; child' (CMLSpn02, A13). Drawing by author (after drawing by Zender 2004:Figure 70).

The other major type of a diacritic—and the focus of this article—is 22A, first identified by Stuart (2014), who noted that the TWO.DOTS sign, typically used to represent the numeral 'two', proto-Ch'olan *chaʔ=, appeared to function as a "doubler": it allowed scribes to "indicate that a sign is to be read twice without having to render a doubled element" (Stuart 2014:1). In fact, a few of the Early Classic examples do not exhibit the common "affix" or "diacritic" look, but instead resemble a common spelling of the numeral 'two' (e.g., MHD abbreviations: COLSGBc173:B, YAXLnt22:A06, COLPoPan:D04 and E02). One of Stuart's (2014) examples was the spelling for the term *k'uk'* for 'quetzal', which could be spelled with two syllabograms: **k'u-k'u**, as in in Figure 3a, or as ^{2}**k'u**, with 22A affixed to a single syllabogram **k'u**, as in Figure 3b, with the superscripted "2" indicating the presence of the duplication diacritic in epigraphic transcriptions. It can also be seen in Figure 3c with a spelling **ʔu-ne**^{2} for *ʔunen* 'baby; child'. This proposal is universally accepted, and it has proven to be a crucial piece of evidence in the decipherment of various expressions (e.g., Stuart 2001).

The use of the 22A diacritic, as observed by prior authors (e.g. Zender 1999), was both *rare and optional*. Figure 4 provides three different types of spellings of *käkäw 'cacao': the first (Figure 4a) shows the use of 22A, rendering ^{2}**ka-wa**; the second shows only **ka-wa** (Figure 4b), with no need for the diacritic; and the third shows a full syllabic spelling **ka-ka-wa** (Figure 4c). Using the MHD, it is possible to study the variety of contexts of use of 22A in a systematic and comprehensive fashion. The MHD (as of August 2022) contains 28 examples of ^{2}**ka-wa**, approximately 233 examples of **ka-wa**, and 35 examples of **ka-ka-wa**. Consequently, out of 261 examples that were the ideal target for the duplication diacritic (i.e., those with only one instance of the syllabogram **ka**), only 10.7 percent employed it. Some of the potential targets of 22A (other morphemes or lexemes containing a C_1VC_1 sequence) employed 22A even less frequently than *käkäw*.

Prior literature on 22A

Following Stuart's (2014) notes, distributed to other scholars after December 7 of 1990 and mentioned at the *Texas Maya Meetings* in March of 1991 by Schele (1991) (cf. Schele and Wanyerka 1991:97), the earliest didactic materials mentioning this convention consist of Harris (1993:ix) and Schele and Grube (1994:19). Harris (1993) was a supplement to

https://doi.org/10.1017/S0956536123000317 Published online by Cambridge University Press

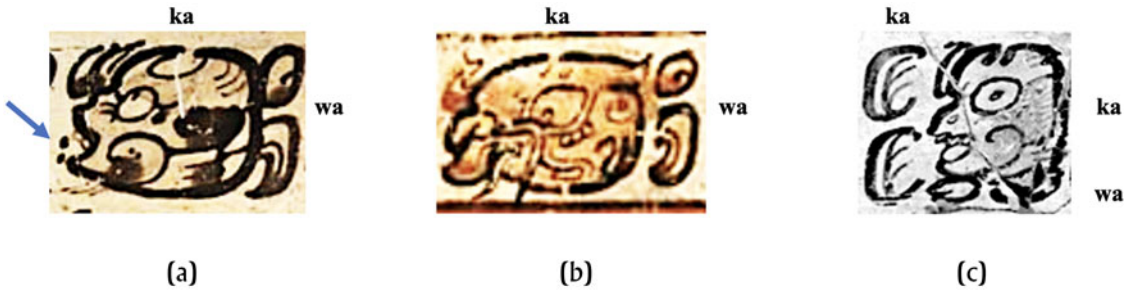

Figure 4. The duplication diacritic in spellings of *käkäw 'cacao'. (a) Spelling **[2]ka-wa** using 22A on K3230. (b) Spelling **ka-wa** without 22A on K554. (c) Spelling **ka-ka-wa** on K1837. All examples are details of photographs by Justin Kerr from the Maya Vase Database at http://research.mayavase.com/kerrmaya.html.

Harris and Stearns (1992). By the time Harris and Stearns (1997) came out—a revised edition of Harris and Stearns (1992)—its authors had already incorporated a basic description (1997:35) and a few examples (1997:Figures 3:13 and 3:16) of the duplication diacritic adopted mainly from Stuart (2014) or Stuart and Houston (1994).

Stuart's (2014) seminal identification of 22A reviewed 12 sets of examples. One of his examples (2014:3) was pointed out to him by Stephen Houston: an instance in which 22A was applied to the syllabogram **ʔu** to indicate that it had to be read twice, once with each of the two nouns present in the glyph block, in order to mark them for possession with *u-* 'third-person singular ergative/possessive'. Stuart took this example to be a case of 22A applying to a logogram, based on the common assumption at the time (e.g., Fox and Justeson 1984; Justeson 1989)—one held by some to this day (e.g., Houston et al. 2001)—that syllabograms used to spell grammatical morphemes behaved logographically. Indeed, Stuart (2014:2) did not see this case as "indicating a doubling of a syllabic element," but instead, as a case indicating that "each of the logographic mainsigns carry the prefix *u*"—that is, the third-person singular agreement marker. In other words, he considered the syllabogram **ʔu** (his **u**) to function logographically (**U**) in the spelling of *u-*. Here, as in previous work by Stuart (1987) and others (Mora-Marín 2005, 2010, 2019, 2021, 2022c), I assume that grammatical morphemes are spelled phonographically by means of syllabograms.

Finally, Stuart innovated the already introduced convention of representing 22A by means of a superscript "2." Henceforth, I apply the superscript to the left of the intended target at the beginning of a collocation, and to the right when it is at the end of the collocation.

Stuart and Houston (1994:46, 49, Figures 56–57) soon followed with a single paragraph and a footnote describing what they knew or suspected about 22A at the time. Both the observations and illustrations were largely based on Stuart's (2014) notes. They suggested that 22A was generally rare to begin with. They also proposed that it could be applied to both syllabograms and logograms to repeat their reading. With regard to the latter case (i.e., logograms), like Stuart (2014:3), these authors were, for the most part, referring to cases of syllabograms used to spell grammatical morphemes. They also observed that 22A tended to be placed on the upper left or upper right of the intended target—what I will refer to in this article as the locus of 22A. Although they do not discuss its discovery and decipherment, they do comment on cases where its function is unclear (1994:46, n13): "We do not yet understand why some spellings use this convention. Perhaps they signal a particular spelling when two are possible: **chi-K'IN** in place of **K'IN-chi**, or **k'a-k'a** instead of **BUTS'**, respectively."

Unfortunately, those authors do not provide relevant illustrations or otherwise reference specific inscriptions. The cases of **chi-K'IN** for Lowland Mayan *chik'in* 'west' and **K'IN-chi** for Yucatecan *k'ihnich/k'iinich* 'Sun God' would be cases of logosyllabic collocations that could be potentially ambiguous, especially given the nonsequential sign orders that were allowed in Mayan writing. Given these facts, **chi-K'IN** for *chik'in* could have been confused with **K'IN-chi** for *k'ihnich*. Stuart and Houston (1994) did not, regrettably, illustrate which of the two spellings would take the diacritic to disambiguate the possible readings. (Below, I suggest the most likely example they were alluding to, and I offer an alternative explanation.) Similarly, regarding the possible ambiguity between T122/ZBBa **K'AK'** for proto-Ch'olan *k'ahk' 'fire', on the one hand, and T122/ZBBa **B'UTZ'** for proto-Ch'olan *b'utz' 'smoke', on the other, the authors did illustrate which spelling would take 22A for the purposes of disambiguation. More importantly, it is no longer generally accepted that T122/ZBBa could be read both as **K'AK'** and as **B'UTZ'**; it is only read as **K'AK'**.

Zender (1999:102–130) provided a more detailed account of the various contexts and uses of 22A than had been attempted up until that time—or for that matter, since. He set out to test four hypotheses (or observations) derived primarily from (Stuart and Houston 1994): (1) that 22A indicates the necessary repetition of a syllabogram; (2) that a logogram can also be "doubled in this manner"; (3) that the loci for placement of 22A were on the upper left (most commonly) and upper right (less commonly); and (4) that 22A was in general "relatively rare." Zender supported (1) and (4). He also rejected (2), concluding that its "canonical function is to double syllables" (1999:118). With regard to (3), Zender also noted cases where 22A is applied in front (i.e., to the left) or below the target grapheme, and more importantly, he showed that it need not be juxtaposed

https://doi.org/10.1017/S0956536123000317 Published online by Cambridge University Press

to the target sign. It could be "attached" to a logogram or to a syllabogram it was not meant to repeat, as long as it was within the same glyph block as the intended target.

Following in Stuart's (2014:3) footsteps, Zender (1999:124–126) also discussed cases where the diacritic was used in the spelling of condensed couplets, expressions that require unfurling, and in such cases, the syllabogram marked with the diacritic would be read twice, in "non-serial" fashion, at the beginning of each paired phrase. Such cases constitute a different type of repetition than the more common cases, and they likely represent an extension of the original function of sequential repetition to a novel function of nonsequential repetition. Below, I distinguish these two functions in my proposed classification. Of the four examples of this function of 22A in my data set, three apply to word- or phrase-initial grammatical morphemes (e.g., **ʔu** for *u-* 's/he/it; his/her/its', or **ʔa** for *a-* 'you (singular); your (singular)', **ma** for *maʔ* 'no/not'), requiring the application of the same morpheme at the beginning of a noun phrase in sequences of two noun phrases functioning as a couplet. The case in Figure 5a shows **[2]ʔu-KAB'-CH'EN** for *u-kab' u-ch'en* 'his/her/its land, his/her/its well/cave; his/her/its settlement/town'. An unfurled version of this couplet is seen in Figure 5b. As Zender (1999) had observed, the fact is that this "non-serial" duplication of syllabograms was also common practice even without 22A: there are cases of couplets where a single **ʔu**, without 22A, must be read twice within the collocation, such as the **ʔu-to-k'a-pa-ka-la** expression (Yaxchilan Lintel 46, F08), also represented **ʔu-to-k'a=ʔu-pa-ka-la** (Yaxchilan Lintel 45, C06), both of which spell *u-tok' u-pakal* 'his/her flint, his/her shield'. This same duplication without diacritic was expected in cases where the same syllabogram could be used at the end of two words that belonged to the same common sequence of nouns, as with **CHAN-na-CH'EN-na** (e.g., COLK1398:R4, Copan Stela 13 Altar:H01) versus **CHAN-CH'EN-na** (Tikal Stela 31:H23), both for *chan ch'en* 'sky cave/well'. Typically, such instances of required duplication of the reading of a syllabogram without 22A involved visual overlap: the syllabogram to be repeated had to be visually adjacent to spellings of both words.

Interestingly, logograms could also experience "duplication" without 22A. The logogram **K'UH(UL)** 'god(ly)' could be similarly applied jointly to two separate nouns by means of the convention of visual overlap, as in Figure 5c, where the expression **K'UH(UL)-CHAN-nal-ʔla-KAB'-ʔla**, for *chanal k'uh, kab'al k'uh* 'celestial god, earthly god', is shown. Or it could be applied individually to each of the nominal expressions, as in Figure 5d, showing **K'UH(UL)-CHAN K'UH(UL)-KAB'**. Given this scribal practice of duplicating the reading of a logogram, without 22A, in contexts requiring "unfurling," it is to be expected that at least one case of 22A applying to a logogram would require a duplication function. Although *not one example* of such a case has surfaced to date, I would not be surprised if at least one were to do so in the near future.

Zender (1999:128) proposed that in instances in which "doubled signs were not meant to be read serially," the diacritic would be found, with "no exceptions," in the top-left corner of the collocation. There are very few examples of

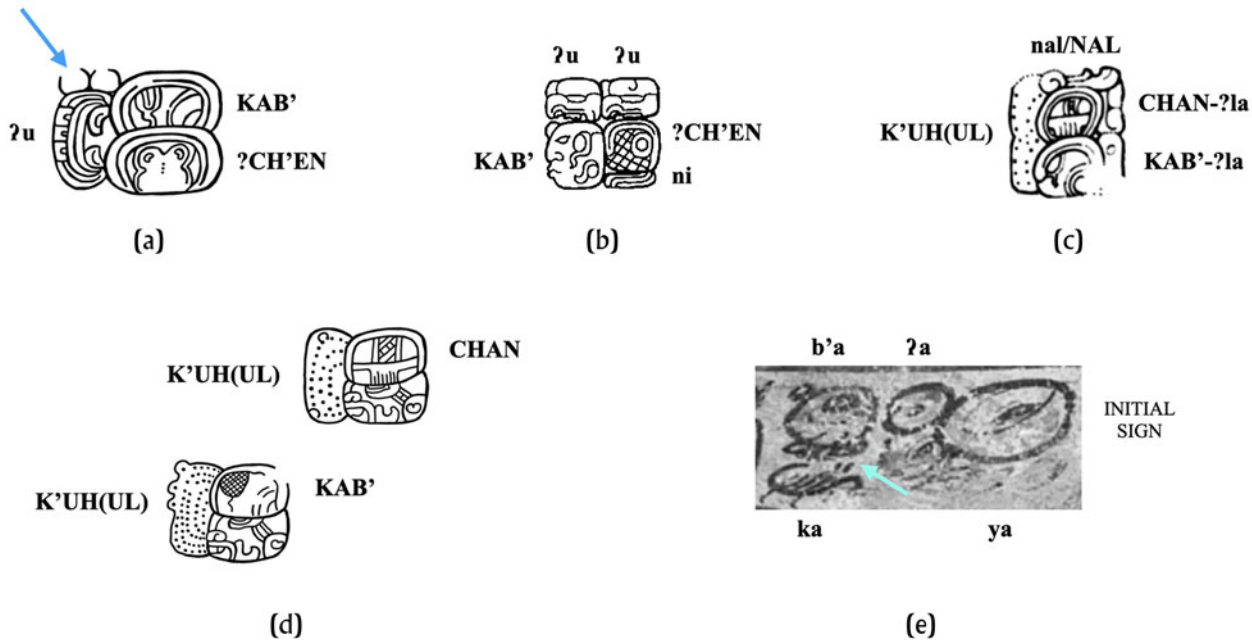

Figure 5. Examples of nonsequential duplication. (a) Glyph block C01 from Copan Stela 49 (CPNSt49). Drawing by author (after drawing by David Stuart in Schele 1990:Figure 19a). (b) Glyph block G'01 from Quirigua Zoomorph G. From drawing by Matthew Looper (2007:96–97, Figure 3.38). (c) Glyph block E08 from Copan Stela 7. From Drawing #1031 from the Linda Schele Drawings Collection at http://research.famsi.org/schele.html. (d) Glyph blocks F25-E26 from Tikal Stela 31. Drawing by author (after drawing by William R. Coe [Jones and Satterthwaite 1982:Figures 51 and 52]). (e) Excerpt from K1552. Photograph from Justin Kerr from Maya Vase Database at http://research.mayavase.com/kerrmaya.html.

https://doi.org/10.1017/S0956536123000317 Published online by Cambridge University Press

22A in this nonsequential duplication function: of the four examples of nonsequential duplication in my data set, three are, *generally*—although not always strictly—on the top-left corner of the collocation. Note that in Figure 5a, 22A is in the top locus, not the top left. Nevertheless, in the example in Figure 5e, the spelling **^{2}b'a-ka** for *b'ah=kab'* 'head/top of land'—a title—22A is found on the bottom of the target grapheme, placing the diacritic in the middle of the collocation.

