RESEARCH REPORT

HISTORICAL SOCIOLINGUISTICS OF THE CLASSIC MAYA LOWLANDS: THE **GENERIC PREPOSITION VARIABLE**

David MORA-MARÍN
Institute for the Study of Americas - The University of North Carolina at Chapel Hill (UNC-Chapel Hill)
Chapel Hill, North Carolina, United States

ABSTRACT

This paper applies a historical sociolinguistic framework to the study of Epigraphic Mayan (ISO 639-3 code emy), a logosyllabic writing system from southeastern Mexico and northeastern Central America. The subject is the Generic Preposition variable (conservative *tä*, innovative *ti*) and the focus of the paper is the linguistic and social factors that may have influenced its temporal and geographic distribution, as well as the reconciliation between the epigraphic data and the historical linguistic reconstructions. Two datasets for quantitative analysis were compiled by means of the Maya Hieroglyphic Database (Looper; Macri, 1991–2025). The study takes advantage not only of the absolute dates associated with a majority of the records, but also proxies for social and demographic factors (Text Type, Site Rank Size, Interaction Strategies) that are necessary due to the scarcity of information on the social profiles of the ancient Mayan scribes. The results show the Generic Preposition was an unstable variable, with the innovative *ti* exhibiting a real-time spread from the Southeast region (Copan, Quirigua) to the rest of the Maya lowlands, and that it was likely a sociolinguistic marker ("a change from above"). The results also detect the signature of a significant influence of Ch'olan scribes from the West region, who retained the conservative *tä* variant, on Yucatecan scribes from the Northern region, perhaps even a case of imposition of a Ch'olan superstratum onto a Yucatecan substratum. The paper concludes with a series of desiderata for future research.

OPEN ACESS

The whole content of *Cadernos de Linguística* is distributed under Creative Commons Licence CC - BY 4.0.

EDITORS
– Joe Salmons (UW–Madison)
– Josh Brown (UW-Eau Claire)
– David Natvig (UiS)

REVIEWERS
– Mary Kelly (MRU)
– Joe Salmons (UW–Madison)

Received: 09/03/2024
Accepted: 08/11/2025
Published: 08/29/2025

HOW TO CITE
MORA-MARÍN, D. (2025). Historical Sociolinguistics of the Classic Maya Lowlands: The Generic Preposition Variable. *Cadernos de Linguística*, v. 6, n. 1, e794.

CHECK FOR UPDATES

RESUMEN

Este trabajo aplica un marco sociolingüístico histórico al estudio del maya epigráfico (código ISO 639-3 emy), un sistema de escritura logosilábico del sureste de México y el noreste de Centroamérica. El tema es la variable Preposición Genérica (variante conservadora *tä*, variante innovadora *ti*) y el enfoque del trabajo son los factores lingüísticos y sociales que pudieron haber influenciado su distribución temporal y geográfica, así como la conciliación entre los datos epigráficos y las reconstrucciones lingüísticas históricas. Se compilaron dos conjuntos de datos para el análisis cuantitativo mediante la Maya Hieroglyphic Database (Looper y Macri 1991-2025). El estudio aprovecha no solo las fechas absolutas asociadas con la mayoría de los registros, sino también indicadores de factores sociales y demográficos (tipo de texto, tamaño del rango del sitio, estrategias de interacción) que son necesarios debido a la escasez de información sobre los perfiles sociales de los antiguos escribas mayas. Los resultados muestran que la Preposición Genérica fue una variable inestable, con la variante *ti* exhibiendo una propagación en tiempo real desde la región sureste (Copán, Quiriguá) al resto de las tierras bajas mayas, y que probablemente fue un marcador sociolingüístico ("cambio desde arriba"). Los resultados también detectan el rastro de una influencia prominente de los escribas ch'olanos de la región occidental, quienes conservaron la variante conservadora tä, sobre los escribas yucatecanos de la región del norte, quizás incluso un caso de imposición de un superstratum ch'olano sobre un substratum yucatecano. El artículo concluye con una serie de desiderata para investigaciones futuras.

KEYWORDS

Historical Sociolinguistics; Mayan Writing; Morphological Variable; Logosyllabic Writing; Ch'olan-Tzeltalan; Yucatecan.

PALABRAS CLAVE

Sociolingüística Histórica; Escritura Maya; Variable Morfológica; Escritura Logosilábica; Ch'olano-Tzeltalano; Yucatecano.

> Once we recognize, though, that historical figures such as Augustus were real people, and that the inscriptions cut in stone and preserved for so many centuries reflect the words of real speakers, then it is not so surprising that evidence should be available that demonstrates a linguistic sensitivity on the part of Romans to social differences or that the Uniformitarian Principle should be applicable here and should lead to the view that ancient Rome, in a sense, was no different sociolinguistically from urbanizing situations readily accessible to study today.
> –Joseph and Wallace (1992, p. 117).

INTRODUCTION

This paper investigates the historical sociolinguistics of Epigraphic Mayan (henceforth EMY, after its ISO 639-3 code, emy). EMY is a "logosyllabic" writing system that was innovated and used between *ca.* 400 BCE–CE 1700 by Ch'olan(-Tzeltalan) and Yucatecan speakers, two of the subgroups of the Mayan language family, primarily in the Maya lowlands region (in parts of southeastern Mexico, northern Guatemala, Belize, eastern Honduras).[1] More specifically, as part of a broader project investigating multiple scriptal and linguistic variables (Mora-Marín, 2011, 2017, 2019, 2020, 2021a, 2021b, 2023a, 2025a, 2025b, n.d.), this paper studies the spread of the innovative variant of a morphological variable —the Generic Preposition (GP), *tä (tə)* ~ *ti*— in the Maya lowlands during the Classic period (*ca*. CE 200–900), and its possible sociohistorical associations and motivations.

Historical sociolinguistics can be traced to the articulation of Weinreich, Labov, and Herzog's (1968) "structured" or "orderly heterogeneity," one that is not only linguistically but also socially motivated, and one that considers the problems of transition, embedding, and evaluation of innovations (Romaine, 2005, p. 1696; Roberge, 2006, p. 2310). To investigate such problems, the temporal dimension is crucial. Assuming that "synchronic variation of the type investigated by sociolinguists represents a stage in long term change" (Romaine, 2005, p. 1696), a real-time variationist study would contribute to both sociolinguistics and historical linguistics, and in the process further contribute to our understanding of the sociocultural history of a society. Ironically, Nevalainen and Raumolin-Brunberg (2003, p. 56) have argued, it is such a real-time approach that has been neglected in historical linguistics. Those authors have even defined historical sociolinguistics as "the real-time dimension of sociolinguistics" (Nevalainen; Raumolin-Brunberg, 2012, p. 26), a discipline that could very well fill in the "real-time" gap.

In this regard, EMY texts offers an exciting testing ground for historical sociolinguists: 1) there is a readily accessible and comprehensive online database of EMY texts, the Maya Hieroglyphic

1 In the English literature, Mayanists have long followed the arbitrary convention of using *Mayan* to refer to peoples and languages, and *Maya* to refer to archaeological remains and material culture (e.g. Maya sites, Maya pottery). Epigraphers, scholars who study Mayan writing, typically use the term *Maya writing*. This paper uses *Mayan*, as I treat the ancient writing system not only as evidence of language, but also as a linguistically structured system in its own right, rather than merely an archaeological artifact.

Database (MHD) by Looper and Macri (1991–2025); 2) a majority of texts bear absolute dates correlated with the Gregorian calendar, allowing for precise characterizations of the real-time distributions of variables; 3) though texts tend to be brief, they span a wide temporal and geographic range, allowing for comprehensive treatments of regional developments; 4) despite ongoing debates and uncertainties about linguistic affiliations, historical stages, and orthographic conventions, a great deal of scriptal and linguistic variation is attested in the texts, making for a fruitful corpus for historical sociolinguistics research; 5) few language families of the continent are as thoroughly studied from a historical linguistic perspective as the Mayan language family; and 6) both the linguistic typology of Mayan languages and the logosyllabic nature of the writing system offer a refreshing counterweight to the abundance of Indo-European and alphabetic case studies that are the norm in the field. Thus, EMY texts offer opportunities to investigate real-time variation and change by means of comprehensive datasets for both scriptal and linguistic variables, while representing a refreshing comparative case. These characteristics entice us to focus on the linguistic variables that are most amenable to study in the datasets, the patterns that such variables exhibit, and their association with factors relevant to sociocultural and political processes revealed by the content of EMY texts, rather than factors imposed, a priori, on the basis of contemporary Western social categories. This is, at least in part, what a data-driven approach calls for, as proposed by Lauersdorf (2018, p. 209–210). There exist already highly fruitful examples of data-driven approaches in the Mayanist literature, such as that by Munson *et al.* (2016) and Munson, Looper, and Scholnick (2024), which employ sophisticated quantitative methods to identifies "ritual networks" and diffusion of ritual terms along such networks based on patterns in the hieroglyphic data, though these examples sometimes conflate graphemic and linguistic variables.

