The Scribe’s Hand Betrays His Tongue: 
Diglossia Among the Ancient Maya

Mary Kate Kelly 
*Tulane University*

1. Introduction

During the Classic period (AD 250-900), the Maya of Southeastern Mexico, Guatemala, Belize, and Western Honduras institutionalized a hieroglyphic writing system. Classic Maya texts detail political rivalries, rituals, historical events, and mythological characters, through which advances in hieroglyphic studies have revealed a detailed portrait of the lives of the elite Maya that eluded archaeological evidence. It has recently become possible to draw linguistic evidence from the texts. Regardless of the fact that the Maya living in these regions spoke a variety of languages, the glyphs record a largely uniform elite language (Wichmann, 2002). Nonetheless, a growing body of evidence suggests that some language variation appears sporadically in the texts (Colas, 2006; Houston, Robertson, & Stuart, 2000; Hruby & Child, 2004; Lacadena & Wichmann, 2002, 2005; Wald, 2004; Wichmann & Davletshin, 2004). Scribes appear to have been writing in a formal, educated language that was not their native tongue – implying a situation of diglossia. This preliminary research lays the foundation for the research I will pursue for my doctoral degree; namely, to determine the geographic distribution of linguistic features in the inscriptions. In synthesizing scholarship on Mayan languages, the hieroglyphic script, as well as related examples from other language families and writing systems, this research seeks to answer questions of linguistic identity using ancient textual evidence.

2. Background

2.1 Mayan Languages

The languages of the Mayan family, approximately 30 in number, are as diverse as languages of the Indo-European family. Most of these languages are known from modern linguistic research, and many were also recorded in colonial times. Chicomuceltec and Ch’olti’ were recorded by the conquistadores but are no longer spoken today, and undoubtedly many more had been lost prior to their ever being recorded (Campbell & Kaufman, 1985, p. 187). Aside from small pockets of speakers in San Luis Potosi, Tamaulipas, and Veracruz, Mexico, the modern Mayan languages are spoken in a more or less contiguous area spanning Guatemala, the Mexican states of Yucatán, Chiapas,
Campeche, Tabasco, Quintana Roo, as well as Belize, and historically were spoken in the western regions of Honduras and likely El Salvador. This segment of Mesoamerica has long been recognized as the home of the Maya, who are not only closely related linguistically, but culturally as well. Archaeological and ethnohistorical records demonstrate continuity of Maya cultural practices in this region from the Middle Preclassic period (ca. 1000-400 BC) through today (Sharer & Traxler, 2006, p. 177).

However, this region is not one monolithic culture group – it is subdivided into numerous distinct ethnic groups. A fundamental division exists between the ecologically and geographically distinct lowland and highland regions. This study focuses on the lowland regions, where languages historically took part in a highly interactive sphere of cultural exchange, termed by Justeson et al. as the “Lowland Mayan Linguistic Area” (1985, pp. 9-20). It is in this lowland region where Classic Maya civilization thrived throughout the first millennium AD. The Ch’olan-Tzeltalan branch of the Mayan languages occupies the southern lowland region. This branch consists of the Tzeltalan sub-branch, including Tzeltal, Tzotzil, and arguably Tojolab’al, while the Ch’olan sub-branch consists of Ch’ol, Chontal, Ch’orti’, and colonially-documented Ch’olti’ (see Robertson, 1977; Campbell, 2013). To the north, the Yukatekan languages – Yucatec, Mopan, Northern Lakantun, Southern Lakantun, and Itzaj – span the Yucatán peninsula and extend south to northern Guatemala and Belize. Dictionaries and grammars produced by modern linguists as well as colonial documents of these lowland languages allow for the hieroglyphic inscriptions to be placed squarely within the Ch’olan languages, though there is some disagreement as to exactly where in the family tree the writing system falls. Each of these lowland languages has been shown to be engaged, to greater or lesser degree, with the scribal preferences in texts (Lacadena & Wichmann, 2005).