Zender (1999:120–121) also offered some interesting remarks on the practical rationale for the innovation of 22A. He noted that because Mayan scribes often duplicated syllabograms for purely graphic purposes (e.g., to fill in more space within a glyph block), duplication of such signs (e.g., **ka**, **la**) would not have been an obvious way to call for their sequential reading. Although I find this idea very appealing, this must not have been a very important concern for the scribes, given the overall rarity of use of 22A. Also, although this explanation makes sense with the highly frequent use of 22A with T25 **ka** (30 cases)—a sign that *was* commonly doubled for graphic purposes (e.g., instances of **ka** in RAZV15)—it does not work as well for other signs mentioned by Zender, such as **la**: the MHD data set of 132 cases of 22A bears only one instance in which 22A *may* have applied to **la**.

In addition, Zender (1999) proposed a subscript convention (e.g., **ka$_2$-wa**) for indicating the placement of 22A, but the subscript convention is already reserved for marking allograms (Fox and Justeson 1984), and therefore I will favor the superscript convention.

Recently, Kettunen and Helmke (2020:20) and Prager (2020:3) have suggested that the diacritic was used "on some rare occasions" with logograms of the shape CVC, where both consonants were identical; henceforth, I refer to such shapes as C_1VC_1. Indeed, 22A can apply to two confirmed logograms: T122/ZBBa **K'AK'** for *k'ahk'* 'fire' (Figure 6a) and **TZUTZ** for *tzutz* 'to finish' (Figure 6b). There is another likely logogram, ZRJ, whose reading remains unconfirmed (Figure 6c). Kettunen and Helmke (2020:20) highlight the case of **K'AK'** for *k'ahk'* 'fire', and they refer to the case of ZRJ, the RUBBER.BALL sign (Helmke, personal communication, 2022), which bears 22A in the texts from Palenque's Temple XIX Platform and Stone Pier. ZRJ, a circular sign depicting some sort of material rolled up onto itself, possibly depicting a rubber ball, has been proposed to bear the value **CH'ICH'/K'IK'** 'blood' (Helmke, personal communication, 2022). Around the same time, however, Prager (2020:3), who also remarked on the use of 22A with ZBBa **K'AK'** and the AW3/AW8/MR6 **TZUTZ** logograms, presented plausible although not definitive evidence in favor of reading **KUK** 'bundle, textile; roll up, wrap up' for ZRJ (Prager 2020:7). Pending confirmation for the reading of ZRJ, I will focus my remarks below on the two clear cases of 22A applied to C_1VC_1 shapes—namely, **K'AK'** and **TZUTZ**.

Most recently, Mora-Marín (2022b) presented evidence for another function of 22A as an abbreviation marker in collocations. As such, it would be analogous to a punctuation marker, akin to the use of the period to mark abbreviations in Latin-derived scripts (e.g., Prof., Dr., etc.). One example is the case on K1670 of the spelling **K'UHUL cha ^{2}TAN** (Figure 7a), which is present on a pottery vessel with a Primary Standard Sequence (PSS) text, immediately before the collocation that begins the inscription **yu-k'i-b'i** for *y-uk'-ib'* 'his/her cup'. This type of PSS text often begins with a possessed noun followed by the name of the intended owner, and the two (possessee and possessor) are often separated by an intervening prepositional phrase that modifies the possessee. In the case at hand, the **K'UHUL cha ^{2}TAN** expression corresponds to a well-known title, corresponding to the intended owner of the vessel. Typically, the title begins with *k'uhul* 'holy/divine', which is followed by *cha... tahn* (an expression that remains opaque) and ends in the logogram **WINIK** for *winik* 'person'. Such a typical case is seen in Figure 7b. In the case in Figure 7a, then, 22A marks the last component of the formulaic collocation, the logogram **TAN** for *tahn* 'chest', which typically precedes the last component, the logogram **WINIK** for *winik* 'person', which was omitted in this case, possibly because the scribe ran out of space before the text—painted around the rim of the vessel—returned to the beginning, the **yu-k'i-b'i** expression. Consequently, a logogram has been omitted from a formulaic collocation, and 22A points to the fact that this component is missing.

A second example, out of a total of five examples where 22A bears this function (Mora-Marín 2022b), occurs in the

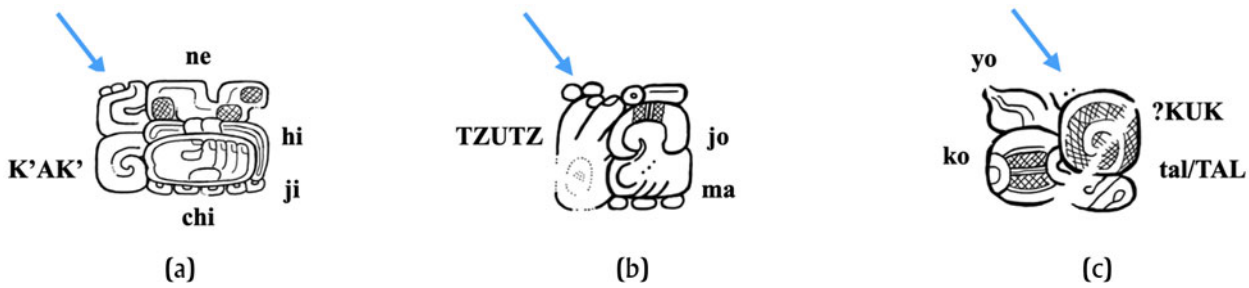

Figure 6. Cases of 22A applied to logograms with C_1VC_1 shapes. (a) Glyph block B02 from Santa Rita Corozal stone bowl. Drawing by author (after drawing by Stuart 2005:131). (b) Glyph block O02 from Tortuguero Monument 6. From drawings #109039 and #109029 from the Linda Schele Drawings Collection at http://research.famsi.org/schele.html. (c) Glyph block Y01 from Palenque Temple 19 Platform. Drawing by author (after drawing by Stuart 2005:123, Figure 92).

https://doi.org/10.1017/S0956536123000317 Published online by Cambridge University Press

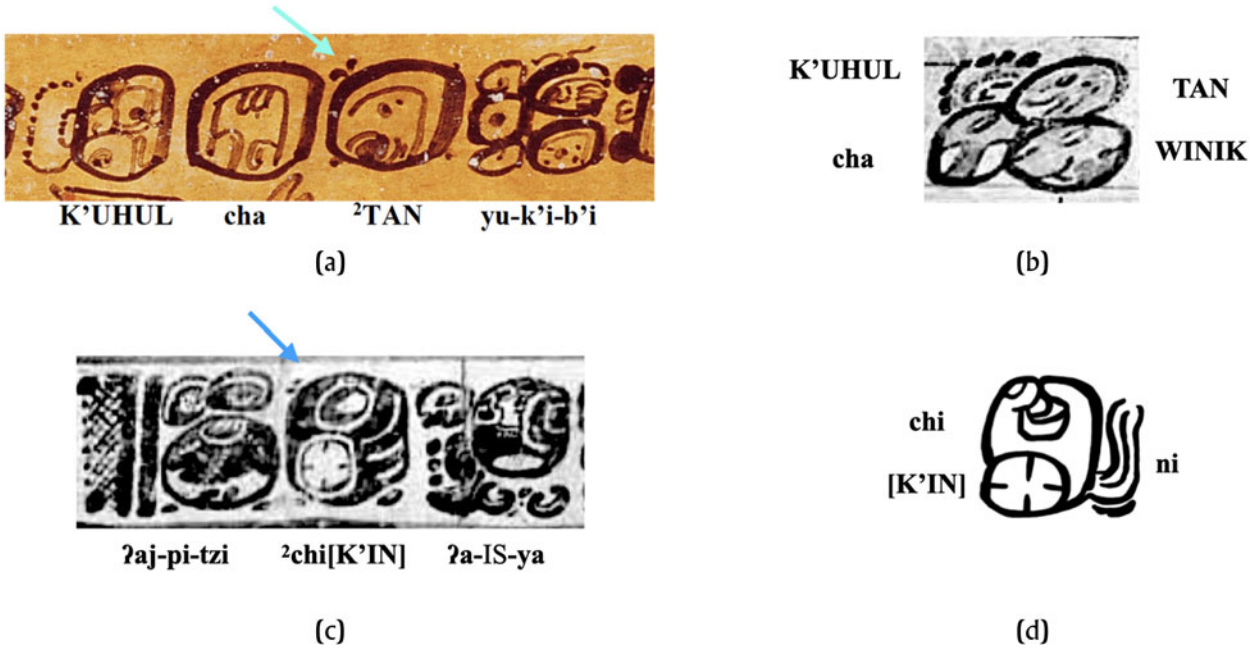

Figure 7. Cases of 22A in its abbreviation function applying to collocations. (a) Excerpt from K1670. Photograph courtesy of Donald Hales (All Rights Reserved). (b) Excerpt from K1810. (c) Excerpt from K2295. (d) Glyph from Simojovel Shell. Drawing by author (after drawing by Peter Mathews in Robertson et al. 1976:Figure 9). (b)–(c) Excerpts from photographs by Justin Kerr from Maya Vase Database at http://research.mayavase.com/kerrmaya.html.

last glyphic collocation of another PSS text painted on a different pottery vessel. In this instance (Figure 7c), the spelling ²**[chi]K'IN** appears as part of a title phrase for an individual, the intended owner of the vessel, and it likely represents the common appellative *k'ihnich/k'iinich* 'radiant/Sun God', referring to the Sun God. The collocation appears at the very end of the text, immediately before the so-called Initial Sign of the PSS, spelled **ʔa**-INITIAL.SIGN-**ya**. This Initial Sign appears in hundreds of PSS texts on pottery vessels, and it unambiguously marks the beginning of such texts. The scribe clearly reached the end of the text (also the beginning of the text) and ran out of room to complete the typical spelling of the 'radiant/Sun God' title, which, when spelled logosyllabically (as opposed to simply logographically), typically bears the **ni** syllabogram (Figure 7d), if the **chi** syllabogram is also present (120 cases in the MHD). Only rarely (17 cases in the MHD) does it show **chi** by itself, without a **ni** in such spellings. Consequently, in the case at hand, the scribe appears to have used 22A to indicate that something was missing—the typical **ni** syllabogram of logosyllabic spellings of *k'ihnich/k'iinich*.

Two more examples support this abbreviation function. They both involve the child-of-father expression. In both cases, 22A is placed between **K'AK'** for *k'ahk'* 'fire' and the T535 (Capped Ajaw) sign, as seen in Figures 8a–b. In this collocation, T535/ZA3 (Capped Ajaw) and T533/ZA1 (Regular Ajaw) may co-occur (Figure 8c). When this happens, T535 always precedes T533. The 22A diacritic only appears in two cases—the two examples already noted—and in both cases, it is T533 that is seemingly absent from the

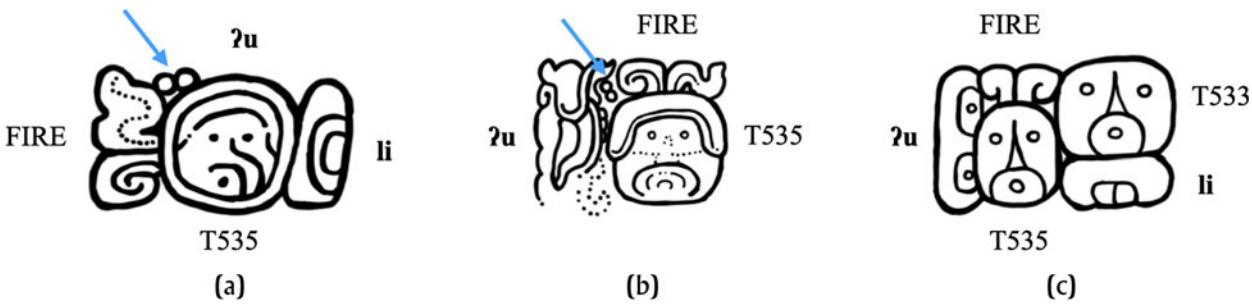

Figure 8. Cases of 22A in its abbreviation function applying to collocations. (a) Itzimte Bolonchen Stela 7 (ITBSt017). Drawing by the author (after drawing by von Euw 1977:19). (b) Excerpt from Tikal Stela 22 (TIKSt22). Drawing by the author (after drawing by William R. Coe in Jones and Satterthwaite 1982:48–50, Figure 33). (c) Excerpt from (BPKSt02). Drawing by the author (after drawing by Peter Mathews 1980:62, Figure 2).

https://doi.org/10.1017/S0956536123000317 Published online by Cambridge University Press

collocation. Also, 22A does not always appear in instances in which T533 is absent; as has already been explained, 22A is optional. But both cases where it does appear are instances in which T533 has been omitted. In principle, examples like those in Figures 8a–b (and similar examples lacking the optional 22A) could be argued to be instances in which T533, Regular Ajaw, has been *infixed* within T535, Capped Ajaw, resulting in the appearance of only the Capped Ajaw. Consequently, 22A could be functioning, in the rare occasion when it is present in the child-of-father collocation, to indicate that something is missing—specifically, T533—or at the very least, not obvious.