Given the aforementioned objectives of the field and the nature of the EMY corpus, this paper has three objectives: 1) to characterize the real-time diachronic and geographic distribution of a the GP morphological variable; 2) to assess to what extent that variable can be correlated with linguistic and social factors, or in the latter case, their proxies; and 3) to reconcile the epigraphic evidence with the results from historical linguistics, as well as the known sociocultural and political processes and events of the Maya lowlands during the Classic. The broader goal of the paper is to illustrate the application of an exploratory historical sociolinguistics framework, and the preparation and quantitative analysis of comprehensive datasets based on the MHD, while testing previous proposals for the temporal and geographic distribution of the GP variable.

The paper is organized as follows. Section (1) provides necessary background to the study of EMY, Classic Mayan society, and the linguistic varieties of relevance. Following this, section (2) introduces the GP variable, the linguistic and orthographic assumptions, and the statistical methods for analyzing them, and the definition of proxies for social factors. Section (3) presents the results of the statistical analyses, beginning with the temporal and geographic distribution of the GP variable, followed by statistical results relevant to linguistic and social factors influencing its distribution. This

section also contains a detailed discussion of some interesting traits of the Northern region with likely significant historical sociolinguistic implications. Section (4) discusses the implications of the results in light of prior research on this variable, the reconciliation with the historical and comparative data. Finally, section (5) offers conclusions and directions for future research.

1. BACKGROUND TO EPIGRAPHIC MAYAN

1.1. CHRONOLOGY, GEOGRAPHY, LINGUISTIC DIVERSITY

Epigraphers divide the history of EMY in three periods, p. Late Preclassic (400 BCE–CE 200), Classic (CE 200–900), Postclassic (CE 900–1521). The Classic period is further subdivided into Early Classic (CE 200–600) and Late Classic (CE 600–900), with the latter –especially the second half of the eighth century CE– constituting the peak of text production (Looper *et al.*, 2015; Looper; Macri, 2022, p. 3). The Terminal Classic (CE 800–950) is another category used in discussions of the decline and collapse of the southern Maya lowland polities. The Postclassic saw a sharp drop in text production, with only a very few stone inscriptions known from this period, and the primary sources being the four surviving paper books, known as *codices* (*códices*).

The present paper applies a regional categorization following that in Munson and Macri (2009, Fig. 5), illustrated in Figure 1. This characterization is not arbitrary: Munson and Macri identified these regions on the basis of frequency of interactions (*i.e.* relative number of interactions among sites), and so, they serve as a preliminary definition of *broad* (multi-site) interaction networks. The following labels will be used: Northern, Central, Eastern, West, Usumacinta, Pasion, Southwest, and Southeast. Given the dearth of data from the Southwest region (with only a single text from the site of Chinkultic represented in the datasets analyzed in this paper), and the high frequency of data from the Southeast region (*e.g.* Copan, Quirigua), the one example from the Southwest has been excluded, and the term Southern has been applied to the Southeast region.

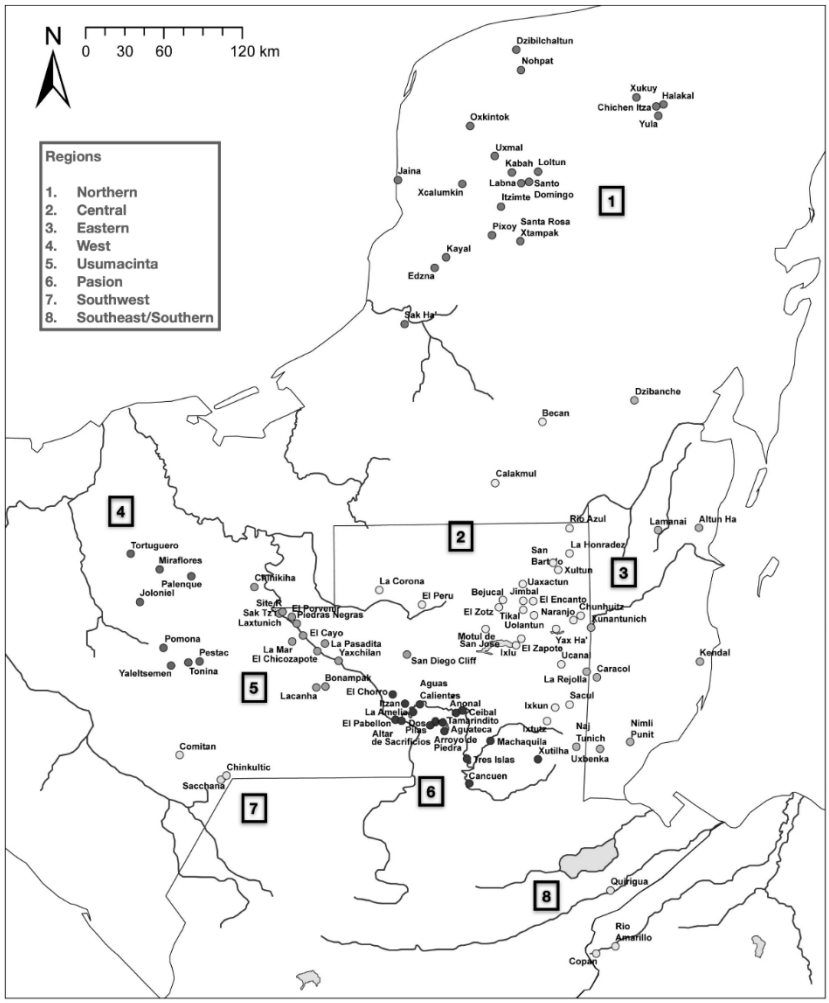

Figure 1. Map of the Maya region. Regional divisions follow those in Munson and Macri (2009, p. 434, Fig. 6b). Used with permission of those authors.

Figure 2, from Josserand (2011, p. 170, Fig. 6.5), presents the distribution *ca.* CE 1500 of all the Mayan languages except Huastec/Wasteko (located far to the northwest, in the Huasteca region of northern Veracruz). The region corresponding to the archaeological Maya sites that can be characterized as part of "Lowland Mayan" society or civilization overlaps primarily with the region where Ch'olan (Ch'ol, Chontal/Yokot'an, Acalá Chol, Manche Chol, Ch'olti', and Ch'orti') and Yucatecan (Yucatec, Lacandon, Itzaj, Mopan) languages are spoken. Nevertheless, as Justeson *et al.* (1985) and subsequent authors have shown, other Mayan languages participated in the Lowland Mayan interaction, resulting in a Greater Lowland Mayan interaction sphere that included also Tzeltalan (Tzeltal, Tzotzil), some Greater Q'anjob'alan (especially Chujean, including Chuj and Tojol Ab'al), and some Greater K'ichee'an (especially K'ichee', Poqom, and Q'eqchi'). During the Classic period the Ch'olan languages likely formed a continuous northwest-to-southeast strip across the

lowlands, with both Ch'olan and Yucatecan speakers along the northern part of the strip, and Tzeltalan speakers in the highlands of Chiapas, in the southwestern part of the strip.[2]

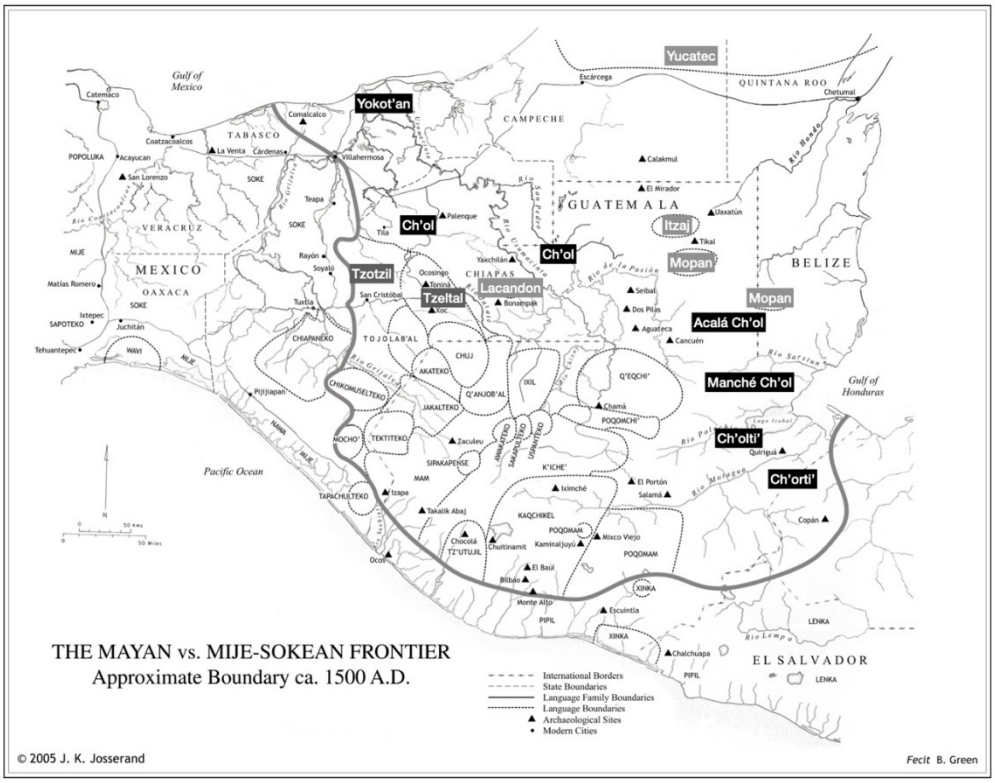

Figure 2. Map of the Maya region showing distribution of Mayan languages *ca.* CE 1500, in relation to some of the major archaeological sites of relevance to this paper. Ch'olan languages are shown in black rectangles, Tzeltalan in dark grey, and Yucatecan in light grey. Used with permission from Nicholas Hopkins, after Josserand (2011, p. 170, Fig. 6.5).