2.2 Mayan Historical Linguistics

As dictionaries and grammars of each of the Mayan languages have become available to the academic community, the understanding of Mayan historical linguistics has made great strides. Early studies focused on the phonological correspondences among Mayan languages in order to reconstruct phylogenetic trees (McQuown, 1956; Robertson, 1977; Campbell & Kaufman, 1985). Once this foundation had been laid based on shared phonological innovation, refinements to the family tree were made based on morphological and syntactic features (Dayley, 1983; Robertson, 1992). Studies offering reconstructions of syntactic structures are relatively sparse, but Hofling (1984) discussed Proto-Yukatekan syntax, and England (1991) based a general reconstruction of Proto-Mayan on the attested syntactic forms in the modern Mayan languages.

As a hieroglyphic interest in morphological structures evolved out of the developments in epigraphy, MacLeod (1987) provided a description of the verb morphology of Ch’olan and Yukatekan languages for students of the hieroglyphs. “An Outline of Proto-Ch’olan Phonology, Morphology, and Vocabulary” (Kaufman & Norman, 1984) was the first extensive effort to explain grammatical features such as split ergativity and complete and incomplete aspect markers, as well as phonological features and their development from Proto-Ch’olan into each of the Ch’olan languages. While some claims made in this paper are still debated today, much of the reconstruction work has provided crucial evidence for further historical reconstruction as well as hieroglyphic decipherments. These reconstructions of proto-Mayan and sub-branches of the Mayan language family have been essential to the decipherment of the glyphs.
However, the incorporation of linguistics in the study of the glyphs is only a recent advancement in the history of the decipherment.

2.3 Hieroglyphic Writing

Discovered among the ruins of ancient Maya sites throughout the lowland region, Maya hieroglyphic inscriptions were first introduced to Western scholars in 1839 and 1840 by explorers John Lloyd Stephens and Frederick Catherwood, and quickly became the focus of much scholarly interest. The texts proved difficult to decipher, and by the middle of the 20th century, the majority of what was legible related to dates, calendrical cycles, and counts of time. It was not until 1952 when Knorosov, a Russian scholar, made the first phonetic decipherment of the glyphs based on the iconography of associated imagery, native Yucatec Maya words, and an “alphabet” of Maya script recorded by Diego de Landa in the 16th century. This groundbreaking finding proved that the script was indeed fundamentally tied to the Mayan languages. Since this early work on the glyphs, much progress has been made in reading the signs such that it is now becoming possible to glean other aspects of language, such as morphological information, syntactic and word order preferences (see Lacadena, 2011) and, as the proposed research aims to do, dialectical tendencies and variation.

The writing system of the ancient Maya is composed of both logographic symbols and phonetic signs. Logograms represent entire words and are often pictographic in nature. Rebus principles may be used to portray abstract concepts with a homophonous word that is more simply illustrated. Glyphs frequently incorporate iconographic elements that are shared between writing and imagery, and may be presented in a variety of forms based on the artistic decisions of the scribe. In addition to whole-word logographic symbols, the texts heavily employ syllabic signs. These may be used to write words that do not have a logogram, or as an alternative spelling for those that can be written logographically. They are also used to phonetically complement logographic signs when the author feels the logogram may need to be specified, or simply for the aesthetics of the glyph block. Syllables are also used to write grammatical suffixes and other elements that are not encoded in logograms. The symbols for syllables represent open consonants of the shape V or CV, although Mayan languages are primarily of closed syllabic structure. Synharmonic and disharmonic vowels in a word written CV-CV can produce the closed CVC syllables, and encode various complex vowels that the CV syllabic elements cannot on their own (Houston, Stuart, & Robertson, 2004). Texts are typically read in double-columns, and the basic word order is VS or VOS, with an ergative-absolutive inflectional pattern on the verbs (Kettunen & Helmke, 2014).