Four examples, then, apply to logograms, one to a syllabogram, and all fall within the Late Classic period. I propose that it was a late innovation—an analogical extension of the more common duplication function—and one that operates at the collocational level, and therefore, at the supragraphemic level. This means that at the graphemic level, either logograms or syllabograms may be the target. Both the examples in Figures 7a and 7c, consequently, constitute cases of the use of 22A in a glyph block at the end of a text. At least one other instance, shown in Figure 5e, also involves 22A in the last glyph block of a text, suggesting that perhaps position within a text could have an influence on its application.

Functional classification of 22A and hypotheses

Given what has been discussed so far, I propose a classification of four functions of the 22A diacritic:

(1) Sequential *duplication* of syllabograms (Stuart 2014; Zender 1999);
(2) *Nonsequential ("non-serial") duplication* of syllabograms (Stuart 2014; Zender 1999);
(3) Marking of logograms representing lexemes based on C_1VC_1 roots/stems (Kettunen and Helmke 2020; Prager 2020); and
(4) *Abbreviation* function, applied to graphemes (logograms, syllabograms) or supragraphemically (entire collocations), proposed here for the first time, which implies that 22A also functioned as a *punctuation* mark (Mora-Marín 2022b).

I propose that functions (2)–(4) constitute analogical extensions of function (1), given that the use of 22A to indicate the need for sequential duplication of a sign within a glyph block would have allowed for scribes to generalize the pattern to the nonsequential duplication of the same sign within a glyph block. In addition, the sequential duplication of a syllabogram to render a C_1VC_1 lexical root could have also led to the reanalysis of 22A as a marker of logograms based on C_1VC_1 roots. And last, given that in the first two functions 22A essentially indicates that something is missing in the spelling, scribes could have extended 22A to function as an abbreviation marker, this time applying at the level of a whole spelling rather than at the level of an individual grapheme within a spelling.

Postclassic codices

A search of 22A in the MHD yielded no cases in the Postclassic codices. One may wonder whether the conditions for any of the functions of 22A just reviewed are present in the codices, and whether the more frequent and likely lexical and morphemic targets of 22A occur as well. I have not conducted a comprehensive search. Instead, I have restricted myself to investigating whether incomplete spellings of *käkäw* 'cacao', *ʔunen* 'baby; child', and *-(V)l-el* 'abstractivizer of nouns' are amenable to the sequential duplication function of 22A—its most frequent function. Figure 9 presents characteristic examples of each one.

The first target (Figure 9a), *käkäw*, occurs 13 times, in all of them showing at least **ka-ka**, and 11 showing **ka-ka-wa**. Not one shows **ka-wa**, the context that would allow for the *optional* application of 22A. The second and third targets, though, do exhibit the right contexts: of the four instances of *ʔunen*, as part of a diphrastic kenning *y-aal y-unen* 'her child, her baby' for 'her children' (Figure 9b), all four show only **ne**, with not a single case of **-ne-ne**; and of the 72 instances of *ʔajaw-(a)lel* 'kingship' (Figure 9c), all cases show only **-le**, with not a single case of **-le-le**. Consequently, the conditions for the application of 22A are present in the Postclassic codices. The Postclassic scribes were just not keen on 22A, if they knew of it at all.

Objectives

Given the foregoing review and discussion, any data set for quantitative analysis of 22A should consider, at the very least, the following scriptal factors variables: the graphemic targets of 22A (syllabogram vs. logogram); its position within a text (dispersion); the function of 22A (duplication,

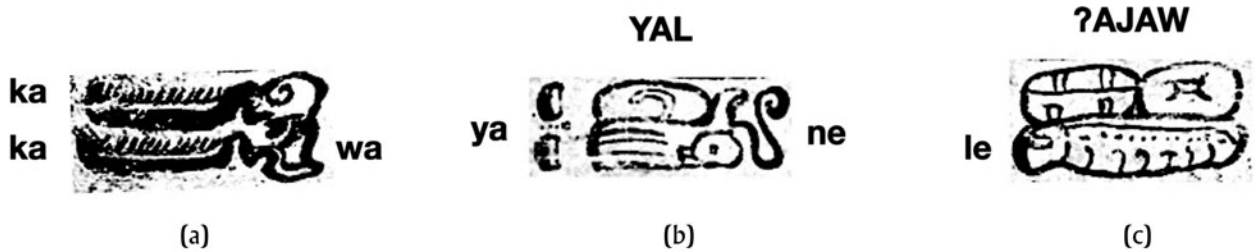

Figure 9. Examples of common lexical and morphological targets of 22A in the Postclassic codices lacking 22A. (a) **ka-ka-wa** for *käkäw* 'cacao' on DRE25:A01. (b) **ya-YAL-ne** for *y-aal y-unen* 'children' on DRE02:D01. (c) **ʔAJAW-le** for *ʔajaw-(a)l-el* 'kingship' on DRE02:B01.

https://doi.org/10.1017/S0956536123000317 Published online by Cambridge University Press

nonsequential duplication, C_1VC_1 shape marking, and abbreviation); and the locus of 22A (where it is placed with respect to the target grapheme). Dispersion here refers to the distribution of 22A within a whole text, whether it occurs near the beginning, middle, or end. Intuitively, one could imagine that as scribes get closer to the end of a text, especially one painted in a rush, and perhaps with less planning time, more information is packed toward the end, and the limited space prompts scribes to abbreviate more. However, to test whether this is an important factor, it would be necessary to have dispersion values for all the spellings of a given lexeme (e.g., *käkäw*), both those that use 22A and those that do not. This is a task for a future study.

Following Kettunen and Helmke (2020) and Prager (2020), who raised the possibility of a linguistic factor (C_1VC_1 shapes), I further propose the following linguistic variables for consideration: phonological classes (manner and place of articulation) and grammatical classes (nouns, verbs, grammatical morphemes) of the graphemic targets. In addition, given that the prior literature has not explicitly discussed nonscriptal and nonlinguistic factors in connection with 22A, I propose the following: media (portable vs. monumental), geography, and time. Studying these variables should allow one to investigate questions regarding the historical development of 22A: when it originated, where, what its most likely functions and targets were early on, and how it became extended to more functions and targets as time went on. More narrowly, it is possible that the data may point to specific lexemes that may have promoted the use of 22A as well as influenced the scribes' preferences in its application.

Procedures and methods

I have employed the MHD (Looper and Macri 1991–2023) to prepare a data set amenable to quantitative analysis. A total of 132 records with 22A were downloaded (in August 2022) as a CSV file and curated with Apple Numbers. This process involved several steps: (1) checking all examples visually for accuracy; (2) culling examples that could not be corroborated; (3) adding a few examples not included in the MHD at the time (of which a couple have since been added); and (4) preparing data categories, which are identified in the Objectives section, to study variables relevant to the current study.

A minor issue pertains to the dates provided by the MHD: (1) some dates are based on calendrical evidence internal to a text; (2) others are based on calendrical evidence and associations (e.g., names of individuals known from a specific period of time); (3) others are based on archaeological associations (e.g., interment in the burial of a historically known individual, ideally one mentioned in the text itself); and (4) others are estimates based on style (generally) and labeled "estimate." This information is captured in the "objcal" field of each record, which needs to be selected prior to downloading data so that the records may be sorted appropriately. For the purposes of studying the chronological development of 22A, I have only included date categories (1)–(3) as part of the independent metric variable of time, which appears in Gregorian years rather than in Mayan Long Count dates in my results. The fourth category has been included in the more general category of all texts, whether dated or not, and coded for time period, as an ordinal variable, either as Early Classic (ca. A.D. 200–600) or Late Classic (ca. A.D. 600–909).

After processing the data, the final data set retained 125 records. Nonetheless, different analyses are based on different total numbers of cases, because it is not possible to classify all cases according to the all the variables of interest in this article. As a case in point, dated texts ($N = 74$) constitute only 59.2 percent of the data set. However, for some tests, time measured metrically (in years) is more useful and accurate than time measured ordinally (in time periods, such as Early Classic, Late Classic), which tends to make up for the lower frequency of dated texts. For a very few tests, texts dated by style may be considered among the dated texts, which raises the total of dated texts to 86 (68.8 percent). In such tests, the inclusion of texts dated only stylistically (12 in total) constitutes a small but nonetheless important source of potential error.

When adding cases of 22A not recorded in the MHD, I have erred on the side of including anything that shows 22A, but for which a canonical duplication function may or may not be ascribed. For example, despite the obvious presence of 22A in two examples from the famous lock-top cacao pot of Rio Azul (MHD code RAZV15), these spellings were not coded in the MHD as containing 22A. The two spellings are identical: they are both cases of **ka-^{2}ka-wa** for *käkäw* 'cacao'. Perhaps the authors of the MHD omitted 22A in these spellings because it seems superfluous: the spellings already contain two instances of the syllabogram **ka,** to which 22A was applied, and therefore there is nothing to duplicate. Nevertheless, I have included these two cases in my data set, because I believe they can teach us something about the way that scribes were thinking about 22A.

Table 1 provides the basic categories and variables of the data set.

Table 2 shows the nature of the variables as metric, nominal, or ordinal. It also clarifies which sets of variables are mutually colinear, and therefore were not included in statistical tests at the same time.

I assessed the grapheme type that was the target of 22A—that is, whether 22A was applied to a syllabogram or logogram—not only by paying attention to the placement of code 22A relative to other graphemes in the column indicating the grapheme codes but also by checking the figures provided in the MHD for each glyph block, or by seeking out figures from other sources. The latter step was necessary whenever the MHD lacked relevant images, or whenever the image provided by the MHD appeared to lack information (some illustrations in the MHD have omitted the 22A diacritic). Cases for which no corroboration was possible were omitted from the data set or from a particular test. Grammatical Class, a category involving nouns and verbs, was ascertained by checking every example in the MHD directly, and in some cases, cross-checking with existing categories in the MHD (e.g., *blsem* for semantic class, *blmaya1* for Mayan transliteration, etc.).

https://doi.org/10.1017/S0956536123000317 Published online by Cambridge University Press

Table 1. Basic dataset categories and variables (125 records).

Variables and Metadata	Variants or Categories	Cases	Variables and Metadata	Variants or Categories	Cases
1. Texts	MHD objabbr codes	108	11. Narrow Phonological Class 1 (manner)	Stop	85
				Affricate	11
2. Time	Dated/Gregorian	74		Fricative	3
	Not dated	51		Nasal	11
				Liquid	10
3. Period	Early Classic	28		Semivowel	5
	Late Classic	97			
			12. Narrow Phonological Class 2 (place)	Alveolar	31
4. Provenience	Usu_West	27		Alveopalatal	2
	Usumacinta	28		Bilabial	22
	Eastern	17		Glottal	3
	Central	36		Palatal	2
	Northern	4		Velar	65
	Southern	2			
	Teotihuacan	1	13. Grammatical Class 1	Nouns	117
	Usu_Pasion	2		Verbs	8
	Unprovenienced	8			
			14. Grammatical Class 2	Nouns	108
5. Object class	Portable	54		Verbs	8
	Monumental	71		Grammatical Morpheme	9
6. Locus of 22A	Left	35	15. Function 1	Duplication	102
	Top left	53		C_1VC_1 shape	14
	Top	21		Abbreviation	5
	Top right	5		NonseqDupl	4
	Right	4			
	Bottom left	2	16. Function 2	Duplication	106
	Bottom right	2		C_1VC_1 shape	14
	[Unchecked]	3		Abbreviation	5
7. Grapheme type	Syllabograms	107	17. Function 3	Duplication	106
	Logograms	18		Nonduplication	19
8. Lexeme	Distinct lexemes	39	18. Date Estimates (nonvariable)	No (not estimate)	58
				Yes	12
9. Broad Phonological Class 1	Obstruents	99		YesAssoc	10
	Sonorants	26		Assoc	6
10. Broad Phonological Class 2	Glottalic obstruent	30			
	Plain obstruent	66			

https://doi.org/10.1017/S0956536123000317 Published online by Cambridge University Press

Table 2. Types of variables (metric, ordinal, nominal) and cases of colinearity.

Variables	Header Title	Colinear	Type
Dependent or independent	Grammatical Class 1	Yes	Nominal
	Grammatical Class 2		Nominal
	Grapheme Type		Nominal
	Broad Class 1	Yes	Nominal
	Broad Class 2		Nominal
	Narrow Class 1		Nominal
	Narrow Class 2		Nominal
	Narrow Class 3		
	Function 1	Yes	Nominal
	Function 2		Nominal
	Function 3		Nominal
	Locus		Nominal
Independent	Gregorian	Yes	Metric
	Period		Ordinal
	Media		Nominal
	Region		Nominal

The main source of error in my data set is the locus variable, which provides location values for 22A relative to the target grapheme. It is likely that different scholars would make somewhat different decisions regarding some of these values (e.g., top left vs. top).

In a relatively few cases, 22A was applied to syllabograms used to spell grammatical morphemes (e.g., **^{2}ʔa** to spell /a-/ 'second person singular ergative/possessive' twice, **^{2}le** to spell /-Vl-el/ 'abstractivizer of nouns and adjectives'). In such cases, I coded two sets of variables: Grammatical Class 1 with variants Noun and Verb, and Grammatical Class 2 with variants Noun, Verb, and Grammatical Morpheme. Of the total 125 cases, only eight unambiguous cases involve grammatical morphemes, only one of which dates to the Early Classic period (A.D. 514), and five cases make up the majority of instances in which 22A was applied to a syllabogram with a liquid consonant (i.e., **^{2}li** or **^{2}le**).

Next, I present the results of descriptive and inferential statistical tests and offer some discussion of their implications. I begin with the results pertaining to the independent variable of time, both for dated texts (in actual years) and then for all texts according to the ordinal distinction of Early Classic versus Late Classic. Then, I consider interactions between pairs and sets of nominal variables, some independent (region, media), and others dependent (grapheme type, grammatical class, phonological class, locus). I have employed DATAtab (DATAtab Team 2022) for the descriptive and inferential statistical procedures, as well as the graphs that make up most of the figures. Several nonparametric tests (Kruskal-Wallis Test, Mann-Whitney U Test) were used for assessing whether the variants or categories of a dependent variable (e.g., syllabogram vs. logogram as variants of the grapheme type variable) differ significantly from each other with respect to time (Gregorian years). Chi-square tests of independence were used for assessing whether two sets of categorical variables (each one nominal or ordinal) exhibit a statistically significant relationship. Last, Logistic Regression Tests were used to assess whether, and to what extent, multiple independent variables (e.g., region, media, Gregorian, period) may influence a nominal dependent variable (e.g., grammatical class), allowing for an assessment of which independent variables may exert a stronger influence than others. The details of the statistics and the data sets are provided as Appendixes 1–3 of the Online Supporting Materials (OSM).