1.2. MAYAN HISTORICAL LINGUISTICS

The classification of the Mayan language family assumed in the present paper is that by Kaufman (1976, 2015, 2017), seen in Figure 3. To understand the lexicon and grammar of EMY texts, the most important subgroups are Ch'olan-Tzeltalan and Yucatecan.[3] Figure 3 also shows the split of the Ch'olan-Tzeltalan into Ch'olan and Tzeltalan. It also shows the split of Ch'olan into Eastern (Ch'olti'-

2 The Yucatecan language known as Lacandón (Lakantun) arrived at its present location during Colonial times, replacing the previously documented Lacadón Ch'ol population (Josserand, 2011, p. 174).

3 I deviate from Kaufman's (2017) terminology and orthography for the names of languages and subgroups slightly. Instead of his Greater Tzeltalan, I will utilize, for the sake of transparency, Ch'olan-Tzeltalan. And instead of spelling Yukatekan, I will employ Yucatecan.

Ch'orti') and Western (Ch'ol-Yokot'an) branches proposed by Kaufman and Norman (1984), supported with additional data in Mora-Marín (2009a, 2009b) and Law (2009). A Ch'olan variety known from a Colonial manuscript called Acalan, closely affiliated with Yokot'an, is also of relevance, but not illustrated in Figure 3.

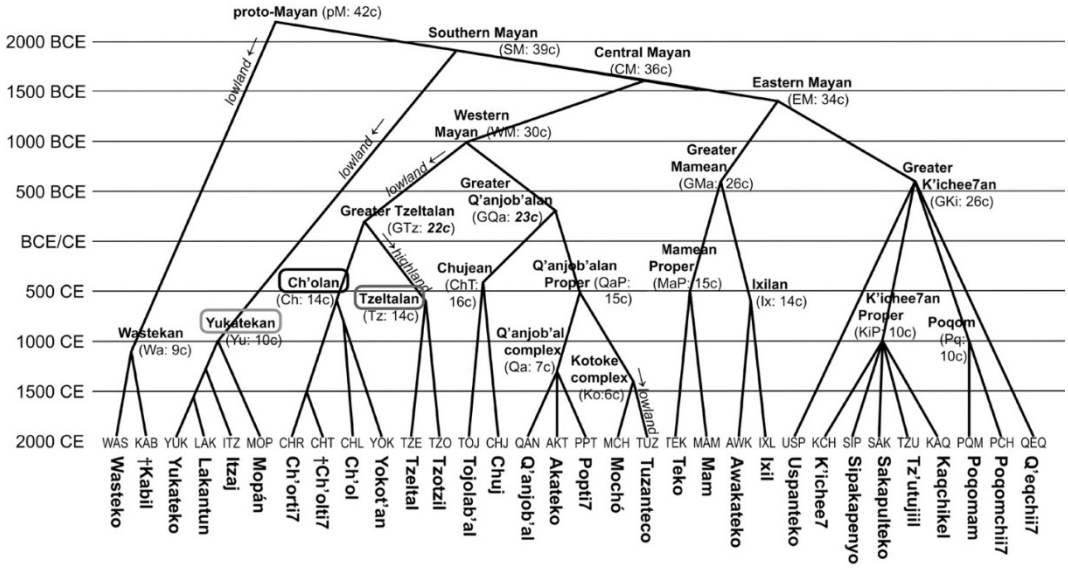

Figure 3. Tree classification of the Mayan languages by Kaufman (2017, pp. 66–67), prepared by John Justeson and available at https://www.academia.edu/37842946/Justeson_Mayan_classification_for_Kaufman_2017_fig_2_pdf.

Kaufman (1976, 2017), Kaufman and Justeson (2007, 2008), and Dahlin, Quizar, and Dahlin (1987) have synthesized a variety of sources of evidence —archaeological, environmental, historical linguistic, lexicostatistic, epigraphic. They generally agree, proposing a differentiation of Ch'olan into its Eastern and Western branches by *ca*. CE 500/600. Dahlin *et al.* (1987, p. 368), correlate this split with the major settlement failures (and associated population movements) that took place during the Terminal Preclassic-to-Early Classic transition (around CE 100–500). Dahlin *et al.* (1987, p. 367–368) further posit another wave of linguistic differentiation events following the Terminal Classic-to-Early Postclassic transition (*ca*. CE 900–1300), following the even more dramatic settlement failures associated with the decline and collapse of centralized rulership throughout the southern Maya lowlands (*i.e.* the Maya lowlands minus the Northern region).

There remains much disagreement among epigraphers regarding the nature of the linguistic varieties that influenced the development of EMY, specifically, whether Classic texts reflect linguistic traits pointing to an undifferentiated Ch'olan language, corresponding to a Proto-Ch'olan stage (Justeson, Fox, 1989; Mora-Marín, 2003, 2009a; Mora-Marín; Hopkins; Josserand, 2005, 2009a), or a post-differentiation variety, whether a Western Ch'olan variety (Hopkins, 1985; Josserand; Hopkins, 2002) or an Eastern Ch'olan ("Classic Ch'olti'an") variety (Robertson, 1998; Houston; Robertson;

Stuart, 2000; Hruby, 2002). Epigraphers generally recognize a high degree of "uniformity" in EMY texts throughout the Maya lowlands, some adopting the concept of a "conservative" or "traditional" basis of EMY writing based on Ch'olan followed by its adoption and adaptation by Yucatecan and possibly also Tzeltalan speakers (Justeson; Fox, 1989), others referring to a "standard" or "prestige" written language based on Eastern Ch'olan referred to as "Classic Ch'olti'an" (Houston; Robertson; Stuart, 2000). Josserand and Hopkins (2002, p. 357) compare the situation in the Maya lowlands during the Classic period to that of "medieval Latin in Europe, where a codified standard was kept from changing while the Latin vernaculars evolved into the Romance family of languages" (2002, p. 358). These authors suggested a diglossic situation was in place, one in which "the older Maya language of Yucatan [Yucatecan] provided a linguistic substratum that was overlaid by a later influx of population that spoke an early form of Cholan Maya" (2002, p. 358); that Yucatecan substratum, Hopkins (1984, 1985) has argued on the basis of morphological traits (*i.e.* ergative and absolutive pronominal agreement markers), influenced the Ch'olan-Tzeltalan superstratum, resulting in the differentiation between Ch'olan and Tzeltalan speakers.

The present paper will not attempt a resolution of the historical stage, the nature of the uniformity of the written language, or the question of superstratum/substratum acculturation. Instead, this paper will offer observations on how the results of the present analysis would be interpreted under a pre-differentiation model versus a post-differentiation model, and if relevant, how they may reflect evidence of such contact between speakers of different varieties.

1.3. LINGUISTIC STRUCTURE OF THE WRITING SYSTEM

Mayan writing reflects the basic structural characteristics of Mayan languages in general: VOS/VOA order in transitive clauses, VS order in intransitive clauses, predicate-initial order in non-verbal clauses, general typological patterning for VO languages (except for the common "exception" of adjectives before nouns), agglutinating word morphology, morphological ergativity (ergative markers on transitive verbs for A arguments, absolutive markers on transitive verbs for O arguments and intransitive verbs for S arguments), some syntactic ergativity (certain constructions apply only to absolutive S/O arguments, excluding A arguments), and evidence of "status" marking on verbs (*i.e.* transitives and intransitives are distinguished by means of portmanteau suffixes that code transitivity, aspect, mood, and main/subordination status all at once), among others.[4]

4 Aissen *et al.* (2017) should be consulted as a comprehensive introduction to the Mayan language family. Mora-Marín (2004) offers a sketch of a few key typological aspects of EMY texts. For grammatical sketches see Bricker (1986, 2004), Law and Stuart (2017), and Baboshkin (2022).