2.4 Linguistic Affiliation of Maya Texts

While the decipherment remains ongoing, the focus of epigraphic work has shifted toward gaining a more profound linguistic understanding of the texts. Originally thought to be tied to Yukatekan languages, it has become increasingly clear that a member of the Ch’olan languages formed the linguistic base for the writing. Generally the language of the writing, termed ‘Classic Ch’olan’ by Wichmann (2006), is linguistically uniform, using identical forms across a broad geographic region, despite the fact that a variety of languages must have been spoken in this region.

However, it has long been evident that regional variation exists in the texts. For example, Yukatekan languages significantly influenced texts of the northern lowlands, and
their effects can be seen in phonetic spellings. Yukatekan languages underwent a *t > ch/CVCV shift while Ch’olan languages retained *t. This shift is visible in an example from Xcalumkin (Column 4), yo-to-che, yotoch, ‘his/her/its house’ using the Yukatekan spelling. This site is located in northern Yucatan, and this spelling reflects the local variety. However, this local variation was not consistently used, as can be seen in nearby Chichén Itzá (Lintel III, Temple of the 4 Lintels) where yo-to-ti, yotoot, was spelled using the standard Classic Ch’olan form. A number of Yukatekan features are evident in northern texts, and yet the majority of forms retain Ch’olan spellings.

During the Late Classic, when the majority of hieroglyphic inscriptions were written, the Ch’olan languages had divided into at least two varieties: Eastern Ch’olan and Western Ch’olan. Teasing these two apart has been more of a challenge than identifying the more diverse differences between Ch’olan and Yukatekan languages. A morphological marker, the positional markers of –laj and –waan, shows clear division along geographic lines between east and west, respectively (Hruby & Child, 2004). Similarly, winal and winik, two terms for ‘man, 20, month’, are divided into east and west (Lacadena & Wichmann, 2002). A number of other markers, such as personal naming preferences (Colas, 2006) and month names (Lacadena & Wichmann, 2005) also show division between Eastern and Western Ch’olan.

Lacadena and Wichmann (2005) provided lexical and morphological evidence for the existence of at least four language varieties during the Late Classic: Eastern Ch’olan, Western Ch’olan, Yukatekan, and Tzeltalan. They enumerated traits that vary in the script and associate them with linguistic varieties, and showed that their geographic distribution matches the expected location of the language varieties based on modern and colonial distribution of languages. One point that they stressed, however, is that the variation in the texts represents elements of the non-prestige languages, not that the texts are written entirely in a non-prestige language. They argued explicitly for an Eastern Ch’olan prestige language that provides the general written standard, and the divergences from this prestige form are relatively few and inconsistent (2005, p. 36).

3. Diglossia

The existence of (at least) four language varieties influencing written Classic Ch’olan suggests that this was a diglossic situation. Diglossia is the coexistence of two or more language varieties in the same language community, and as defined by Ferguson (1959, 1996), one variety carries more social prestige and is used in educated and literary settings, while the other (or others) are used only in informal contexts.

Here, the pertinent feature of diglossia is the existence of “societally held norms governing differential functional allocation of codes” (Hudson, 2002, p. 43). One variety is used as a prestige code, specifically for use in the hieroglyphic writing system and only by the educated elite, and other language varieties are the vernacular languages. The linguistic preferences employed in the texts show sporadic evidence of Yukatekan features in the north, Eastern Ch’olan features in the eastern lowlands, Western Ch’olan features in the western lowlands, and Tzeltalan features near southwestern border of the Maya lowlands, paralleling where these features would be anticipated based on modern and historical homelands of the speakers of these languages (Lacadena & Wichmann, 2005).

However, there is still much left to be discovered about how these linguistic divisions were manifested on the landscape. By mapping all of the traits that bear significance as
linguistic or dialect markers, I will first examine whether these linguistic boundaries appear as clear divisions, with multiple features following the same boundaries, or whether they are a dialect continua, with languages spatially fading into one another. And second, because of the historical information recorded on the texts themselves, I will investigate whether these isogloss borders parallel divisions along known political boundaries. That is, are political affiliations intimately tied to language?