Results

Relationship between 22A and time

Recall that no cases of 22A are attested in the Postclassic codices, according to the data from the MHD, so only Classic-period examples will be of relevance. A Late Preclassic example dating to the first century B.C., the so-called Uaxactun perforator (Kováč et al. 2016), may exist, but its presence has yet to be proven to constitute a diacritic function—much less be correlated with a specific type of diacritic function—given that graphemes to which the TWO.DOTS grapheme was applied, not to mention the text as a whole, remain largely undeciphered and untranslated. It is possible that the two examples of this grapheme on the perforator could constitute cases of the numerical logogram **CHAʔ** for *chaʔ= 'two', or even the syllabogram **ʔu** without a "bracket" element, consisting of only the two dots.

Figure 10 presents a basic chronological breakdown of the data set, with Figure 10a showing the distribution of all texts in their respective Early Classic and Late Classic periods; Figure 10b showing the distribution of the reliably dated texts, with a mean corresponding roughly to the year A.D. 700; Figure 10c showing the geographic distribution by period; and Figure 10d showing the geographic distribution for dated texts.

Figure 11 presents the temporal distribution of texts containing 22A broken down by media, whether portable or monumental. It is noteworthy that the proportions of texts with 22A shift according to media: portable texts make a larger proportion of the Early Classic data set relative to monumental texts, but the proportion is inverted during the Late Classic period. When all dated texts are considered, including the 12 texts dated stylistically, portable texts exhibit a lower mean value corresponding to the year A.D. 588; when those 12 texts are subtracted, then monumental texts exhibit a lower mean value of approximately A.D. 699. The question is whether these apparent differences are statistically significant.

Before addressing this question, a few words about the statistical tests are needed. Table 3 presents the statistically

https://doi.org/10.1017/S0956536123000317 Published online by Cambridge University Press

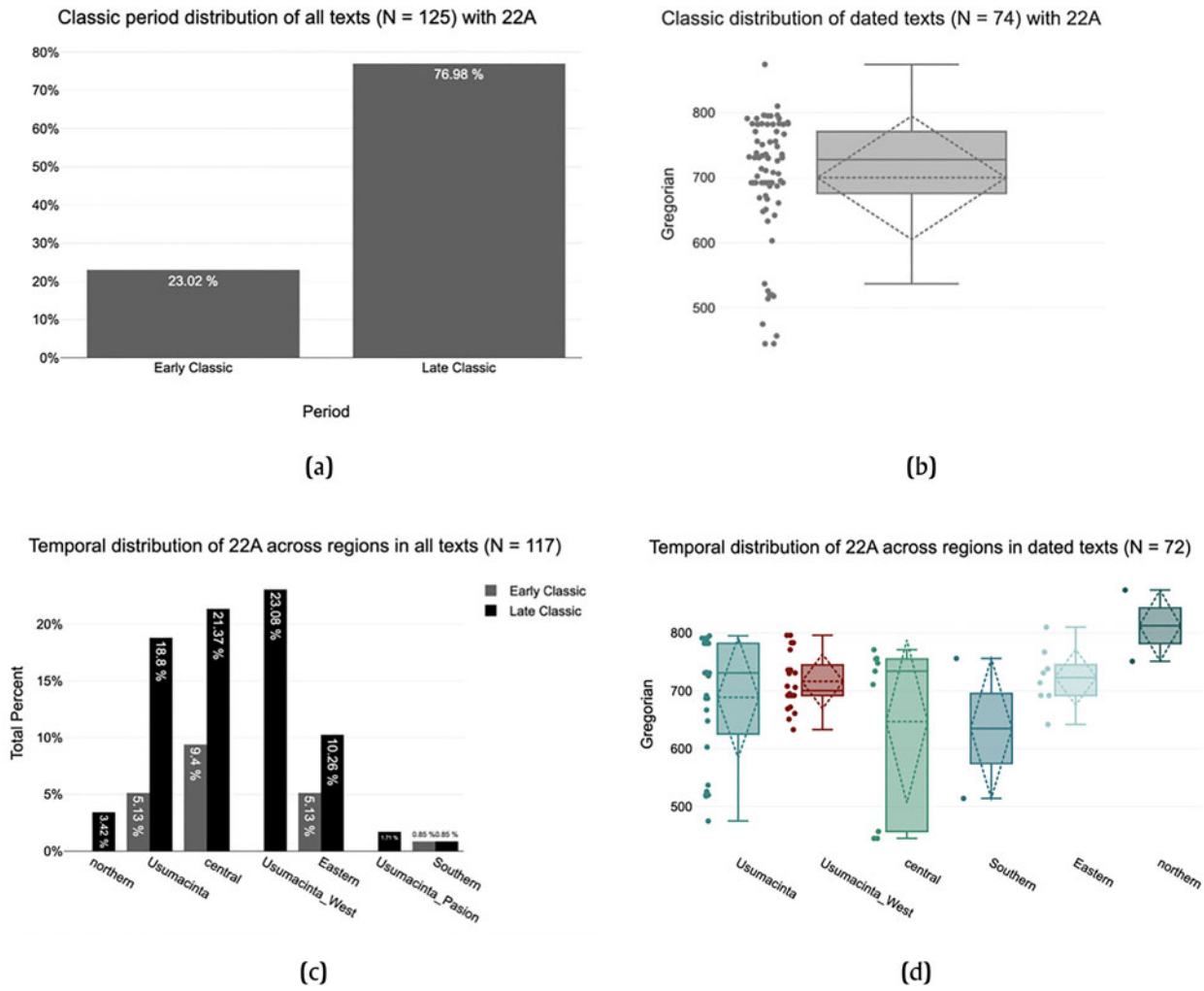

Figure 10. Overall distribution of texts with 22A during Classic period. (a) All texts ($N = 125$), dated and undated, by period. (b) All reliably dated texts ($N = 74$), with mean (699.82) represented as a horizontal dashed line, and median (728) represented as a horizontal solid line. Dashed triangle represents standard deviation (95.37). (c) Distribution of all generally provenienced texts ($N = 117$) across regions by period. (d) Distribution of reliably dated texts ($N = 74$) across regions.

significant results of the variables of relevance with respect to time, considered both as a metric variable (Gregorian years) and as an ordinal variable (period). The final column refers to the tables within Appendix 1 of the OSM (OSM-1), where the detailed statistical results can be consulted. Every test was carried out with both the full data set of dated and undated texts, and the data set of only dated texts. Next, I will review the significant results in a bit of detail to draw out their implications.

Returning to the question of the relative distribution of texts with 22A broken down by media (portable vs. monumental), the results summarized in Table 3a, considering all texts ($N = 125$), indicate that portable texts make up a significantly higher proportion of instances of 22A during the Early Classic period compared to the Late Classic period, as illustrated with Figure 11a. Similarly, the summary in Table 3b indicates that when all dated texts ($N = 86$) are considered, including the 12 dated stylistically, the difference is also significant, with portable texts with 22A being significantly earlier than monumental texts. When only reliably dated texts ($N = 74$) are considered (OSM, Appendix 1, Table A5), the difference between portable and monumental is no longer significant. On the weight of the first test (all texts), I propose the following hypothesis:

Hypothesis 1: 22A may have been innovated, but at the very least was initially popularized, on portable media. During the Late Classic period, 22A became more widely accepted, and in fact favored, on monumental media. This may reflect an important difference (e.g., in style or formality, or simply available space) between portable and monumental media, but one that became blurred over time.

There was a statistically significant result related to geography and time (Table 3c): when all regions are considered together (except for Unprovenienced and Teotihuacan, the latter with only one case), as in Figures 10c and 10d, they are, as a whole, significantly different with respect to time. However, the data are too limited for some regions to allow for more narrow results across the board. When additional regions

https://doi.org/10.1017/S0956536123000317 Published online by Cambridge University Press

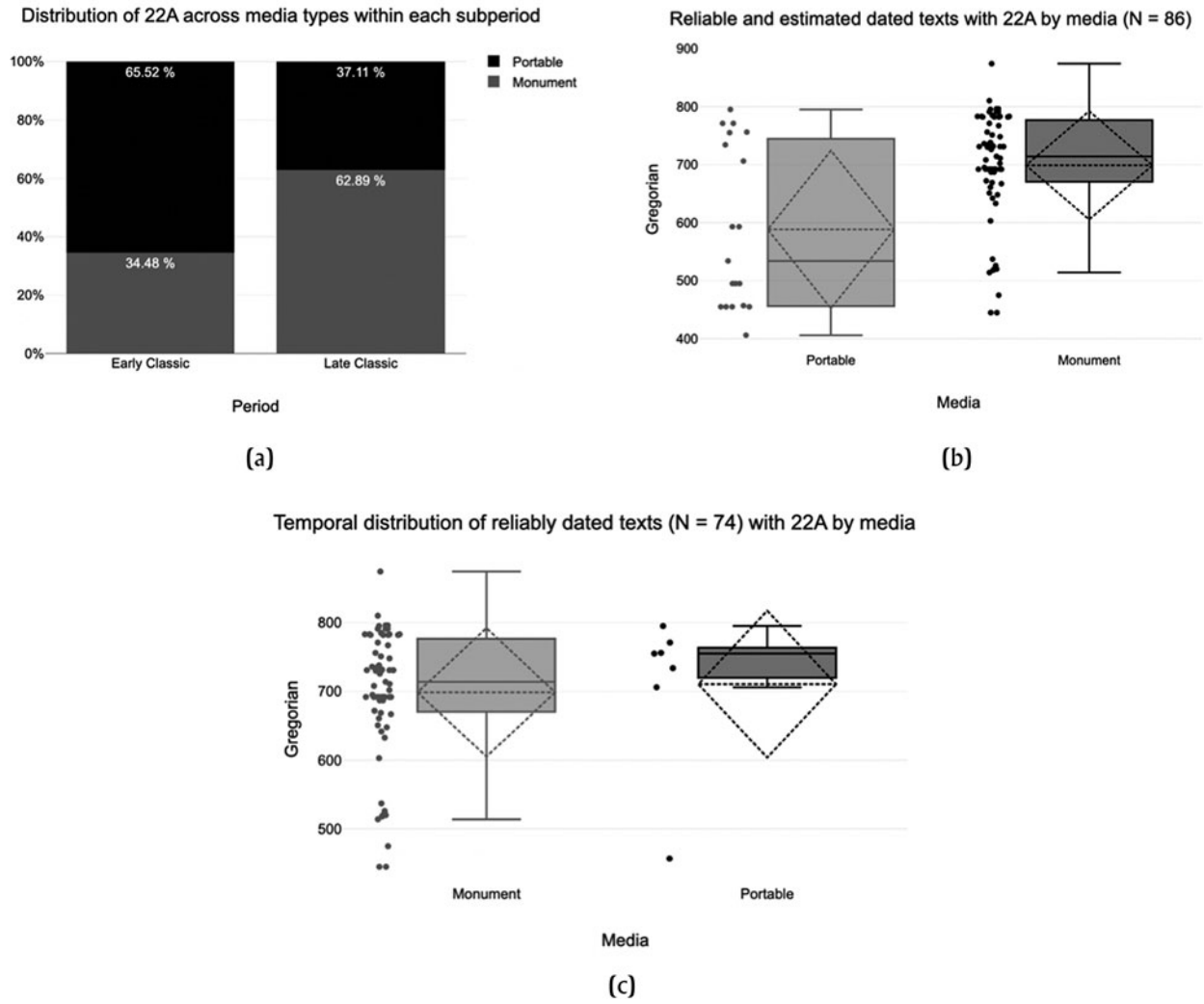

Figure 11. Distribution of texts with 22A during Classic period by media. **(**a) All texts (*N* = 125), dated and undated, by period and media, showing total percentages within period by media. Portable texts make up greater proportion of Early Classic examples of 22A. The distribution is inverted during the Late Classic period. (b) All dated texts (*N* = 86). Portable texts exhibit a lower mean value (588.21) than monumental texts (698.7). (c) Only reliably dated texts (*N* = 74) by media. Monumental texts exhibit a slightly lower mean value (698.7) than portable ones (710.57).

were excluded due to their scarcity of cases (Southern, Northern), the results (Table 3d) point to significantly different pairwise comparisons (Usu_West vs. Central, Usu_West vs. Eastern), with the Central region (lower mean values) showing overrepresentation during the Early Classic, and the Usumacinta_West region (higher mean values) in the Late Classic (Table 3e). Given that the Central and Usumacinta regions have the earliest reliably dated examples, A.D. 445 and A.D. 475, respectively, and given that the Central region is at least significantly different (i.e., earlier) than the Usumacinta_West region, I put forth the following hypothesis:

Hypothesis 2: 22A may have been innovated, but at the very least was initially popularized, in the Central region during the Early Classic.

It is worth exploring whether Hypotheses 1 and 2 may be combined: is there a significant relationship between media and geography? Using only the four regions with ample representation of 22A cases (Central, Usumacinta, Usumacinta_West, and Eastern), a chi-square test of independence with respect to media (portable vs. monumental) was carried out. The results, summarized in Table 3f, show that there is: portable texts are overrepresented in the Central region with respect to monumental, and by comparison with the other regions. Consequently, the following hypothesis can be proposed:

Hypothesis 3: 22A may have been initially popularized on portable media in the Central region during the Early Classic.

Next, I examine scriptal and linguistic variables pertaining to 22A in relation to time. The significant statistical tests are summarized in Table 4.

First, I will consider the scriptal variable of the graphemic target. Despite the difference in frequency between

https://doi.org/10.1017/S0956536123000317 Published online by Cambridge University Press

Table 3. Summaries of significant statistical tests, Part 1: Media and Region variables.