Phonologically, the written language agrees very closely with what is known of early stages of the contemporary Ch'olan(-Tzeltalan) and Yucatecan languages, as systematically laid out by Fox and Justeson (1982) and Justeson and Fox (1989), and supported by many studies since then. Table 1 provides the Proto-Ch'olan sound inventory as reconstructed by Kaufman and Norman (1984), mostly representative of the phonological structure of EMY writing. Nevertheless, the script lacks evidence for a sixth vowel, **ä*, which means that it was probably innovated prior to the shifts of pre-Ch'olan **a:* > Proto-Ch'olan **a*, and pre-Ch'olan **a* > Proto-Ch'olan **ä*.[5] Also, so far, the script lacks evidence of a distinct set of **p'V** syllabograms, suggesting that EMY was innovated prior to the development of */p'/ from instances of both /b'/ and /p/, a development that was likely an instance of areal diffusion involving distinct Ch'olan, Tzeltalan, and Yucatecan speech communities (Kaufman; Norman, 1984, p. 127, 130; Campbell, 1996; Wichmann, 2006).

Consonants						**Root shapes**
p	t	ts <tz>	tʃ <ch>	k	ʔ <7, ’>	CVC CVhC CVjC CVʔV$_i$C
#p’	t’	ts’ <tz’>	tʃ’ <ch’>	k’		
ɓ <b’>		s	ʃ <x>	x <j>	h	
	l					
m	n					
w			y			
Vowels						
i	ə <ä>	u				
e		o				
	a					

Table 1. Proto-Ch'olan sound inventory (Kaufman; Norman 1984, p. 85–89). Angled brackets correspond to the practical orthography from the PLFM for the Mayan languages of Guatemala, and the addition of <7> for the glottal stop by Kaufman (2015), while <#> is used to mark areal diffusion.

EMY texts are visually organized on the basis grids, each grid cell called a "glyph block," a squarish or rectangular arrangements of signs that generally correspond to a word or small syntactic constituent (cf. Knudsen, 2023).[6] Mayan graphemes include three basic types: logograms, syllabograms, and diacritics/determinatives. *Logograms* are graphemes representing lexemes, sets of words based on the inflection and derivation of a specific root or stem. A *syllabogram* is a grapheme with a <CV> value (*e.g.* **Ca** values like **ʔa**, **b'a**, **cha**, **ch'a**, etc.); though most Proto-Ch'olan roots are /CV(:/h)C/ in shape, when a suffix is added, such as a /-VC/ suffix (the most frequent

5 Bricker and Orie (2014) have argued that the alternation between **Ca** and **Ci** syllabograms in some Classic spellings was a means for scribes to represent /ä/. Evidence against this argument is briefly outlined in Section (3.2.2).

6 Mayan graphemes have been cataloged by Thompson (1962) and Looper et al. (2022), the latter a revised version of Macri and Looper (2003) and Macri and Vail (2009). T-numbers will cue codes from Thompson's catalog, and typically alphanumeric codes will cue Looper et al.'s (2022) catalog, which is the catalog employed in the Maya Hieroglyphic Database (MHD) by Looper and Macri (1991–2025). Whenever an inscription is cited, it will be cited with the "object abbreviation" codes utilized in the MHD.

suffix shape), a stem with the syllabification /CV(:/h).C-VC/ results, so that an open <CV> syllabogram is very well suited for syllabification, especially since the complex codas (*i.e.* /CV:C, CVhC, CVjC/) were not distinguished from simplex codas (*i.e.* /CVC/) directly.[7] *Diacritics* or *determinatives* are graphemes that cue a deviation or disambiguation of the value of another grapheme: the duplication dots, a grapheme consisting of two dots, generally tells the reader to read a grapheme twice, and is represented in transliterations by means of a superscript <2> (*e.g.* **ʔAJAW-le²** for *ʔajaw-(a)l-el*); lexical determinatives (Mora-Marín, 2022a, 2023b), more generally known as *semantic determinatives* (Hopkins 1994; Hopkins; Josserand, 1999; Mora-Marín, 2008), combine with a polyvalent grapheme to determine its specific lexical value, and may also be superscripted in transliterations (*e.g.* cartouche**ʔAJAW(AL)** for *ʔajaw(al)* 'Lord (20th day name)'), though most scholars do not transliterate them.

Figure 4 presents examples of EMY spellings, with logograms rendered in uppercase, bold letters, and syllabograms in lowercase, bold letters. **Figures 4a–c** show spellings of the same word, *k'ay-om* (sing-agentivizer) 'singer', with the first spelling (**Figure 4a**) showing a lexographic spelling **K'AYOM**, the second (**Figure 4b**) a lexosyllab(ograph)ic spelling **K'AYOM(-ma)**, and the third (**Figure 4c**) a syllabic spelling **k'a-yo-m(a)**. The logogram stands for the derived stem *k'ay-om*.

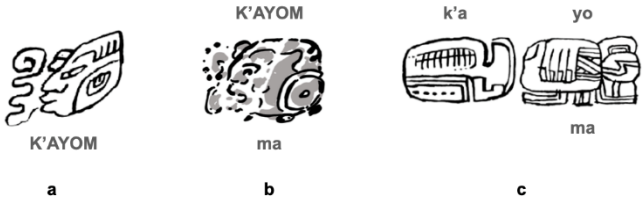

Figure 4. Illustration of logograms and syllabograms, and the use of syllabograms as phonographic determiners ("phonetic complements"). The abbreviations for specific texts correspond to the unique object codes used for the "objabbr" field queries of the MHD (*e.g.* COLK0519). a) Glyph at D6 on conch shell trumpt (COLK0519). Drawing by the author after . b) Glyph J on polychrome pottery vessel from Tikal Burial 196, Structure 5D73 (TIKMT176). Drawing by the author based on photo #8008 by Justin Kerr (http, p. //research.mayavase.com/kerrmaya.html). c) Glyph at A5–B5 on conch shell trumpt (COLK0519). Drawing by the author. Drawings in a) and c) after photograph in Coe (1982, p. 120–123, Fig. 63).

The variety of textual genres in Mayan writing included, in order of increasing grammatical and lexical complexity, the following: object-tags, proprietary statements, dedicatory statements, brief

7 Houston et al. (1998, 2004) and Lacadena and Wichmann (2004) have proposed that root- or word-closing syllabograms could be used to indicate, indirectly, whether the vowel of the preceding syllable was simplex (*i.e.* /V/) or complex (*i.e.* /V:, Vh/). Mora-Marín (2005, 2010, 2022b) has presented arguments against these proposals, including not only a plethora of counterexamples, including systematic ones, but also alternative approaches (*e.g.* obligatory synharmony, consonant deletion of final /-VC/ suffixes, and typical "suffixing" or complementation based on most frequent /-VC/ suffixes) that were previously discussed in the literature (*e.g.* Bricker, 1989; Hofling, 1989; Justeson, 1989; Hopkins, 1997) but not addressed by Houston *et al.* (1998, 2004) or Lacadena and Wichmann (2004).

quotative texts, ritual almanacs, political narratives, and cosmological narratives. Two datasets were compiled for this study: the Generic Prepositions Dataset, which is a comprehensive compilation of the GP variable (across genres), and the Accession Statements Dataset, which consists of political narratives commemorating key events in the political career of a ruler. Whatever their degree of distance with respect to the spoken varieties of the time (Schneider, 2004), such texts exhibit patterned variation, and they should be studied in spite of their thematic, discoursive, and social biases, to the best of our abilities (cf. Nevalainen; Raumolin-Brunberg, 2003:26).

1.4. SOCIAL STRUCTURE AND LITERACY

Maya society, throughout the Maya lowlands and the entirety of the Classic period, displayed a wide variety of social and political organizational structures. At the very least, two distinct groups, elites and commoners, can be differentiated, but at some sites, the distinction in wealth between the low-status elites and the most successful commoners (*e.g.* some military specialists, artists, merchants) may have been blurred (Martin, 2020, p. 325–326), so that at some sites at least, one can speak of a rising "middle class" (Chase, D.; Chase, A., 2004).

By the beginning of the Classic period, the region of relevance was organized into a few dozen kingdoms of varying sizes, each governed by a ruling dynasty based on hereditary kingship with a *k'uhul ʔajaw* 'holy king' at the top. A half dozen long-enduring kingdoms exhibited enormous sway over others through conquests or alliances of various types. Diplomatic strategies included royal visits on the occasion of major events (such as the accession to power of a local ruler) and intermarriages between dynasties. The authority of the royal dynasties and holy kings began to decline and collapse by the end of the eighth century, and in the process the inscriptional record of the southern Maya lowlands came to an end (*cf*. Ebert *et al*., 2014). By the beginning of the tenth century the system had collapsed in the southern lowlands, marking a major depopulation of the major cities, coinciding with population movements to the Northern region, where a different political system, a more decentralized system, took hold during the Postclassic period.