If the goal is to better understand geolinguistic variation in the texts, we must first question in what ways linguistic diversity might be encoded in a written medium. With an understanding of the languages under consideration, it will help to look at how written sources in other areas show these kinds of linguistic diversity while still under one umbrella language used as the basis for the written tradition.

4. Language Variation in Written Sources

As detailed in §2.3, the task of wresting linguistic variation from texts has yielded successful, though preliminary, results for the Maya area. Similar success can be seen in the study of language variation in the written sources of Romance languages and Middle English. Considering these efforts can yield beneficial suggestions for methodology and warnings against specific pitfalls for a dialect study of Mayan texts.

4.1 Written Evidence of Romance Languages

The linguistic situation in Europe during and immediately following the Roman Empire show a number of parallels to the Maya case. Latin was both the prestige language and the vehicle for the writing system, and over time evidence of other languages and dialects emerged in texts as the Romance languages diversified. The first explicit evidence for language variation in the Roman Empire comes from Spain in AD 101 in the Historica Augusta in which the Iberian peninsula native, Emperor Hadrian, is described as having a ‘rustic accent’ in an address to the Roman senate, suggesting that Spanish was beginning to diversify from Latin (Adams, 2007, pp. 231-232). Yet, the texts show very little evidence of this linguistic diversity until the beginning of the 8th century. The persistence of Classical Latin in written sources is likely due to the elevated status of Latin over the vulgar dialects, as well as the small number of literate individuals (Pei, 1976, p. 73). The Appendix Probi from the 3rd or 4th century AD lists spelling rules for scribes, imploring them to use “correct” (i.e. non-vulgar) Latin forms in their writing (Baehrens, 1922). This prescriptivism depicts the direct influence of the standardization of Latin in writing, and also reveals that the spoken form was straying away from the written standard.

Vulgar Latin features began to appear in central Italian texts in earnest during the 8th century, in which phonological and morphological features used by authors were “foreshadowing Italian conditions” (Pei, 1976, p. 83), yet they still use many of the features that typify Classical Latin writing. The first record written in an unmistakably Romance language is the Oaths of Strasbourg of 842, a speech given by Charlemagne’s son, Louis the German, as an oath to the people ruled by his brother, Charles the Bald, in the contemporary language of modern France (Pei, 1976, p. 91). This explicit, intentional use of language not only provides unique insight into the language situation of the development of French language, but also the emerging political rejection of elite, learned language practices associated with Latin and emerging preference for vernacular language in order to gain support, confidence, and allegiance from the unlearned population.
One challenge of deciphering language variation in written sources is “the Latin that survives, being by definition written, consists almost exclusively of the forms of the educated standard, and one is not likely to find much sign of primary dialects in the literary language” (Adams, 2007, p. 14). Yet, it is possible in the study of Romance languages to identify nonconformities and associate them with the incipient forms of the diversifying Romance languages. There are three primary lines of evidence both Pei and Adams used in their determination of language varieties in the inscriptions: phonetic evidence, based on differences in orthography; morphological evidence; and lexical evidence (Pei, 1976; Adams, 2007). This study proposes to follow a similar methodology using these lines of evidence, and, where possible, adding syntactic and onomastic data.

There is one significant way the study of Romance languages differs from Mayan: Latin diffused into Romance languages as a result of the political dispersal of the Roman Empire, and thus the linguistic diversity of the Romance languages was instigated by the spread of Latin language, culture, and political rule. In contrast, the Mayan languages were already diffused before the Classic era, and the writing system evolved within one of the genetically related lowland Mayan languages. We see extensive lexicographical evidence of the influence Ch’olan languages had on the other languages in the region, but Ch’olan held a more distant relationship with Yukatekan and Tzeltalan than Latin did to its daughter Romance languages.