	Dependent Variable	Relevant Category	Time[a]	Result	Test	*p*-Value	Effect Size	OSM-1
a.	Media	Portable	EC	***Overrepresented***	Mann-Whitney U test, Chi-square test	0.011	Cramér's V = 0.23 (small)	Table A4a–b
		Monument	LC	***Overrepresented***				
b.	Media	Portable	Gregorian	Lower values	Mann-Whitney U test	0.007	*r* = 0.29 (small)	Table A4c
		Monument		Higher values				
c.	Region	All categories (except excluded)	EC/LC	Significant difference with respect to time	Kruskal-Wallis test	0.035		Table A60
		Central vs. Usu_West		Usu_West almost significantly higher mean than Central	Dunn-Bonferroni tests	0.065		
d.	Region	All categories (except excluded)	EC/LC	Significant difference with respect to time	Kruskal-Wallis test	0.014		Table 61
		Central vs. Usu_West		Usu_West higher mean than Central	Dunn-Bonferroni tests	Adj. *p* = 0.02		
		Usu_West vs. Eastern		Usu_West higher mean than Eastern	Dunn-Bonferroni tests	Adj. *p* = 0.045		
e.	Region	Central	EC/LC	Overrepresented in Early Classic	Chi-square test	0.008		Table 62
		Usu_West		Overrepresented in Late Classic				
f.	Region	Central	Media	Overrepresented with portable	Chi-square test	<0.001	Cramér's V = 0.75 (large)	Table 63
		Usumacinta		Overrepresented with monumental				
		Usu_West		Overrepresented with monumental				

[a]EC = Early Classic, LC = Late Classic, Gregorian = Gregorian calendar years.

https://doi.org/10.1017/S0956536123000317 Published online by Cambridge University Press

Table 4. Summaries of significant statistical tests, Part 2: Scriptal and linguistic variables.

	Dependent Variable	Relevant Category	Time[a]	Result	Test	*p*-Value	Effect Size	OSM-1
a.	Locus	All categories (unchecked and bottom right excluded)	Gregorian	Significant difference with respect to time	Kruskal-Wallis test	0.02		Table A25
		Top left vs. top		Top left higher mean values than top	Dunn-Bonferroni tests	Adj. $p = 0.046$		
b.	Broad Phonological Class 1	Obstruent	EC	***Overrepresented*** with obstruents	Mann-Whitney U test	0.011	$r = 0.23$ (small)	Table A9a
		Sonorant	LC	***Overrepresented*** with sonorants				
c.	Broad Phonological Class 1	Obstruent	Gregorian	Lower values	Mann-Whitney U test	0.019	$r = 0.27$ (small)	Table A9b
		Sonorant		Higher values				
d.	Narrow Phonological Class 1	All categories (including semivowel)	EC/LC	Significant difference among classes with respect to periods	Kruskal-Wallis test	0.018		Table A11
e.	Narrow Phonological Class 1	All categories (except semivowel)	Gregorian	Significant difference among classes with respect to time	Kruskal-Wallis test	0.016		Table A12
		Stops vs. liquids		Stops show lower mean values than liquids	Dunn-Bonferroni tests	Adj. $p = 0.012$		
f.	Broad Phonological Class 2	Plain obstruent	Gregorian	Plain obstruents take 22A earlier than glottalic ones	Mann-Whitney U test	0.049	$r = 0.27$ (small)	Table A17a–b
		Glottalic obstruent						
g.	Broad Phonological Class 2	Plain obstruent	EC/LC	Plain obstruents take 22A earlier than glottalic ones	Mann-Whitney U test	0.008	$r = 0.27$ (small)	Table A17c–d
		Glottalic obstruent						
h.	Grammatical Class 1	Nouns	EC/LC	—	Chi-squared specified distribution test	0.012		Table A23
		Verbs		Underrepresented				

[a]EC = Early Classic, LC = Late Classic, Gregorian = Gregorian calendar years.

the two graphemic targets (108 syllabograms, 18 logograms), the statistical tests did not yield any statistical significance in their temporal distribution as targets of 22A. For all texts, considering time as an ordinal variable (Early Classic vs. Late Classic), the results of the Mann-Whitney U test ($U = 773.5$, $p = 0.185$, $r = 0.17$) support the null hypothesis (no significant difference). The same was the case when time was treated as a metric variable (Gregorian years) (1) with all dated texts ($N = 86$), with results ($U = 388.5$, $p = 0.304$, $r = 0.11$) showing no significance; and (2) with all reliably dated texts ($N = 74$), with results ($U = 316$, $p = 0.417$, $r = 0.1$) also showing no significance. This is despite the fact that, so far, only one Early Classic text bears a case of 22A applied to a logogram (**^{2}K'AK'** for *k'ahk'* 'fire', present on a vessel from Santa Rita Corozal, dated stylistically and archaeologically to ca. 9.3.0.0.0 [ca. A.D. 495]). In contrast, the earliest *reliably* dated texts with 22A applying to syllabograms appear by ca. A.D. 445 (on Tikal Stela 31). These results could suggest that future findings may yet yield earlier cases of 22A applied to logograms.

Another scriptal variable, the functions of 22A (sequential duplication, nonsequential duplication, C_1VC_1 shape, abbreviation), was also considered. As with the case of the graphemic target variable, the statistical tests also failed to yield evidence of a significant difference among these various functions with respect to time, whether measured metrically (Gregorian years) or ordinally (Early Classic, Late Classic). Despite such results, it is worth observing that, thus far, all cases of the abbreviation function of 22A are found on Late Classic pottery vessels, and that only one case each of the nonsequential duplication and C_1VC_1 shape-marking functions is known from the Early Classic period.

The last scriptal variable is the locus of 22A—in other words, where it was placed in relation to its graphemic target. For these tests, the cases that could not be confirmed ("unchecked") were eliminated, as was the single instance of the bottom-right locus among dated texts. The results of the statistical tests summarized in Table 4a show significance overall and, more specifically, point to a significant difference between the top-left and top loci, with the former having higher mean values than the latter, despite the earliest attestation for both loci being on the same text (Tikal Stela 31) in A.D. 445. These results suggest the following hypothesis:

Hypothesis 4: 22A was innovated, or at the very least was initially popularized, on the top or top-left locus of the target grapheme.

Next, I will consider linguistic variables, starting with phonological variables. The Broad Phonological Class 1 variable, referring to whether 22A was applied to graphemes that contained an obstruent consonant (stops, affricates, fricatives) or a sonorant consonant (nasals, liquids, semivowels) is first. The statistically significant results summarized in Table 4b (all texts) and 5c (reliably dated texts) suggest that the earlier targets of 22A, in general, were graphemes with an obstruent consonant (e.g., **pa**, **ta**, **ka**, etc.), and that sonorant consonants (e.g., **ma**, **nu**, **ne**, **le**, etc.) became targets later. In fact, with one exception—the spelling **^{2}ma-si** on Piedras Negras Panel 12—22A begins to target sonorants during the Late Classic period. Given that obstruents and sonorants are broad classes of consonants, I also devised a Narrow Phonological Class 1 variable that breaks up consonants according to manner of articulation (stops, affricates, fricatives, nasals, liquids, semivowels). Although the statistical test applied to all texts (dated and undated, summarized in Table 4d) yielded significant results overall, the simple ordinal distinction of time into two periods (Early Classic, Late Classic) was not sufficient to resolve significant differences among subclasses of consonants. In contrast, the test applied to all reliably dated texts (summarized in Table 4e) yielded significant results overall, as well as a significant difference specifically between stops (e.g. **ka**, **ta**) and liquids (e.g., **le**, **li**). Additionally, given that 10 obstruent consonants exist in the varieties represented in Epigraphic Mayan that can be distinguished by plain (pulmonic) or ejective (glottalic) articulation, I devised a Broad Phonological Class 2 variable that distinguishes between Plain and Glottalic obstruents. The relevant statistical tests (see summaries in Table 4f and 4g) suggest that 22A was applied to plain obstruents (e.g., **ka**) significantly earlier than to glottalic ones (e.g., **k'a**). I therefore propose the following hypothesis:

Hypothesis 5: 22A may have been innovated, but at the very least was initially popularized, with lexical targets that begin with plain (nonglottalic) obstruents, more generally, and with plain (nonglottalic) stops, more narrowly.

The last linguistic variable worth considering is grammatical class. This variable posed a more complex problem. First, two grammatical class variables were defined: Grammatical Class 1, consisting of a three-way distinction between nouns, verbs, and grammatical morphemes; and Grammatical Class 2, a two-way distinction between nouns and verbs. Neither version yielded significant results—that is, it would seem that there is no significant difference between nouns and verbs (and grammatical morphemes) with respect to time, whether the latter variable is measured ordinally or metrically. However, such results seem counterintuitive for two reasons: (1) it cannot be assumed that nouns and verbs occur in equal proportions in a language (they usually do not); and (2) it cannot be assumed that C_1VC_1 shapes, which constitute the majority of the targets of 22A (whether we are dealing with the sequential duplication or C_1VC_1 shape-marking function), occur with equal frequency in nouns and verbs in a language. Regarding the first point, the proto-Ch'olan vocabulary by Kaufman and Norman (1984) contains 361 (60.7 percent) nouns and 234 (39.3 percent) verbs. And regarding the second point, the proto-Ch'olan vocabulary yields 43 lexical roots or stems with a C_1VC_1 sequence (i.e., C_1VC_1, C_1VC_1VC, CVC_1VC_1, C_1VC_1CV), of which 37 are nouns and verbs, broken down into 31 (72.1 percent) nouns (7.2 percent

https://doi.org/10.1017/S0956536123000317 Published online by Cambridge University Press

of nouns overall) and 6 (14.0 percent) verbs (2.14 percent of verbs overall).

Taking into account the total numbers of nouns and verbs with and without C_1VC_1 sequences, a chi-square test was carried out, with results ($\chi^2(1) = 8.98$, $p = 0.003$) pointing to a significant difference between nouns and verbs, with C_1VC_1 shapes overrepresented among nouns with respect to verbs. Assuming these proto-Ch'olan proportions are representative of the varieties spoken by the Epigraphic Mayan scribes, and considering again the total number of distinct expressions to which 22A was applied in the Late Classic and Early Classic periods—34 (33 nouns, 1 verb) and 7 (7 nouns, no verbs), respectively—we would expect to see approximately 28.1 distinct nouns and 5.9 distinct verbs for the Late Classic, and 5.8 and 1.2 for the Early Classic, respectively. These then constitute the "specified expected frequencies." If one now attempts another chi-square test of distribution to see how much the *observed* frequencies deviate from the *specified* expected frequencies, considering the numbers of distinct lexemes per period, the results summarized in Table 4h ($\chi^2(1) = 6.37$, $p = 0.012$) point to a statistically significant difference, essentially to a bias —an overrepresentation of nouns relative to verbs. Consequently, the following hypothesis can be proposed:

Hypothesis 6: 22A may have been innovated, but at the very least was initially popularized, with nouns (with C_1VC_1 sequences), and only later, and to a much lesser extent, extended to verbs (with C_1VC_1 sequences).

Relationships between nominal variables

In this section, I will report on the logistic regression tests that assess relationships between one dependent nominal variable and two or more independent nominal variables at once. Tables 5 and 6 present statistically significant results ($p = \leq 0.05$), as well as one case that was close to the conventional alpha value (0.05). Table 5 presents the results relevant to media and region as independent variables, whereas Table 6 presents the results relevant to scriptal and linguistic variables as independent variables.

An important result of Table 5 is that when the media and region variables are both tested as independent variables for a relationship with a scriptal or linguistic variable as a dependent variable, it is only the former (media) that appears to be an important predictor. Starting with scriptal factors, it can be observed, regarding the Function variable (Table 5a), that the C_1VC_1 shape-marking function exhibits a significant positive correlation (with Odds Ratio of 7.03) with monumental texts, whereas the duplication function exhibits a significant negative correlation with monumental texts. In addition, regarding the grapheme type variable (Table 5b), the results indicate that logograms are positively correlated with monumental media (Odds Ratio of 4.47), whereas syllabograms are negatively correlated with monumental media. These results are likely related, given that the C_1VC_1 shape-marking function applied to logograms (e.g., **^{2}K'AK'** and **^{2}TZUTZ**). Consequently, it is no surprise that both are positively correlated with monumental media.

Table 5. Important results of LR analysis of scriptal and linguistic variables in relation to media and region for all texts ($N = 125$)[a].

	Dependent Variable	Relevant Variant	Significant Relationships	*p*-Value	Coefficient B	Standard Error	Odds Ratio	z	95% Conf. Interv.
a.	Function	Duplication	Media: Monument	0.041	-1.31	0.64	0.27	2.05	0.08–0.95
		C_1VC_1 Shape	Media: Monument	0.019	1.95	0.83	7.03	2.35	1.38–35.67
b.	Grapheme Type	Syllabogram	Media: Monument	0.03	-1.5	0.69	0.22	2.17	0.06–0.87
		Logogram	Media: Monument	0.03	1.5	0.69	4.47	2.17	1.16–17.28
c.	Grammatical Class 2	Noun	Media: Monument	0.047	-2.4	1.21	0.09	1.99	0.01–0.97

[a]Statistical significance involves *p*-values <0.05.

https://doi.org/10.1017/S0956536123000317 Published online by Cambridge University Press

Table 6. Important results of LR analysis of scriptal and linguistic variables in relation to other scriptal and linguistic variables for all texts ($N = 125$).

Dependent Variable	Relevant Variant	Significant Relationships	*p*-Value	Coefficient B	Standard Error	Odds Ratio	z	95% Conf. Interv.
Grammatical Class 2	Noun	Broad Phonological Class 1: Obstruent	0.043	1.42	0.7	4.13	2.03	1.05–16.27
Narrow Phonological Class 1	Nasal	Locus: Left	0.026	-3.24	1.46	0.04	2.22	0–0.68

[a]For these tests, when a scriptal variable was the dependent variable, only Broad Phonological Class 1 was used as a phonological variable. In cases where locus of 22A is the dependent variable, the unchecked instances were omitted ($N = 122$). Statistical significance involves *p*-values <0.05.

The following hypothesis can be proposed on the basis of these results:

Hypothesis 7: Scribes favored the application of 22A with C_1VC_1-shaped logograms on monumental media from the beginning, and consequently, such practice may have been perceived as formal and prestigious.

Finally, when tested for a relationship with the media and region variables, Grammatical Class 2 (nouns vs. verbs vs. grammatical morphemes)—specifically the noun category—shows a significant negative correlation with monumental media (Table 5c). This is probably due to the already mentioned overrepresentation of nouns with respect to verbs: because all the cases of verbs that take 22A occur on monumental media, and because nouns that take 22A occur on both monumental and portable media, the apparent negative correlation between nouns and monumental media probably has more to do with the absence of cases of 22A applying to verbs on portable media.