There were communication routes of two types: inland routes, whether terrestrial or riverine, including in the former case road networks, some of them quite elaborate and extensive closer to major centers; and the circumpeninsula coastal route. The Maya region was never under the hegemonic control of a single polity, and yet, the high degree of cultural uniformity across the lowlands indicates an intense level of dissemination of information, including language, writing, artistic styles, etc., as noted by Martin (2020, p. 304–306).

Only about 1.6% of EMY texts contain scribal signatures; of these, only a few provide explicit evidence of the scribes' social profiles (*i.e*. gender, age, rank, place of origin). Generally, scribes were elites, some of them bearing the title *ʔajaw* 'lord, ruler' (but likely with the meaning of 'high-ranking noble'). Several important works pertain to the identification of scribal hands and signatures (Stuart,

1989; Tate, 1994), and the distribution of intrasite and intersite authorship, including the diffusion of scribal art and writing between polities, typically between primary centers and their satellites (Houston, 1993; Montgomery, 1995; Van Stone, 2000, 2005), a topic that has been given a thorough recent review (Houston, 2016), as well as an extremely detailed case study (Matsumoto, 2021). Some scribes and artisans were almost certainly attached to specific kings or dynasties, who served as their royal patrons. Under such patronage, scribes likely functioned as a means of exchange of information between overlords and vassals at different sites, as the evidence appears to indicate for at least some Maya sites (Houston, 1993, p. 135, 2016, p. 403; Martin; Houston; Zender, 2015; Houston, 2016).

These lines of research open avenues for understanding the nature of scribal practices, their institutionalization, and their sociopolitical significance. Nevertheless, for now, this approach is unlikely to provide statistically significant clues to the relationship between the social profiles of scribes and the spread of scriptal and linguistic innovations, except perhaps for a very few sites (*e.g.* Piedras Negras) during a very brief period of time (*e.g.* late eighth century).

It must be assumed that EMY texts were, at the very least, representative of the linguistic practices and ideologies of the uppermost elite groups of Classic Maya society (Justeson, 1985, p. 326–334), in the sense that such groups were the ones commissioning their creation, and also the ones who had a vested interest in their reception among other elites, minimally, and possibly within the larger population, given the likelihood that texts were performed orally and publicly (*e.g.* Houston; Stuart, 1992, p. 591). Extreme evidence that this was the case is provided by the common and recurring practice of destruction of inscriptions at some sites (*cf*. Moholy-Nagy, 2003, 2016). The great investment in the production of art and writing by ancient Maya kings and other elites presupposes the existence of a significant audience, but we simply do not know much about literacy rates during the Classic period (cf. Houston; Stuart, 1992, p. 591–592).

2. ASSUMPTIONS AND METHODS

2.1. SOCIOLINGUISTIC VARIABLES

Linguistic variables are cases of variation, two or more ways of saying "the same thing," predictably constrained by independent factors, linguistic or otherwise. Change, for example the spread of an innovative variant of a linguistic variable, is not abrupt, but instead, a continuum (Chambers, 2013, p. 316), and is characterized by "a period of variation and coexistence between new and old forms in the process of change" (Wolfram and Schilling-Estes, 2004, p. 715), preventing disruptions in communication. The present paper explores both linguistic and non-linguistic factors using a variationist model in an attempt to deal with the transition from one linguistic form to another, its

linguistic embedding, and its social evaluation and embedding (Weinreich; Labov; Herzog, 1968, p. 184; Labov, 1982, p. 27–28, 60; Roberge, 2006, p. 2310).

Due to the paucity of explicit information about the ancient scribes social profiles, linguistic variation can be initially approached on the basis of regional and stylistic variation (cf. Winter, 1999, p. 75). The first step for any such approach is to assume a version of the Uniformitarian Principle, stated for historical linguistics as "the understanding that basic mechanisms of linguistic change in the past (e.g., phonetic change, reanalysis, extension, etc.) were not substantially different from those observable in the present" (Rankin 2003, p. 186), and as "the linguistic processes taking place around us are the same as those that have operated to produce the historical record" (Labov 1972, p. 101). It was reformulated for historical sociolinguistics by Romaine, somewhat vaguely as "the present is the key to the past, the past is the key to the present" (Romaine, 1982, p. 122, 127; 2005, p. 1697), and more concretely as "sociolinguistically speaking, [Uniformitarianism] means that there is no reason for believing that language did not vary in the same patterned ways in the past as it has been observed to do today" (Romaine, 1988, p. 1454). Joseph and Wallace (1992, p. 117) seem to abide by this version of Uniformitarianism in connection with ancient Rome.

This is of course where the historical paradox comes in, as articulated by Labov, "[t]he task of historical linguists is to explain the differences between the past and the present; but to the extent that the past was different from the present, there is no way of knowing how different it was" (1994, p. 21). Given this paradox, a historical sociolinguist who assumes that language in the past exhibited "the same" type of patterning with regard to social factors as in the present must define what they mean by "the same." Also, Labov's resignation ("there is no way of knowing how different it was") seems to negate the validity of any historical enterprise; a historical sociolinguist should instead acknowledge the sources of historical information —social, political, cultural, linguistic— and how they will be used to glean the past. In this regard, a much more constrained discussion of Uniformitarianism as applied to linguistics in general, and historical linguistics in particular, is presented by Walkden (2019, p. 5), who notes that Uniformitarianism is, or should be at best, a methodological assumption, a kind of null hypothesis, one that is open to the possibility of significant differences between the past and the present, or presumably, across different social and cultural contexts. Nevalainen and Raumolin-Brunberg (2003, p. 54) appear to assume such a version of the Uniformitarian Principle, keeping an open mind to major disjunctions; they even highlight a "chief difference between Tudor and Stuart England and the present day: late medieval and early modern Englishwomen did not promote language changes that emanated from the world of learning and professional use, which lay outside their own spheres of 'being'."

Once Uniformitarianism is assumed, heuristically, it can be proposed, following Labov (1972), that stylistic variation may reflect social differentiation in the past, much as it does today, and that such a relationship could offer the means for elaborating a more principled framework for "uncovering social context in historical records" (Romaine, 1982, p. 122–124). In other words, it may

be possible to utilize stylistic variation to infer the presence of variation defined by social factors (cf. Roberge 2006, p. 2311), even if the details of such factors are unknown or unclear or different from their particular permutations in present-day case studies.

A more detailed framework for analyzing linguistic variation has been proposed and elaborated over the years by Labov (1972, p. 314, 1994, p. 78, 2001, p. 196), who defines three types of sociolinguistic variables according to the parameters of social awareness, stylistic variation, and social stratification, p. indicators, markers, stereotypes. These can be characterized as in **Table 2**, generally following Romaine's (1982, p. 265–266) schematization, with examples for each type borrowed from the literature. As argued below, the evidence from the GP variable in EMY texts likely points to a sociolinguistic marker at work.

Type	Social awareness	Social stratification	Stylistic variation	Interpretation	Examples
Indicator	-	+	-	(relatively) ***stable variation*** or beginning of a linguistic change (***change from below***)	Fronting of (ɑ:) in Norwich more common with working class speakers (Chambers; Trudgill, 2004, p. 70–72)
Marker	+	+	+	a linguistic change in progress (***change from above***)	Increase of Postvocalic /r/ in New York City (Labov, 1966); Casual variant of (-ing) in Norwich (Chambers; Trudgill, 2004, p. 70–72)
Stereotype	++	++	++	subject to folklorization and/or stigmatization, can result in reversal of change (***change from above***)	H-dropping in Cockney; *thoidy-thoid* for *thirty-third (street)* in New York City

Table 2. Labov's three major types of sociolinguistic variables (indicators, markers, stereotypes).

These types of patterns, which point to shared communal norms and valuations (Labov, 1972, p. 120–121; Chambers, 2012, p. 300), are not static or fixed, but instead vulnerable to reevaluation and shift, as evidenced in the social re-evaluation of postvocalic "r" in New York City after World War II described by Labov (1972, p. 64–65), as well as the case of "t-glottaling" in Glasgow (Fabricius, 2002), the latter cited in Chambers (20, p. 300). Given the difficulty of assessing social differentiation directly, this paper will pay attention to deviations from expected patterns as clues to possible instances of behavior resulting from social awareness, whatever social factors may underlie it. Labov of course employed the "crossover" phenomenon that he labeled hypercorrection (cf. "Labov-hypercorrection" in Chambers and Trudgill (2004, p. 82)) to confirm the relationship between stylistic variation and social differentiation. As Kerswill (2004, p. 23) notes, "The symptom of change is the "crossover" pattern, by which, in more "monitored" styles [...] the group leading the change exceeds the usage by the next higher group in the social hierarchy." Such unusual or deviant patterns could be identified as evidence of social awareness and socially motivated linguistic behavior.