4.2 English Variation in Written Sources

A closer parallel may be made by looking at how the Latin alphabet was employed outside of the Romance language family, in Middle English. In the 1930’s two studies looked at Middle English dialect variation: Oakden (1930) used phonological and morphological evidence from texts and toponyms, and Moore, Meech, & Whitehall (1935) did the same, but without onomastic evidence (Studer-Joho, 2014, p. 6). These studies found that contrary to the highly standardized Old English, Middle English showed much more variation that fell along geographically bound regions. Moore (1957) was able to create isogloss maps marking the boundaries of different dialect features. Later studies by McIntosh et al. (1986) compiled Late Middle English dialect information from all of the texts from known locations and created a database which allowed for previously unprovenanced texts to be fitted to a geographic location according to their linguistic features (Studer-Joho, 2014, p. 9). The same was done for Early Middle English, and used many of the same lines of evidence but also incorporated syntactic and word class evidence into the discussion of the distribution of features (Studer-Joho, 2014).

There are some issues that arise in these studies, as identified by McIntosh (1989), which may be relevant to the application of these strategies to Maya glyphic studies. First, there is an ever-present possibility of standardization eliminating evidence of local variation. Tolkien recognized similarities in the language of two 12th century Middle English texts, *Ancrene Wisse* and *Hali Meidhad*, and proposed that the similarities were due to a shared scribal tradition, likely a “literary center” in which scribes were trained in a uniform written language (Zettersten, 2011, p. 168). This standardization can arise from the institutionalization of writing, and may significantly reduce the amount of linguistic variation recorded in texts. Second, “even in the best of conditions, certain types of spoken language phenomena are not going to be revealed” (McIntosh, 1989, p. 82). Elements such as allophones which would have been represented by the same grapheme will appear more similar than was in fact the case. Standardization was clearly in play in the Maya case, as it was for Old English, but the structure of the Maya system may make it difficult to
identify phonetic variation. Despite these problems, dialects of Middle English were visible in the texts, and it was possible to determine isogloss lines (Moore, 1957) as well as the provenience of many texts (McIntosh et al., 1986).

5. Isogloss Studies

Similar to Moore’s (1957) Middle English study, I propose to use isogloss lines to divide geographic regions based on multiple linguistic features. Two potential linguistic situations can be anticipated from this research. The first would parallel Trudgill’s recent survey of dialects across England in which dialects “merge into one another, without any abrupt transitions” (1999, p. 6). Trudgill focuses on phonological (20+) and lexicographic (10+) features, but incorporates a section on grammatical features including: deixis, pronouns (personal, reflexive, possessive, gender), present and past tense of verbs, negative verb forms, and conjugations of to be. Individual features have distinct isogloss lines delineating the regions of their use, but in combining data from a number of isogloss lines, they do not trend together. This results in a blending of dialects, creating a dialect continuum. Trudgill’s study is by no means exhaustive, but rather uses examples that typify the kinds of features seen in the dialects of England.

A second, more comprehensive study by Labov and Ash (2006) provides an example of the second foreseeable outcome of this dialect study, in which discrete dialect regions of North America are identifiable. The survey used for this study asks participants to answer hundreds of questions in order to elicit vocabulary and sentences. This extensive study produced a series of isogloss maps and resulted in bundles of isogloss boundaries that generally co-occurred. A map of the United States and Canada shows geographically discrete dialects, where the highest concentrations of features co-occur (Labov and Ash, 2006, p. 250). The major regions do themselves have subdivisions, but as whole entities are relatively distinct, as opposed to shading into one another as in the Trudgill study. For the Maya area, either of these two geolinguistic patterns is possible. The objective of my dissertation will be to determine which model is more representative of the Maya case.

6. Methods

As mentioned above, research has begun to reveal language variation in the texts that align with contemporary or historically reconstructible languages. However, these studies are all partial in their analyses, recording only a limited set of features. I will begin with these studies, and collect data on instances of the two dozen or so features already discussed in the literature. I will then proceed to look for new features that have promising implications for representing linguistic diversity, and search for geographic correlations.