Table 6 summarizes fewer significant results. Table 6a suggests that the left locus of 22A is negatively correlated with nasals (Narrow Phonological Class 1). Regarding Grammatical Class 2, Table 6b points to a positive correlation between nouns and obstruents (Broad Phonological Class 1).

22A and the lexicon

Another interesting question pertains to the distribution of 22A with respect to the lexicon. The data set of 125 examples (tokens) includes 39 distinct lexemes and grammatical morphemes (types). Figure 12a presents the frequencies of all 39. The frequencies can be best described in terms of a power trendline ($R^2 = 0.9612$), as in Figure 12b: the most frequent target of 22A, *käkäw*, is roughly twice as frequent as the second most frequent, *k'ahk'*, which is roughly twice as frequent as the next most frequent, *tzutz*, and so on.

Also interesting is whether one can predict how many spellings (tokens) of a lexeme or morpheme (type) will take 22A. Given that the majority of instances of 22A apply to syllabograms with the function of duplicating their reading, I have opted to focus on such instances only. Figure 13 provides a tentative answer: it shows a close agreement between the linear (Figure 13a) and polynomial solutions (Figure 13b). Because of this close agreement, the following linear equation should suffice: $x = (y - 8.5675)/13.183$. For example, the number of instances of *käkäw* in

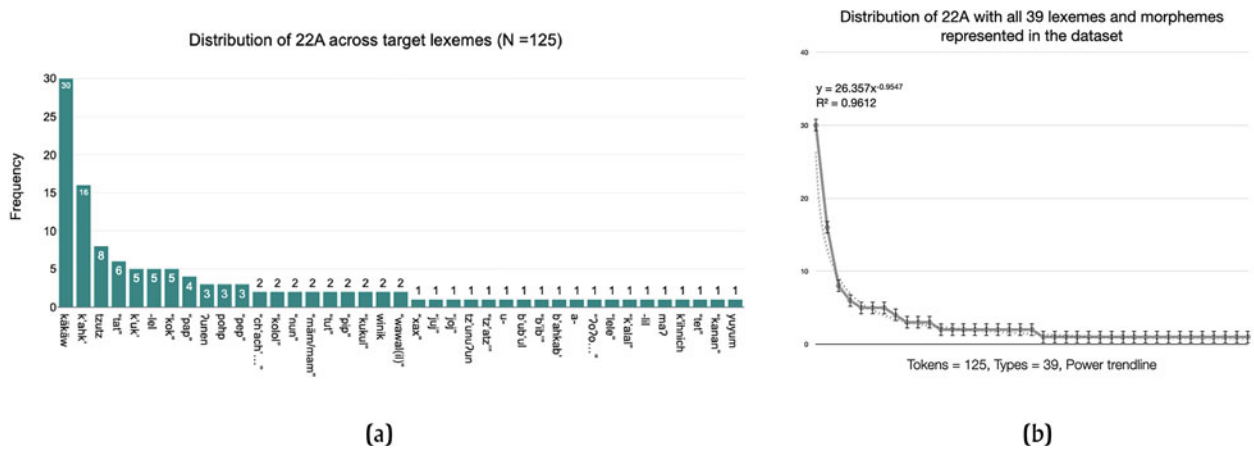

Figure 12. Lexical distribution of 22A. (a) Distribution of all 39 lexemes and grammatical morphemes. Prepared with DATAtab (DATAtab Team 2022). (b) Linear model (power trendline) of distribution of 22A among all 39 types.

https://doi.org/10.1017/S0956536123000317 Published online by Cambridge University Press

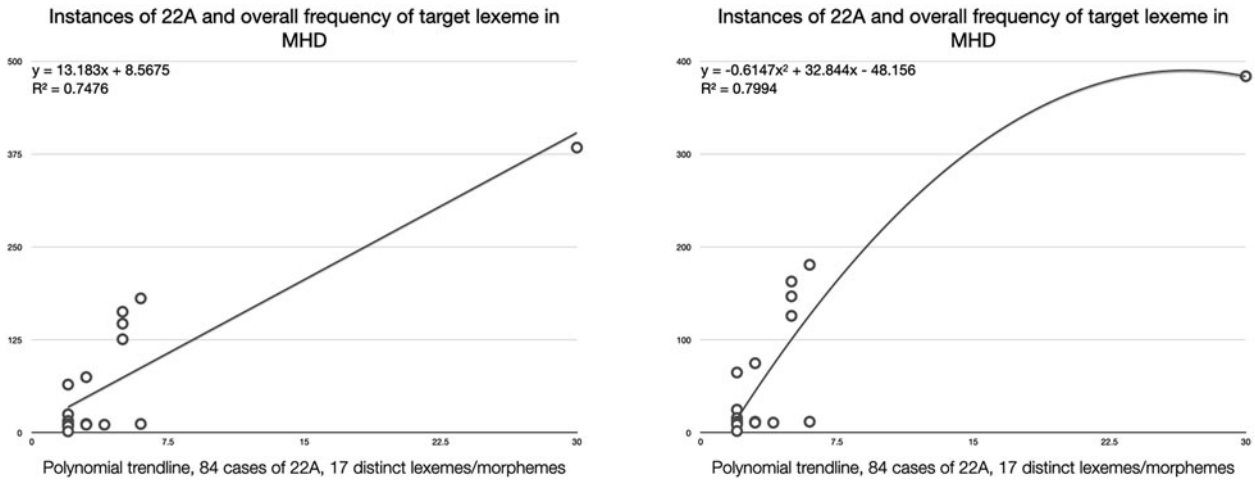

Figure 13. Linear relationships between overall attestations with syllabograms as targets and frequency of use of 22A. (a) Linear trendline. (b) Polynomial trendline. Prepared with Apple Numbers.

the MHD is 384; hence, $x = (384 - 8.5675)/13.183 = 28.5$. This is close to the actual number of cases of with *käkäw* attested with 22A (30). The much more complex and precise polynomial equation yields $x = 29.98361$, but this equation is less practical.

Revisiting käkäw 'cacao'

It is worth repeating the observation (cf. Figure 12a) that the most frequent use of 22A involves the *käkäw* expression, with 30 out of 125 examples, or 24 percent. This observation calls for testing to what extent cases of *käkäw* dominate the use of 22A according to temporal distribution. A chi-square test was conducted to compare the observed and expected frequencies of *käkäw* spellings with 22A against those of all other lexemes in the data set with respect to period (Early Classic, Late Classic). The results (Table 7a) point to a statistically significant relationship, characterized by an overrepresentation of 22A applied to *käkäw* during the Early Classic period, with 15 observed attestations, but 6.72 expected attestations. At the same time, it is underrepresented in the Late Classic period, with 15 observed attestations, but 23.28 expected attestations. This is especially striking when one considers that, overall, the MHD contains about 300 spellings of *käkäw* during the Classic period, 49 of which date to the Early Classic, and 251 of which date to the Late Classic. In other words, 30.6 percent of *käkäw* spellings during the Early Classic bear 22A, whereas only 5.9 percent of *käkäw* spellings during the Late Classic period do. The results summarized in Table 7b support the proposition that such difference is significant.

Consequently, it *is* worth repeating that *käkäw* is the most frequent target of 22A overall. This is remarkable for two reasons. First, the use of 22A with *käkäw* starts early, perhaps earlier than for any other *proven* example of 22A. In fact, based on ceramic styles, some of the inscribed pots with **²ka-wa** may predate the earliest dated occurrence of 22A on Tikal Stela 31 (A.D. 445) by as much as two centuries, as in the case of a lidded tripod vessel, catalog #109 in (Fields and Reents-Budet 2005:215), estimated to approximately A.D. 250–400. And second, given the likely motivation for the use of TWO as a duplication diacritic—based on proto-Mayan *ka?= 'two', which is also typically used to mean 'again' or 'twice' when incorporated as an adverbial modifier with a verb (e.g., Yucatec Maya ***ka?a*** *súunajen* 'I returned again') or combined with a numerical classifier for 'times' (e.g., Yucatec Maya *ka?a=téen* 'two times' and *ka?a=máal* 'twice') (Bricker et al. 1998:120–121, 273)—it seems that *käkäw* would have been an ideal first target for

Table 7. Summaries of statistical tests relevant to use of 22A with *käkäw*, Part 1.

	Dependent Variable	Relevant Category	Time	Result[a]	Test	*p*-Value	Effect Size	OSM-1
a.	Lexeme	*käkäw*	Period	Overrepresented in EC	Chi-square test	<0.001	Cramér's V = 0.37 (medium)	Table A51
		Other		Overrepresented in LC				
b.	22A with *käkäw*	Yes	Period	Overrepresented in EC	Chi-square test	<0.001	Cramér's V = 0.27 (small)	Table A52
		No		Overrepresented in LC				

[a]EC = Early Classic, LC = Late Classic.

https://doi.org/10.1017/S0956536123000317 Published online by Cambridge University Press

Table 8. Summaries of statistical tests relevant to use of 22A with *käkäw*, Part 2: Removal of *käkäw* from data set.

	Dependent Variable	Relevant Category	Time[a]	Result	Test	*p*-Value	Effect Size	OSM-1
a.	Broad Phonological Class 1	Obstruents	Gregorian	Lower mean values	Mann-Whitney U test	0.02	$r = 0.28$	Table A54
		Sonorants		Higher mean values				
b.	Narrow Phonological Class 1	All categories	Gregorian	Significant difference with respect to time	Kruskal-Wallis test	0.013		Table A56
		Stops vs. liquids		Stops lower values than liquids	Dunn-Bonferroni tests	0.011		
		Affricates vs. liquids		Affricates lower values than liquids	Dunn-Bonferroni tests	0.06		
c.	Broad Phonological Class 2	Glottalic	Gregorian	Higher median and mean values	Mann-Whitney U test	0.03	$r = 0.3$ (medium)	Table A60
		Plain		Lower median and mean values				

[a]Gregorian = Gregorian calendar years.

22A because of its /$C_1V_1C_1V_1$w/ structure. Unlike a large number of instances where 22A calls for duplication, in which the result is a /C_1VC_1/ sequence requiring that the vowel of the syllabogram not be read "aloud" the second time (e.g., **tz'u-nu**^{2} for *tz'unun* 'hummingbird'), with *käkäw*, the duplication triggered by 22A requires that the vowel be read "aloud" the second time, allowing for a simpler, more straightforward application (with no special "fictitious vowel" rule needed). In other words, the coincidence of ideal structure for the application of 22A, the chronological priority in its demonstrated use of 22A, and the significant overrepresentation during the Early Classic of 22A with *käkäw* all support the possibility that *käkäw* may have been 22A's first target—the lexeme that motivated its innovative use as a duplication diacritic—or at the very least, that it may have been quickly regarded by scribes as the prototypical target of 22A.

What would happen if *käkäw* were removed from the data set? Would there be evidence for a possible influence of its phonological traits (obstruent, stop, nonglottalic), suggesting that the early, frequent use of 22A with *käkäw* could in fact have had an impact and served as a type of prototype for extension of 22A to spellings of other lexemes? Recall that the variable categories obstruents, stops, and plain stops appear to be significantly earlier than other consonant types (Tables 4b, 4c, 4e, 4f, 4g, 4h). To test this possibility, several tests were carried out, and they are summarized in Table 9. As far as Broad Phonological Class 1 is concerned, the data set with reliably dated texts (minus those with *käkäw*) shows a significant difference (Table 8a), supporting the earlier use of 22A with obstruents. Similarly, the data set with all reliably dated texts (minus those with *käkäw*) shows a statistically significant difference across Narrow Phonological Class 1 (manner) categories (Table 8b) and supports the earlier use of 22A with stops, and the distinction between stops and liquids, with stops exhibiting lower mean values. Finally, the data set with all reliably dated texts (minus those with *käkäw*) shows a statistically significant difference across Broad Phonological Class 2 (plain vs. glottalic obstruents) categories (Table 8c) and supports the earlier use of 22A with plain stops, which exhibit lower mean values.

Given the results just presented, and recalling that *käkäw* is an obstruent/plain/stop-initial lexeme, and that it is overrepresented on portable media, the following hypothesis can be proposed:

Hypothesis 8: 22A may have been initially applied with significant frequency to *käkäw*, resulting in the use of 22A, with this target becoming a phonological prototype for scribes, which led to the preferential application of 22A to duplicate syllabograms with initial obstruent/plain/stops, especially on portable media, during the Early Classic.

An interesting example of 22A applied to *käkäw* is seen in two Early Classic spellings of *käkäw* from Rio Azul (Figure 14)—more specifically, from the famous lock-top cacao pot (Stuart 1988). The spellings (Figures 14a–b) are rendered as **ka-**^{2}**ka-wa**. This spelling obviously seems redundant, and as noted at the beginning of this article, the MHD does not even code the presence of 22A in these examples (or it did not when I prepared the data set).

https://doi.org/10.1017/S0956536123000317 Published online by Cambridge University Press

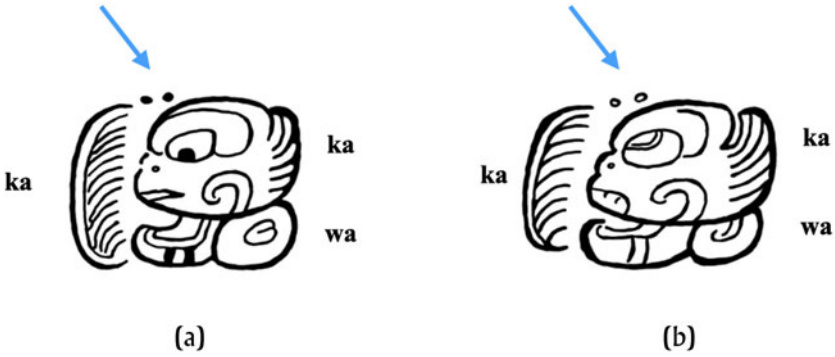

Figure 14. Examples of seemingly superfluous 22A. (a) Glyph C on RAZV15. (b) Glyph F on RAZV15. Drawing by author (after drawings by Stuart 1988:153–157, Figure 30).