Instability in the distribution of a variant, whether identified by means of an apparent-time or a real-time approach, is often characterized as a so-called S-Curve pattern, the typical trajectory

inferred (apparent time) or documented (real time) for the spread of an innovation. This S-Curve pattern has been described as composed of three stages by Chambers (2013, p. 312), including initial stasis, rapid rise, and tailing off; Labov (1994, p. 67, 79–83) posits five stages, including incipient (below 15%), new and vigorous (15–35%), mid-range (36–65%), nearing completion (66–85%), and completed (above 85%). The goal in the present paper will be to describe the relative temporal stability or instability of the GP morphological variable, at different regional scales, along with its pattern of spatial diffusion. The regional categories adopted from Munson and Macri (2009) are thus assumed correspond to nested speech communities (Kerswill, 2004, p. 30), and the goal will be to trace the spread of the innovative GP variant, *ti*, assuming that diffusion across space recapitulates diffusion within a social group, with both showing the characteristic stages of change depicted by the S-curve (Bailey *et al.*, 1993, p. 366). Thus, rather than attempting to infer patterns of change by highlighting the first appearances of innovative variants in EMY texts, as attempted by Grube (2004, p. 79–81) for the case of the **h:*j* > /j/ merger, or by Lacadena and Wichmann (2000, 2002, 2005) for the cases of the 'intransitivizer of positionals' and 'abstractivizer of nouns', the present paper will investigate spatial diffusion by means of the overall proportions of the innovative variant of the generic preposition variable in the various subregions of the Maya lowlands, assuming that it takes time for an innovation to spread both within and between communities at various levels.

Lastly, since there is a great deal of information about the historical events and processes that transpired in the Maya lowlands between CE 300–909, such evidence can be adapted to serve as a proxy for social factors, as described in Section (3.4). Additional evidence of general historical processes will also be considered, especially with regard to the discussion of the Northern region in Section (4.4).

2.2. EPIGRAPHIC VARIABLES

2.2.1. TYPES OF VARIABLES AND ORTHOGRAPHIC RESOLUTION

There is no shortage of evidence of variability in EMY texts. The problems lies, at times, in determining what type of variability is at work. Mora-Marín (2019, 2020, 2021a, 2021b, 2022c) has distinguished four types of variables: graphic (different designs of the same grapheme), graphemic (different graphemes with the same value, *i.e.* allography), orthographic (different spellings of the same word), and linguistic (different variants of the same phoneme or same morpheme, for example). The variable of interest in this paper is linguistic, and more specifically, morphological: *tä (tə) ~ ti* 'generic preposition'. Mora-Marín (2020, 2021a, 2021b, 2022c, 2023a) has also introduced a distinction between high and low orthographic resolution variables. *High-resolution variables* are those whose orthographic representation is straightforward, allowing for an unambiguous identification of the phonological shape of each variant. *Low-resolution variables* are those whose orthographic representation is not straightforward due to the common abbreviatory spelling practices of the scribes. The GP variable is thankfully a high-resolution variable: being a grammatical

particle of /CV/ shape, <CV> syllabograms can be used to unambiguously distinguish the two variants, **ta** for *tä (tə)* and **ti** for *ti* (though potentially also *ti?*). It can also be studied as a graphemic variable: five allograms, different graphemes with the same value, could be used to spell it, three allograms with the value **ta** (**Figure 5a**), and two with the value **ti** (**Figure 5b**). However, this paper does not address this graphemic variable, a task left for a future treatment.

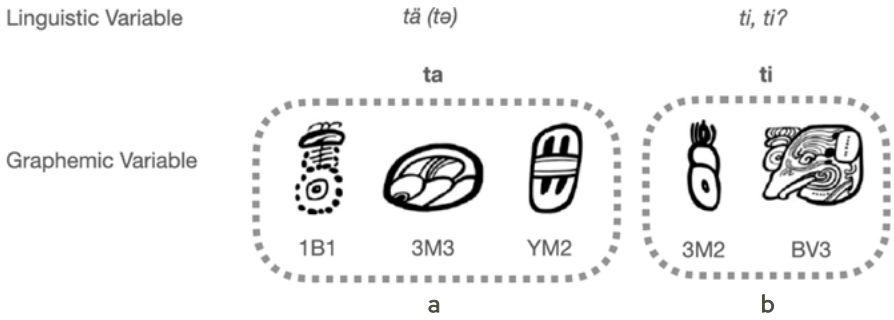

Figure 5. Allomorphs of **ta** and **ti** used to spell the GP variable. a) **ta** allomorphs. b) **ti** allomorphs. Drawings by Matthew Looper (, 1991–2025), used with permission. Alphanumeric codes from sign catalog in Looper *et al.* (2022).

2.2.2. GP VARIABLE

The GP Dataset consists of a total of 1,074 cases of the GP variable, spread across a total of 773 texts; of these, 182 contain two or more cases, and of those, 38 (20.9%) exhibit intratext variation. The Accession Statements Dataset consists of 161 *non*-null cases spread across a total of 119 texts; of these, 18 contain two or more cases of the GP variable, and only one of those exhibits intratext variation. Null cases (10.44% of GP variable cases in Accession Dataset) are those where the scribe omitted the spelling of the GP variable despite its being grammatically required in a particular context.

Based on the data available to them at the time, Kaufman and Norman (1984, p. 81–82) argued that the comparative evidence for this morpheme could not be reconciled with the Eastern Ch'olan/Western Ch'olan differentiation model. The fact is both variants are present in both branches, as the more detailed documentation that followed those authors' work has shown (**Table 3**). It is now known that both variants are widely represented across the Ch'olan languages, though in some cases a variant is preserved only in a highly idiomatic or grammaticalized context. The evidence now suggests, as proposed here, that Proto-Ch'olan had **tä ~ *ti*.

Language/Stage		GP	Sources
Proto-Mayan		**tya*	Kaufman and Norman (1984, p. 139), Kaufman (2015, p. 707)
Proto-Yucatecan		**ti?*	This paper based on data in Hofling (2017)
Proto-Yucatecan		**ti(?)*	Mathews and Justeson (1984, p. 227)
Proto-Ch'olan		**tä*	Kaufman and Norman (1984, p. 139), Mora-Marín (2009a, p. 127)
		**tä ~ *ti*	This paper
Western Ch'olan	Ch'ol	*ti; tyä*	Kaufman and Norman (1984, p. 82), see Wald (2004, p. 34) for idiomatic *tyä*
	Acalan	<ta>, <ti>	Smailus (1975, p. 168, 171)
	Yokot'an	*tä*	Knowles (1984, p. 232)
Eastern Ch'olan	Ch'olti'	*ti*	Kaufman and Norman (1984, p. 82)
	Ch'orti'	*ta, ti*	Kaufman and Norman (1984, p. 82), Mora-Marín (2009a, p. 127)
Proto-Tzeltalan		**ta*	Kaufman (1972, p. 117)
Proto-Ch'olan-Tzeltalan		**ta*	Justeson (1985, p. 470)

Table 3. GP variable attestations and reconstructions.

Table 4 presents the GP variable as attested in EMY texts (cf. Figure 5). It is a clear example of a high-resolution variable, though it is possible that in the Northern region **ti** may have been intended to spell Proto-Yucatecan **ti?*, in which case the final /ʔ/ would not have been made explicit.

Variant	Spellings	Graphemes	Frequency	%
tä	**ta**	1B1, 3M3, YM2	545	50.7%
ti	**ti**	3M2, BV3	529	49.3%
			1074	100%

Table 4. GP Variable as high-resolution variable. Prepositions Dataset (no null cases).

The *tä* (*tə*) variant is earlier than the *ti* variant, appearing in Late Preclassic (400 BCE–CE 200) texts between *ca.* 100 BCE–CE 120, originally spelled with T51/T53/3M3 **ta** (Mora-Marín, 2001, p. 167, 248, 267, 282–288). The earliest dated examples of innovative *ti* are found on Tikal Stela 4, dated to CE 379, and the Tikal Ballcourt Marker, dated to CE 416, both of which are also the earliest cases of intratext variation between *tä* and *ti*. The two cases of the GP variable on the Ballcourt Marker are seen Figure 6, where it appears as *tä* (Figure 6a), and as *ti* (Figure 6b). A few decades prior to this, also on a text from Tikal (Stela 39) dated to CE 376, the first confirmed example of the syllabogram **ti** in a purely syllabic function is found, in the spelling **ʔu-ʔUH(T)-ti** for the verbal expression *ʔu[h]t-i-Ø* (finish[mediopassive]-completive:intransitive-third.person.singular.absolutive) 'it got finished/made; it happened'.