A primary assumption of my research is that one prestige language (of the Ch’olan branch) was the principal language used by ancient Maya scribes. This was particularly true through the Classic period, as there was a high overall level of linguistic uniformity among the texts. Aspects of the writing system that show variation are subject to be deviations from this prestige language. Anomalies, alternations, and divergences will be used as the primary evidence for influence from other linguistic groups on the script. This is based on a second assumption: that choices made by scribes that do not comply with the standard form stem from the scribe having a native language that was not the prestige (Wichmann & Davletshin, 2004). Where possible, these anomalies will be directly linked to modern or reconstructed lowland Mayan languages. Isogloss lines and possibly also isogloss bundles will be drawn based on the geographic distribution of these features.
Isogloss maps will help to clarify the linguistic environment during the Late Classic, demonstrating geographic boundaries of the contemporary languages or dialects.

7. Sources

The source material for this research will be limited to texts from the time period of approximately AD 650-850. Only texts that can be securely dated either by the inscription or the context in which the artifact was excavated will be used. Texts from across the Maya area will be considered. As the geographic location of the texts is of central importance in creating isogloss maps, this study will only consider texts that have been excavated, or texts inscribed on artifacts that can be securely traced to an origin at a specific site. Looted texts will be considered if they are known to come from a region, even if the exact site is not known, but they will be noted as of uncertain origin. The material and artifact types will not be limited. All media with inscriptions will be considered – stone, ceramic, wood, stucco moldings, painted murals, bone, and shell being the most prevalent. Artifact types will range from personal jewelry engraved with name phrases to public stone monuments. Each material type and artifact type will be classified and analyzed separately in order to see if there are significant differences in the linguistic variation based on material. Ceramics, for example, as portable and personal objects may show a higher level of regional variation than the more official stone stelae. The analysis will primarily come from studying illustrations of texts available in published materials. In the event that a diagnostic element is not clearly visible in the drawing, the use of photographs and vising the original monument or casts will be considered.

8. Projected Conclusions

Research from the early 2000’s illuminated a situation of diglossia among the ancient Maya – in which writing primarily used a formal, prestige language, and yet scribes’ native languages subtly influenced their writing. In the decade that has passed since Wichmann synthesized this research on the distribution of lowland languages in a comprehensive map (2006, p. 287), several dozen lengthy new texts have been archaeologically excavated – all of them with the potential to shed light on historical language distribution. The nature of such a historical quest inherently has its limitations: a finite corpus that cannot respond to questions, an extinct writing system, and, of course, 1000 years of separation. Looking to successful studies of historical geolinguistics, such as the diversification of Romance languages and Middle English, as well as modern studies of dialects on the landscape provide much insight and direction. I hypothesize that a detailed linguistic analysis of texts from across the Maya region will show statistically significant variation of features, which can be aligned with features of known modern or historically reconstructed Mayan languages. In addition, I hope to provide insight into the role of the state in the process of language institution in diglossic situations. Law remarks, “writing needs the state more than the state needs writing” (2015, p. 178). In essence, it is the infrastructure of the state that allows for the creation and implementation of a writing system, for it could not be a system without institutional support. Language can serve as a powerful tool, and the inscriptions may provide insight into this power structure.

References

Lacadena, A., & Wichmann, S. (2002). The distribution of lowland Mayan languages in the Classic period. In V. Tiesler, R. Cobos, & M. G. Robertson (Eds.), La organización social entre los mayas préhispanicos, coloniales, y modernos: Memoria de la Tercera Mesa Redonda de Palenque (pp. 277-319). Mexico: Instituto Nacional de Antropología Universidad Autónoma de Yucatán.
archaeology's sake, Material culture and style across the disciplines (pp. 32-48). Calgary: University of Calgary.


Department of Anthropology, Tulane University
6823 St. Charles Ave, 101 Dinwiddie Hall
New Orleans, LA 70118
mkelly20@tulane.edu