Nevertheless, it is worth not only commenting on but also speculating about. This spelling could suggest that the use of 22A was common enough by this time, at least in spellings of *käkäw*, that scribes were beginning to see it as a fixed component of the spelling—in fact, as a type of lexical determinative (recall examples in Figure 2) for the term *käkäw* (Justeson, personal communication, 2022)—perhaps to distinguish it from other possible terms that could be spelled with a sequence **ka-wa** or even **ka-ka**. If so, then the spelling **ka-^{2}ka-wa** could be analyzed as **(ka-)**[$^{\text{KAKAW}}$**ka-wa**], with the initial **ka-** functioning as a phonetic complement (and 22A transliterated as superfixed **KAKAW** to indicate its lexical determinative function). If additional cases of 22A functioning in this manner were found, they could allow for the definition of yet another function, a fifth function (lexical determinative) of 22A.

An alternative is that, by this point, the spelling **^{2}ka-wa** may have been so common that some scribes may have begun to think of it as a (pseudo-)logographic spelling—one that could take a phonetic complement—which would explain the seemingly unnecessary **ka** syllabogram preceding it. This would be consistent with Zender's (1999:122–123) suggestion that in the context of pottery vessels, at least, the **ka** syllabogram by itself (or in conjunction with 22A) may have been understood as a type of logogram for **KAKAW**. In this way, the spelling **ka-^{2}ka-wa** could be analyzed as follows: **(ka-)KAKAW**. This is consistent with other instances in which syllabic or logosyllabic spellings have been observed to function as (pseudo-)logograms (cf. Bricker 1986; Mora-Marín 2010; Tokovinine and Davletshin 2001; Zender 1999) in a process that Matsumoto (2017) has referred to as "orthographic semantization."

The C_1VC_1 shape–marking function and k'ahk' 'fire'

It is a worthwhile exercise to review the question of the C_1VC_1 shape-marking function of 22A. As was mentioned before, lexical roots and stems with C_1VC_1 shapes are relatively rare: 5.3 percent of the approximately 813 etyma reconstructed to proto-Ch'olan and, more specifically, five adjective roots, six transitive verb roots, 31 noun roots, and one verbal noun root. Given that the duplication function of 22A is empirically the earliest attested function, even if the difference among the four proposed functions is not significant statistically; given that such function is also by far the most frequent function of 22A; and given that such function results in sequences of the shape /C_1VC_1(V)/, it is likely that one of the following two processes took place:

(1) Scribes began to spell *k'ahk'* and *tzutz* with 22A applied to the relevant CV syllabogram (i.e., **^{2}k'a** for *k'ahk'*, **^{2}tzu** for *tzutz*), and soon after, they extended the use of 22A to the logographic spellings of such roots.
(2) Scribes began to apply 22A directly to logograms based on roots of the shape C_1VC_1 by analogy with the general duplication function of 22A, which results in phonological /C_1VC_1.../ shapes, without *necessarily* spelling such sequences phonographically with 22A first (i.e., **$^{2}C_1V$**).

The question now is whether the temporal distribution of 22A with spellings for 'to finish' and 'fire' favors either scenario. As it turns out, cases of 22A applying to *logographic* allograms (MR6, AW3) with the value **TZUTZ** for 'to finish', as well as the logogram ZBBa **K'AK'** for 'fire', *precede* instances in which 22A is applied to the syllabogram that would allow for the spelling of the root —ZT1s **tzu** or MZ2 **k'a**, respectively. In the former case, the difference is not great: the earliest spelling of *tzutz* as **^{2}TZUTZ** only precedes the earliest as **^{2}tzu** by 23 years. However, the earliest spelling of *k'ahk'* as **^{2}K'AK'**—though not reliably dated and only assigned an estimated dating of approximately 9.3.0.0.0 (ca. A.D. 495)—does precede the earliest spelling as **^{2}k'a** by well over a century and perhaps close to two centuries.

Still, I decided to test these dated texts to see whether there is a statistically significant difference between them with respect to time, both with and without the estimated date of circa A.D. 495 for the earliest example of **^{2}K'AK'**. This involved testing only cases of 22A applying to *k'ahk'* and *tzutz* to determine whether cases of 22A applied to syllabographic spellings are significantly earlier than those applied to logographic spellings. Although the syllabic targets do exhibit lower mean and median values (A.D. 716, A.D. 692) than the logographic targets (A.D. 719/737.67, A.D. 738/744.5), the results—summarized in Tables 9a and 9b—of two Mann-Whitney U tests applied to all dated texts ($N = 86$) and to all reliably dated texts ($N = 74$) suggest that the answer is negative.

https://doi.org/10.1017/S0956536123000317 Published online by Cambridge University Press

Table 9. Summaries of results of Mann-Whitney U tests applied to cases of 22A with *k'ahk'* and *tzutz* spelled logographically or syllabographically.[a]

	Dependent Variable	Relevant Category	Time[b]	Result	Test	*p*-Value	Effect Size	OSM-1
a.	Grapheme type	Syllabogram / Logogram	Gregorian ($N = 20$)	No significant difference with respect to time	Mann-Whitney U test	0.877	$r = 0.04$	Table A49
b.	Grapheme type	Syllabogram / Monument	Gregorian ($N = 19$)	No significant difference with respect to time	Mann-Whitney U test	0.65	$r = 0.12$	Table A50

[a]Relationship of time (Gregorian, period) to dependent variables in texts with 22A: significant results.
[b]Gregorian = Gregorian calendar years.

This means that both scenarios remain viable, and of the two, I would favor the first scenario, which predicts that spellings such as **^{2}k'a** and **^{2}tzu** should precede spellings such as **^{2}K'AK'** and **^{2}TZUTZ**. This leads to the following hypothesis:

Hypothesis 9: Scribes began to apply 22A to the relevant CV syllabogram (i.e., **^{2}k'a** for *k'ahk'*, **^{2}tzu** for *tzutz*) to spell *k'ahk'* and *tzutz* before extending its use to the logographic spellings of such roots; findings of **^{2}k'a** and **^{2}tzu** temporally preceding cases of **^{2}K'AK'** and **^{2}TZUTZ** should be expected in the future.

Finally, given that *k'ahk'* is the second most frequent lexeme taking 22A, it would be interesting to know whether it may have exerted an influence similar to that of *käkäw* on scribal preferences. Given that it is an obstruent and a stop, like the initial consonant of *käkäw*, the only way to distinguish its possible influence from that of *käkäw* would be in terms of the glottal stricture feature: Broad Phonological Class 2 (glottalic vs. plain). First, all instances of *k'ahk'* were removed from the data set, and then statistical tests were carried out on all reliably dated texts (minus those with *k'ahk'*), and to all dated texts (minus those with *k'ahk'*), including those dated stylistically. The results are summarized in Table 10.

As far as Broad Phonological Class 2 is concerned, the data set with reliably dated texts (minus those with *käkäw*) does not show a significant difference between glottalic and plain obstruents (Table 10a). Nevertheless, the data set with all dated texts (minus those with *käkäw*), including a few dated stylistically, shows a statistically significant difference between Broad Phonological Class 2 categories (Table 10b) and supports the possibility that glottalic obstruents (stops, affricates) can be significantly distinguished from their plain obstruent counterparts (stops, affricates), with cases of 22A applied to glottalic targets being later than plain targets. Consequently, the following more tentative hypothesis can be proposed for future testing:

Hypothesis 10: *k'ahk'* may have served as a second prototype, after *käkäw*, possibly influencing scribes in the application of 22A to lexemes with initial glottalic consonants over time.

Discussion: The evolution of 22A

It is now possible to review the results and hypotheses and to elaborate a model that explains the development of 22A.

(1) 22A originated in the logogram TWO (proto-Mayan *ka?=, proto-Ch'olan *cha?=), as originally proposed by Stuart (2014), probably during the Late Preclassic (300 B.C.–A.D. 200) or early Early Classic (A.D. 200–600) period.
(2) Its initial function was likely the sequential duplication function, not only because it is the most frequent by far of all four functions but because the other three

Table 10. Summaries of results of Mann-Whitney U tests applied to cases of 22A according to Glottalic/Plain consonantal target, excluding *k'ahk'*.[a]

	Dependent Variable	Relevant Category	Time[b]	Result	Test	*p*-Value	Effect Size	OSM-1
	Broad Phonological Class 2	Glottalic / Plain	Gregorian ($N = 44$)	No significant difference with respect to time	Mann-Whitney U test	0.325	$r = 0.15$	Table A58
	Broad Phonological Class 2	Glottalic / Plain	Gregorian ($N = 55$)	Glottalic shows significantly higher values than Plain	Mann-Whitney U test	0.035	$r = 0.29$	Table A59

[a]Relationship of time (Gregorian, period) to dependent variables in texts with 22A: significant results.
[b]Gregorian = Gregorian calendar years.

https://doi.org/10.1017/S0956536123000317 Published online by Cambridge University Press

functions (nonsequential duplication, C_1VC_1 shape marking, abbreviation) can be explained as analogical extensions based on it. The overall chronological distribution also suggests this, even if the statistical tests with dated texts did not point to a significant difference. The origin of this diacritic in the logogram for TWO also supports this, as the term for 'two' can also mean 'twice' or 'again' in the spoken Mayan languages.

(3) It is likely that *käkäw* became the *most frequent* early target of 22A and, as a result, became a functional (sequential duplication), contextual (portable media), grammatical (noun), phonological (obstruents, stops), orthographic (syllabogram), and positional (left locus) *prototype* for the use of the 22A in general.

(4) The lexeme *k'ahk'* 'fire', the second most frequent lexeme to take 22A, *may* have become a second prototype for the use of 22A, leading to the extension of the sequential duplication function of 22A to logograms with C_1VC_1 shapes (*tzutz*, *k'ahk'*, *kuk*), but it remains to be seen whether such application preceded explicit phonetic spellings of such roots.

The absence of 22A in the Postclassic codices suggests the possibility that 22A was employed too infrequently, overall, to survive what might have been a significant bottleneck effect at the end of the Classic period, when text production dropped drastically. It would be worth comparing this decline and cessation to other Classic-period traits that similarly did not survive into the Postclassic period. Perhaps, alternatively, as a result of a bottleneck effect at the end of the Classic period, only one regional scribal subtradition, or a few such subtraditions that already exhibited extremely limited use of 22A (e.g., Northern, Southern, Pasion), continued into the Postclassic period. The obvious choice would be the Northern subtradition, but I suspect that other factors are likely at play.

Conclusions

The most significant conclusion of this article is that the MHD (Looper and Macri 1991–2023) offers many advantages for the study of Epigraphic Mayan, especially for scholars interested in studying broader patterns relevant to the script and its relationship to linguistic and nonlinguistic factors via quantitative methods. If a data set of a mere 125 records can offer insights into the history of Mayan writing, the 5,000 texts and 85,565 records contained within the MHD (which is continually updated) promise to open up new avenues for epigraphic, linguistic, art historical, and archaeological discoveries. This has already been demonstrated by Munson and Macri (2009), Looper et al. (2015), and Munson et al. (2016) years prior to the inauguration of the MHD online.

Another important conclusion of this article is that the statistical methods applied to a relatively small data set were useful in discerning patterns of interest for assessing the evolution of an ancient scribal practice, in terms of not only scriptal and linguistic factors but also temporal, geographic, and contextual ones. The following conclusions can be offered:

(1) For the most part, 22A was applied to syllabograms to trigger sequential duplication, supporting the prior literature on the topic (Kettunen and Helmke 2020; Mora-Marín 2022b; Prager 2020; Stuart 2014; Stuart and Houston 1994; Zender 1999). This function is more strongly correlated with portable media, both because of the high frequency of use of 22A to duplicate the initial syllabogram in the syllabographic spellings of *käkäw* on pottery vessels, and due to the fact that the majority of cases that applied to logograms (especially in the spelling of *k'ahk'* 'fire') occur on monuments.

(2) Four functions of 22A can be supported, with functions (b)–(d) likely derived from (a):
   (a) Sequential duplication
   (b) Non-sequential duplication
   (c) C_1VC_1 shape marking
   (d) Abbreviation punctuation

(3) The most significant temporal, phonological, graphemic, grammatical, and media distribution patterns attributable to 22A can be explained on the basis of its use with *käkäw* and, to a lesser extent, with *k'ahk'*.

(4) It is possible to employ the overall incidence of a lexeme or morpheme with a C_1VC_1 shape in the corpus of texts to estimate the incidence of the application of 22A, characterized by the (practical) linear equation $x = (y - 8.5675)/13.183$, where y = overall frequency of expression in text corpus.

There are still many questions that can be addressed regarding 22A, including its discontinuation sometime after the end of the Classic period. More examples of 22A undoubtedly remain to be included in the MHD, or remain to be discovered, and a larger database will allow us to discern more patterns and, hopefully, carry out further tests of the hypotheses presented here. Perhaps the more obvious next step would be to prepare a database of the more common lexical targets of 22A (e.g., *käkäw*, *k'ahk'*, *k'uk'*, *ʔajawlel*), including cases both with and without 22A, to attempt to discern evidence of factors, such as dispersion (position within a text), that may have influenced variable scribal behavior 22A.

Acknowledgments. I am indebted to Matthew Looper for his answers to several crucial questions and comments regarding the MHD and specific examples. I am also grateful to Nick Hopkins, John Justeson, Barbara MacLeod, Gaby Vail, and the rest of the Q&G participants, for their comments and feedback during and following a presentation on this topic in September 2022. Special thanks to Søren Wichmann for extremely helpful observations and comments on a more recent version of my manuscript. Finally, I want to express my gratitude to the anonymous reviewers, whose feedback allowed me to fill in some important gaps, and to the editorial staff at *Ancient Mesoamerica* for their guidance and patience.

Data availability statement. All data are available as supplementary online materials.

Supplementary Material. To view supplementary material for this article, please visit https://doi.org/10.1017/S0956536123000317

Supporting information.
Appendix 1.
Appendix 2.
Appendix 3.

References cited

Bricker, Victoria
1986 *A Grammar of Mayan Hieroglyphs*. Middle American Research Institute, Publication 56. Tulane University Press, New Orleans.

Bricker, Victoria, Eleuterio Poʔot Yah, and Ofelia Dzul de Poʔot
1998 *A Dictionary of the Maya Language as Spoken in Hocabá, Yucatán*. University of Utah Press, Salt Lake City.

DATAtab Team
2022 DATAtab: Online Statistics Calculator. DATAtab e.U. Electronic document, https://datatab.net, accessed January 12, 2024.