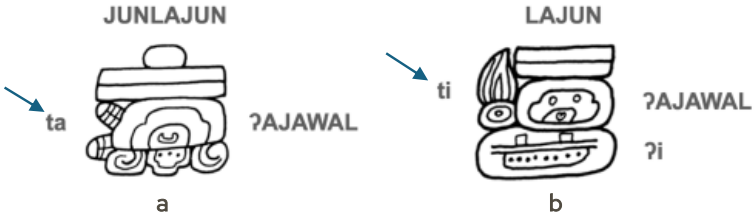

Figure 6. Examples of prepositional temporal phrases headed by the GP variable. a) TIKBCM:E01. Excerpt from drawing #2059 by Linda Schele (http://research.famsi.org/schele.html). b) TIKBCM:F07. Excerpt from drawing #2059 by Linda Schele (http://research.famsi.org/schele.html).

Bricker and Orie (2014, p. 197–198, Fig. 4) have proposed that ancient scribes, like the later Colonial Yucatec and Acalan (Yokot'an) scribes, may have alternated between **ta** and **ti** as a means of attempting to indicate the vowel /ə/: thus, those authors would analyze the **ta** ~ **ti** variation as spellings of a form *tä (tə)*. If so, such alternations would instead point to the Proto-Ch'olan change of ***a* > **ə*, a fascinating possibility. Nevertheless, most other items where Proto-Ch'olan */ə/ was expected are represented exclusively with **Ca** syllabograms (*e.g.* **ya-k'a-wa** for *y-äk'-aw-Ø* 's/he gives/gave/put it', never **yi-k'a-wa***; **b'a-la-ma** for *b'ahläm* 'jaguar', never **b'a-li-mV***; **ka-ka-wa** for *käkäw* 'cacao', never **ki-ki-wV***; **ma-ka** for *mäk* 'to cover', never **mi-kV***; **pa-ta-wa-ni** for *pät-wän-i-Ø* 'it formed', never **pi-ti-wa-ni*** or **pi-ti-wi-ni***; **ʔu-tz'a-pa-wa** for *u-tz'äp-aw* 's/he planted it', never **ʔu-tz'i-pi-wa***; **ta-ta** for *tät* 'thick (liquid)' never **ti-ti***, etc.). This fact supports the notion that the **ta ~ ti** alternation could simply be representing the expected *tä (tə) ~ ti* alternation.[8]

The epigraphic scholarship on this variable is significant (Mathews; Justeson, 1984, p. 187–203, 221–223, 226, 229; Justeson, 1985, p. 470; Justeson; Fox, 1989, p. 15–16, 24–25; Macri, 1988, 1991, 2021; Carter, 2009, p. 6–8, 17–21; Kelly, 2022, p. 101–107). Justeson (1985, p. 470) had already argued for the earlier use of *tä* relative to *ti*, and following Mathews and Justeson (1984), supported the notion that *ti* was likely diffused, likely from Yucatecan. Macri (1988) had observed a strong preference of **ta** spellings in the at Palenque and Tortuguero, with Carter (2009) agreeing and adding Tonina to the group, and Kelly (2022, p. 101–107) further supporting this distribution. More will be said below, in Section (4.4), regarding the distribution of this variable in the Northern region, particularly in connection with Macri's (2021, p. 11) and Kelly's (2022, p. 101–107) observations of the frequency of *ta* in that region despite the fact that *ti* would be expected to be canonical, given the exclusive presence of *ti?* among the Yucatecan languages. Lastly, Carter (2009, p. 20–21) has also suggested that innovative *ti* may have spread due to the influence of the Kan Dynasty (Snake Kingdom), following up on Lacadena and Wichmann's (2002, p. 309–310) suggestion that this dynasty promoted the spread of Western Ch'olan traits in particular. More recently, Kelly (2022, p. 239–243) also has examined the possibility of a prominent role by the Snake Kingdom in the spread

8 Another objection to the Bricker and Orie's (2014) scenario is that most of the examples of **Ca** ~ **Ci** spelling alternations from Classic inscriptions are cases of CV syllabograms used in word-closing spellings in which the vowel of the syllabogram was either not read ("silent") or represented a vowel-initial suffix that is *not known* to have had a value /ə/. For example, Bricker and Orie (2014, p. 201–202) argue, assuming the Commutativity Principle whereby syllabograms with CV values are analyzed by epigraphers as representing VC values (Closs, 1986), that the **-wa** and **-wi** spellings of antipassive inflections represented a suffix *-əw*. Nevertheless, no Mayan language attests to a suffix *-əw* 'antipassivizer'. Instead, a few Mayan languages attest to the form *-aw…* (e.g. Tzeltal, Tzotzil) or *-o(w)* (K'ichee', Tz'utujil, Kaqchikel, Q'eqchi') or *-w* (Kotoke, Tojol Ab'al, Chuj, Teko, Mam, Awakatek, Ixil, Uspantek, Poqomchi'), and multiple Mayan languages attest to the suffix sequences *-w-a ~ -w-i* (e.g. Popti', Q'anjob'al, Akatek) (cf. Mora-Marín, 2001, p. 393–397; Kaufman, 2015, p. 324–325). Other than the case of **ta** and **ti** signs representing the GP variable, none of Bricker and Orie's examples are cases of CV syllabograms spelling the initial /CV/ sequence of a root or stem known to be reconstructible to Proto-Ch'olan as */Cə…/. Last, Mora-Marín (2005, 2010, 2022b) has presented arguments against the Commutativity Principle, following Stuart's (1987) approach.

of linguistic and orthographic traits, though not specifically the GP variable. This idea that will be reviewed and discussed in Section (3.3).

This paper supports prior suggestions that the *ti* variant in Ch'olan may have been innovated as a result of influence from Proto-Yucatecan **ti?* (cf. Mathews; Justeson, 1984, p. 187–203), but not necessarily as a direct loan, since Ch'olan speakers should have easily borrowed such a form as /ti?/. I offer two alternatives to account for this discrepancy: 1) perhaps it was borrowed as *ti* to avoid homophony with Proto-Ch'olan **ti?* 'mouth; speech'; and/or 2) perhaps it was borrowed as *ti* because Yucatecan scribes were spelling it with **ti**, and thus, Ch'olan scribes may have borrowed it through the filter of spelling pronunciation. In either scenario, this form can be added to the inventory of grammatical morphemes that Hopkins' (1984, 1985) proposed Ch'olan-Tzeltalan superstratum borrowed from the Yucatecan substratum, facilitating the linguistic differentiation of Ch'olan from Tzeltalan.

2.3. QUANTITATIVE METHODS

Descriptive and inferential statistics have been employed in this paper, the latter type with the goal of determining whether certain variables exhibit a statistically significant association with each other that could point to influential/predictive factors. The inferential tests include hypothesis and correlation tests (*e.g.* parametric and non-parametric, including Analysis of Variance, Hierarchical Cluster Analysis, Friedman Test, Kruskal-Wallis Test, Spearman Correlation, Mann-Whitney U-Test, Logistic Regression), and almost all have been carried out with DATAtab (DATAtab Team, 2025), but a very few with StatPlus for Mac. Initially, to assess the likelihood of a relationship between a linguistic variable (nominal) and one of the potential independent variables, a Chi-Square Test of Independence (nominal vs. nominal), Kruskal-Wallis Test, Pearson Correlation, or Mann-Whitney U-Test was carried out. If fruitful, the independent variables in question would then be used in a Logistic Regression analysis, to assess to what extent, if any, such independent variables were influential in the distribution of each linguistic variable when considered at the same time with other independent variables. The Logistic Regression summaries presented below are interpreted on the basis of each independent factor: the summarized results indicate which categories (*e.g.* portable or monumental) of an independent variable (*e.g.* Text Type) were more influential on the dependent variable (GP variable), and if significant (p-value $\leq .05$), whether it the influence was positive or negative (Coeff. B), and what the odds (Odds Ratio) are favoring that category over the reference category. This paper reports primarily the results from this last step.

In addition, to illustrate the distribution of variables with respect to time, measured in Gregorian years based on correlations between the Mayan calendar and the Gregorian calendar, raw frequencies per arbitrary units of time (50 Gregorian years) were used to produce charts showing combined relative cumulative frequencies over time. (A future study could attempt to calculate more appropriate periodizations according to the amount of data.) This is preferred over raw frequencies to make up for

the temporally imbalanced inscriptional record (cf. Munson; Macri, 2009, p. 430, Fig. 3b). The cumulative frequencies are proportional, allowing one to compare across regions more faithfully.