Fields, Virginia, and Dorie Reents-Budet
2005 *Lords of Creation: The Origins of Sacred Maya Kingship*. Scala, London.

Fox, James, and John S. Justeson
1984 Appendix C: Conventions for the Transliteration of Mayan Hieroglyphs. In *Phoneticism in Mayan Hieroglyphic Writing*, edited by John Justeson and Lyle Campbell, pp. 363–370. Institute for Mesoamerican Studies Publication 9. State University of New York, Albany.

Gelb, Ignace
1963 *A Study of Writing*. Rev. ed. University of Chicago Press, Chicago.

Graham, Ian
1978 *Naranjo, Chunhuitz, Unsatanic*. Corpus of Maya Hieroglyphic Inscriptions 2.2. Peabody Museum of Archaeology and Ethnology, Harvard University, Cambridge.

Gronemeyer, Sven
2014 *The Orthographic Conventions of Maya Hieroglyphic Writing: Being a Contribution to the Phonemic Reconstruction of Classic Mayan*. PhD dissertation, Department of Archaeology, La Trobe University.

Harris, John F.
1993 *New and Recent Maya Hieroglyphic Readings: A Supplement to Understanding Maya Inscriptions*. University of Pennsylvania Museum of Archaeology and Anthropology, Philadelphia.
1994 *A Resource Bibliography for the Decipherment of Maya Hieroglyphs and New Maya Hieroglyph Readings*. University Museum of Archaeology and Anthropology, University of Pennsylvania, Philadelphia.

Harris, John F., and Stephen K. Stearns
1992 *Understanding Maya Inscriptions: A Hieroglyph Handbook*. University Museum of Archaeology and Anthropology, University of Pennsylvania, Philadelphia.
1997 *Understanding Maya Inscriptions: A Hieroglyph Handbook*. Rev. ed. University of Pennsylvania Museum of Archaeology and Anthropology, Philadelphia.

Henderson, Leslie
1985 On the Use of the Term "Grapheme." *Language and Cognitive Processes* 1(2):135–148.

Hopkins, Nicholas A.
1994 *Days, Kings, and Other Semantic Classes Marked in Maya Hieroglyphic Writing*. Paper presented at the 93rd Annual Meeting of the American Anthropological Association, Atlanta.

Hopkins, Nicholas A., and J. Kathryn Josserand
1999 *Issues of Glyphic Decipherment*. Paper presented at the 17th Annual University Museum Maya Weekend, "Maya Epigraphy—Progress and Prospects," University of Pennsylvania, Philadelphia.

Houston, Stephen D., John Robertson, and David S. Stuart
2000 The Language of Classic Maya Inscriptions. *Current Anthropology* 41:321–356.
2001 Quality and Quantity in Glyphic Nouns and Adjectives. *Research Reports on Ancient Maya Writing* 47. Center for Maya Research, Washington, DC.

Jones, Christopher, and Linton Satterthwaite
1982 *The Monuments and Inscriptions of Tikal: The Carved Monuments*. University Museum, University of Pennsylvania, Philadelphia.

Josserand, Kathryn, and Nicholas A. Hopkins
2002 Classic Maya Social Interaction and Linguistic Practice: Evidence from Hieroglyphic Inscriptions and Mayan Languages. In *La organización social entre los mayas; memoria de la Tercera Mesa Redonda de Palenque*, vol. II, edited by Vera Tiesler Blos, Rafael Cobos and Merle Greene Robertson, pp. 355–372. Instituto Nacional de Antropología e Historia, México, D. F. and Universidad Autónoma de Yucatán, Mérida.

Justeson, John S.
1986 The Origin of Writing Systems: Preclassic Mesoamerica. *World Archaeology* 7:437–458.
1989 The Representational Conventions of Mayan Hieroglyphic Writing. In *Word and Image in Maya Culture: Explorations in Language, Writing, and Representation*, edited by William F. Hanks and Don S. Rice, pp. 25–38. University of Utah Press, Salt Lake City.

Justeson, John, and James Fox
1989 *Hieroglyphic Evidence for the Languages of the Lowland Maya*. Unpublished manuscript in possession of author.

Justeson, John S., William M. Norman, Lyle Campbell, and Terrence Kaufman
1985 *The Foreign Impact on Lowland Mayan Language and Script*. Tulane University, New Orleans.

Kaufman, Terrence, and William Norman
1984 An Outline of Proto-Cholan Phonology, Morphology, and Vocabulary. In *Phoneticism in Maya Hieroglyphic Writing*, edited by John S. Justeson and Lyle Campbell, pp. 77–166. Institute for Mesoamerican Studies Publication No. 9. State University of New York, Albany.

Kaufman, Terrence, with John Justeson
2003 *Preliminary Mayan Etymological Dictionary*. Electronic document, http://www.famsi.org/reports/01051/index.html, accessed January 12, 2024.

Kelley, David H.
1976 *Deciphering the Maya Script*. University of Texas Press, Austin.

Kettunen, Harri, and Christophe Helmke
2020 *Introduction to Maya Hieroglyphs*. 17th ed. Electronic document, https://www.mesoweb.com/resources/handbook/IMH2020.pdf, accessed January 12, 2024.

Knudsen, Andrew
2023 *The Role of Glyph Blocks in Punctuating Mayan Texts*. Master's thesis, Linguistics Department, University of North Carolina, Chapel Hill.

Kováč, Milan, Eva Jobbová, and Guido Krempel
2016 The Legacy of Early Maya King: Text, Imagery, and Ritual Contexts of a Late Preclassic Cache from Structure H-XVI, Uaxactun, Guatemala. *Mexicon* 38: 9–29.

Lacadena, Alfonso, and Søren Wichmann
2002 The Distribution of Lowland Maya Languages in the Classic Period. In *La organización social entre los mayas: Memoria de la Tercera Mesa Redonda de Palenque, Volume* 2, edited by Vera Tiesler Blos, Rafael Cobos, and Merle Greene Robertson, pp. 275–319. Instituto Nacional de Antropología e Historia, Mexico City, and Universidad Autónoma de Yucatán, Merida.

Looper, Matthew
2007 *Quirigua: A Guide to an Ancient Maya City*. Editorial Antigua, Guatemala City.

Looper, Matthew G., and Martha J. Macri
1991–2023 *Maya Hieroglyphic Database*. Department of Art and Art History, California State University, Chico. Electronic document, https://www.mayadatabase.org, accessed January 12, 2024.

Looper, Matthew G., Martha J. Macri, Yuriy Polyukhovych, and Gabrielle Vail
2022 MHD Reference Materials 1: Preliminary Revised Glyph Catalog. *Glyph Dwellers Report* 71. Electronic document, http://glyphdwellers.com/pdf/R71.pdf, accessed January 12, 2024

Looper, Matthew, Jonathan B. Scholnick, Jessica L. Munson, Yuriy Polyukhovych, and Martha J. Macri.
2015. Measuring grapheme innovation in Classic Maya Writing. Poster presented at the Society for American Archaeology Meetings, San Francisco, CA. April 17.

Macri, Martha J., and Matthew Looper
2003 *The New Catalog of Maya Hieroglyphs. Volume One: The Classic Period Inscriptions*. University of Oklahoma Press, Norman.

Macri, Martha J., and Gabrielle Vail
2009 *The New Catalog of Maya Hieroglyphs. Volume Two: The Codical Texts*. University of Oklahoma Press, Norman.

Martin, Simon, Marc Zender, and Nikolai Grube
2002 *Notebook for the XXVIth Maya Hieroglyphic Forum at Texas March 2002*. University of Texas, Austin.

https://doi.org/10.1017/S0956536123000317 Published online by Cambridge University Press

Mathews, Peter
1980 Notes on the Dynastic Sequence of Bonampak, Part I. In *Third Palenque Round Table, 1978, pt. 2*, edited by Merle Green Robertson, pp. 60–73. University of Texas Press, Austin.

Matsumoto, Mallory E.
2017 From Sound to Symbol: Orthographic Semantization in Maya Hieroglyphic Writing. *Writing Systems Research* 9(2):99–122.

Montgomery, John
2000 The Montgomery Drawings Collection. Electronic document, http://research.famsi.org/montgomery.html, accessed January 12, 2024.

Mora-Marín, David F.
2005 *Affixation Conventionalization Hypothesis: Explanation of Mayan Spelling Practices*. Paper presented at the 43rd Conference on American Indian Languages, Oakland.
2008 Full Phonetic Complementation, Semantic Classifiers, and Semantic Determinatives in Ancient Mayan Hieroglyphic Writing. *Ancient Mesoamerica* 19:195–213. https://doi.org/10.1017/S0956536108000345.
2009 A Test and Falsification of the "Classic Ch'olti'an" Hypothesis: A Study of Three Proto-Ch'olan Markers. *International Journal of American Linguistics* 75(2):115–157.
2010 Consonant Deletion, Obligatory Synharmony, Typical Suffixing: An Explanation of Spelling Practices in Mayan Writing. *Written Language and Literacy* 13:118–179.
2019 Orthographic Conventions in Mayan Writing: Problems with Disharmony and Morphosyllables, Benefits of More Plausible Alternatives. Paper presented at the South-Central Conference on Mesoamerica, Saturday, October 19th, 2019, Louisiana State University, Baton Rouge.
2021 A Survey of Systematic Spelling Practices and Conventions in Mayan Writing. Paper presented at the 13th International Workshop on Writing Systems and Literacy online, "On the systematic nature of writing systems," University of North Carolina, Chapel Hill.
2022a A Study of T670 and Two Lexical Determinatives in Mayan Writing. *Glyph Dwellers* 76:1–23. Electronic document, http://glyphdwellers.com/pdf/R76.pdf, accessed January 12, 2024.
2022b A Possible Collocation Abbreviation Function of the Duplication Diacritic. *Notes on Mesoamerican Linguistics and Epigraphy* 31. Electronic document, https://davidmm.web.unc.edu/2022/12/05/note-31/, accessed January 12, 2024.
2022c Invalidando la hipótesis de (dis)armonía vocálica como convención ortográfica para representar complejidad silábica. Ponencia presentada en el V Seminario Internacional de Epigrafía Maya en Guatemala, 25 de julio del 2022, Universidad del Valle y Museo Popol Vuh, Universidad Francisco Marroquín, Ciudad de Guatemala, Guatemala.
2023 Evidence for Lexical and Phonetic Determinatives in Mayan Writing: The Case of T713. *Ancient Mesoamerica*. https://doi.org/10.1017/S0956536122000335.

Munson, Jessica, and Martha Macri.
2009. Sociopolitical network interactions: A case study of the Classic Maya. *Journal of Anthropological Archaeology* 28:424–438.

Munson, Jessica, Jonathan Scholnick, Matthew Looper, Yuriy Polyukhovych, and Martha Macri.
2016. Ritual diversity and divergence of Classic Maya dynastic traditions: A lexical perspective on within-group cultural variation. *Latin American Antiquity* 27(1):74–95.

Prager, Christian M.
2020 The Sign 576 as a Logograph for KUK, a Type of Bundle. *Textdatenbank und Wörterbuch des Klassischen Maya*, Research Note 15. Electronic document, https://dx.doi.org/10.20376/IDIOM-23665556.20.rn015.en, accessed January 12, 2024.

Robertson, Merle Greene, Marjorie S. Rosenblum Scandizzo, John R. Scandizzo
1976 Physical Deformities in the Ruling Lineage of Palenque and the Dynastic Implications. In *The Art, Iconography and Dynastic History of Palenque, Part III: Proceedings of the Segunda Mesa Redonda de Palenque, December 14–21, 1974, Palenque*, edited by Merle Greene Robertson, pp. 59–86. PreColumbian Art Research, Robert Louis Stevenson School, Pebble Beach.

Sánchez Gamboa, Ángel A., Yuriy Polyukhovych, Marta Cuevas García, Luz de Lourdes Herbert Pesquera, and Guido Krempel
2020 The Hieroglyphic Stairway of House C, Palenque, Chiapas: New Approaches. *Mexicon* 42(1):7–11.

Schele, Linda
1983 *Notebook for the VIIth Texas Maya Meetings Workshop*. Maya Workshop Foundation, University of Texas, Austin.
1990 The Early Classic Dynastic History of Copan: Interim Report 1989. *Copan Notes* 70:1–32.

Schele, Linda, and Nikolai Grube
1994 *Notebook for the XVIIIth Maya Hieroglyphic Workshop at Texas*. University of Texas, Austin.

Schele, Linda, and Phillip J. Wanyerka
1991 *The Proceedings of the Maya Hieroglyphic Workshop, March 9–10, University of Texas at Austin*. University of Texas, Austin.

Smith, A. Ledyard
1950 *Uaxactun, Guatemala: Excavations of 1931–1937*. Carnegie Institution of Washington, Washington, DC.

Stuart, David
1988 The Río Azul Cacao Pot: Epigraphic Observations on the Function of a Maya Ceramic Vessel. *Antiquity* 62:153–157.
2001 A Reading of the "Completion Hand" as TZUTZ. *Research Reports on Ancient Maya Writing*, No. 49. Center for Maya Research, Washington, DC.
2005 *Sourcebook for the 29th Maya Hieroglyph Forum, March 11–16*. Department of Art and Art History, University of Texas, Austin.
2014 "Hieroglyphic Miscellany" from 1990. *Maya Decipherment Blog*. Electronic document, https://mayadecipherment.com/2014/02/25/hieroglyphic-miscellany-from-1990/, accessed January 12, 2024.

Stuart, David, and Stephen Houston
1994 *Classic Maya Place Names*. Studies in Pre-Columbian Art and Archaeology 33. Dumbarton Oaks Research Library and Collection, Washington, DC.

Tokovinine, Alexandre, and Albert Davletshin
2001 Patterned Spellings in Maya Orthography. Paper presented at the Symposium on Early Mesoamerican Writing Systems, March 2001, at the University of Texas at Austin.

von Euw, Eric
1977 *Itzimte, Pixoy, Tzum. Corpus of Maya Hieroglyphic Inscriptions 4.1*. Peabody Museum of Archaeology and Ethnology, Harvard University, Cambridge, MA.

Zender, Marc U.
1999 Diacritical Marks and Underspelling in the Classic Maya Script: Implications for Decipherment. Unpublished Master's thesis, Department of Archaeology, University of Calgary, Calgary.
2004 *A Study of the Classic Maya Priesthood*. Ph.D. dissertation, University of Calgary, Calgary.