2.4. PROXIES

Text Type (portable vs. monumental) will be used, preliminarily, as a proxy for a combination of style (*i.e.* "formal" vs. "informal") and register ("official" vs. "unofficial"), with portable texts likely reflecting less formal and less official language, and monumental texts more formal and more official. The difference may have to do with intended audiences, with many or most portable texts (*e.g.* pottery vessels, jade beads and necklaces, etc.) likely having the intended owner as its audience, and many or most monumental texts intended for a much wider audience (from a few nobles to hundreds or thousands of people). Mora-Marín (2009b, 2025a, 2025b, n.d.) has presented preliminary evidence regarding two lines of evidence: 1) the Proto-Ch'olan reflex of Proto-Mayan **ha?-* 'demonstrative pronoun base' is attested in EMY texts with the highly conservative form **ha?-*, mostly on monumental texts (N = 65) but also in a few portable texts (N = 7), suggesting these forms were more formal, akin to a standard, while the innovative form **hin-*, reconstructible to Western Ch'olan, is attested exclusively on a very few portable texts (N = 6), in most of these appearing in passages representing quoted speech, suggesting that such forms were perhaps more informal, akin to vernacular; and 2) accession statements, referring to the ascent of an individual to a political office (mostly 'kingship'), are found in 8.2% of monumental texts (more than the expected 5.7%) and only 2.4% of portable texts (less than the expected 5.7%), a difference that is statistically significant according to a Chi-Square Test of Independence (χ^2 = 75.87, N = 4865, p = < .00001), suggesting that monumental texts were perceived as more "official" and thus worthy of accession statements than portable texts. If correct, there is a good chance that such a distinction would have correlated also with some sort of social distinction among speakers and scribes (cf. Chambers; Trudgill, 2004, p. 70), though we can only speculate at this point as to what that might have been.

Given the role of demographic factors like population size and density in the spread of innovations (Trudgill, 1974; Bailey *et al.*, 1993; Chambers; Trudgill, 2004; Wolfram and Schilling-Estes, 2005; Britain, 2012), it would be useful to incorporate such information for Maya sites. However, such parameters are exceedingly difficult to assess for ancient settlements, for many reasons, including the fact that settlements change in size and density over time. As a proxy for population size and/or density, the paper uses the site Rank Size classification of ancient Maya sites by Brown and Witschey (Brown; Witschey, 2001, 2002; Witschey; Brown, 2010, 2025; Hausman, 2013). Rank Size is "a summary of the overall size and internal complexity of a site, including the volume and elaboration of the civic and ceremonial architecture, which is usually easier to document than the boundaries of the settlement, and the diversity of buildings and structures (temples, palaces, ball courts, causeways, marketplaces, and so forth, as well as residences)" (Witschey; Brown,

2025, p. XIX). Six Rank Sizes (1–5, 7) are part of this categorization, of which only Ranks 1–3 are of interest here. This paper uses the categories "Top" for Rank Size 1 ("rare and very large" and "housed large populations"), "Middle" for 2 ("smaller cities or large towns"), and "Bottom" for 3 ("villages"). This index is static, unfortunately: it does not take into account the historical development of a site. At the very least, though, it takes into account differences in their final states: some sites ended their primary period of occupation as Rank 1 or Rank 2 or Rank 3, so that presumably a Rank 3 site was not a Rank 1 before, though of course, a Rank 1 site may have started out as a Rank 3, increased to Rank 2, and then Rank 1 over time.

The author also follows the lead of Munson and Macri (2009) and Munson *et al.* (2014), who made use of a set of glyphic expressions referring directly or indirectly to interactions between polities. Such interpolity interaction expressions were first employed systematically to understand the macro-politics of the Maya lowlands by Marcus (1973, 1976), and subsequently, to a much greater extent, by Simon Martin and Nikolai Grube (Martin; Grube, 1994a, 1994b, 1995, 2000, 2008; Grube; Martin, 1998; Martin, 2020). Munson and Macri's (2009, p. 428, Table 2) classification schema of four types of interactions (antagonistic, diplomatic, lineage, and subordination) is adopted here, with some limitations and modifications, aimed at simplifying the quantitative analysis. This paper distinguishes between diplomatic ("co-presence") and non-diplomatic ("hierarchy" and "conflict") interaction types, as well as the absence of either type ("none"). This simplified dichotomy could be considered parallel to Blanton *et al.*'s (1996) inclusionary vs. exclusionary political strategies. Figure 7 presents the categorization employed in this paper, though two of the expressions of "conflict" that were considered are not illustrated (*i.e. u-chan/kan(-ul/al)* 'his/her captor/guardian' and *jub'-uy-i-Ø* 'it became fallen (it fell/it was felled/it was brought down)'). The categorization arrived here is convenient, but not unproblematic. Some authors may differ on how they would categorize some of these expressions with regard to the diplomatic/non-diplomatic or some other dichotomy.[9] This paper opts for interaction strategies, which refer to general types of events that allude to interpolity interactions, instead of explicit alliances (*e.g.* between Snake Kingdom and Caracol, for instance), because such alliances were not static, fluctuating sometimes rapidly over time, while the interaction strategies are more general.

9 Some might argue that *u-kab-(i)j-iy-Ø* 's/he oversaw it' should be categorized as diplomatic. However, from context it is clear that the 'overseers' were invariably politically dominant, which means that the expression conveys hierarchy, one of the two criteria for classifying an interaction type as non-diplomatic.

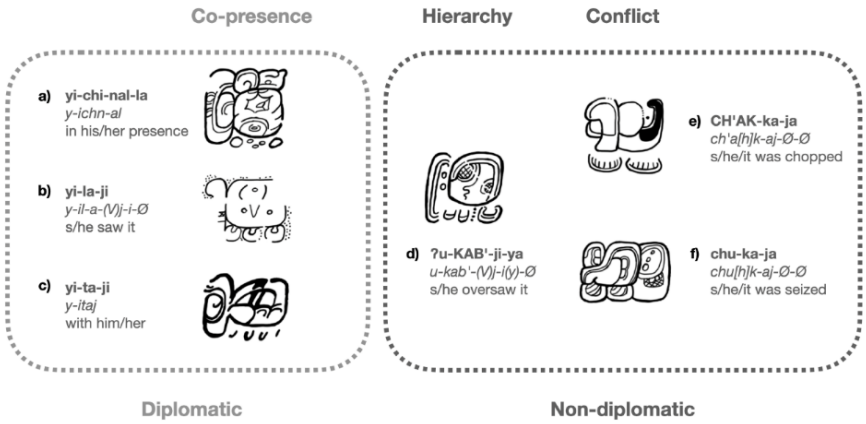

Figure 7. Proxies for social factors: sociopolitical strategy types. Three major types of interpolity interactions are categorized into two major types of sociopolitical strategies: diplomatic and non-diplomatic. Drawings by the author.

The consideration of interpolity interaction strategies could be construed as a combination of a variationist and interactional sociolinguistics approaches, assuming a relationship between political ideology (*e.g.* inclusionary vs. exclusionary) and linguistic choices. In this regard, this paper constitutes a test of this approach, and the results presented below should be considered tentative.

2.5. DATASETS

Datasets were compiled by means of queries using the Maya Hieroglyphic Database (MHD) by Looper and Macri (1991–2025), which contains records from 4,865 inscriptions from the Late Preclassic, Classic, and Postclassic periods. The GP dataset is comprehensive: all Classic-period cases of the GP variable in the MHD were collected. The Accession Statements dataset is likewise comprehensive: all cases of accession predicates were collected, but for the present purposes, only the records containing the GP variable are relevant. The Northern Ch'olan Terms Dataset will be used for the analysis of Ch'olan influence on scribal practices in the Northern region in Section (3.3). The process of dataset curation involved culling: each record was checked against drawings or photos of the inscription, whether available in the MHD itself or elsewhere, and either corroborated or culled if not. Table 5 summarizes the datasets compiled along with their metadata and target variables, whether dependent or independent.

Variable	Type	Values
a) *All Datasets*		
Region	Nominal	West, Usumacinta, Pasion, Central, Northern, Eastern, Southern
Gregorian	Metric/Interval	Gregorian years
Period	Ordinal	Late Preclassic, Early Classic, Late Classic, Terminal Classic
Text Type	Nominal	Monument, Portable object
b) *Generic Preposition Dataset* (All = 1,075, Dated = 681)		
Preposition	Nominal	tä, ti
Context Narrow	Nominal	Complementizer (Progressive), For (content), In/As (Status/Condition), Locative, Oblique Case, Temporal, With (Company), With (Instrument)
Context Broad	Nominal	Complementizer, Locative, Oblique Case, Calendar/Temporal, With
c) *Accession Statements Dataset* (All GP cases = 181, Dated GP cases = 161)		
Rank Size	Ordinal	Top, Middle, Bottom (and Unknown)
Preposition	Nominal	tä, ti, Ø
Interactions	Nominal	Dipl(omatic), Nondipl(omatic), None
d) Northern Ch'olan Terms Dataset (All = 38, Dated = 27)		
Ch'olan Term	Nominal	chab', chahk, chan, tun
Spelling Type	Nominal	Logographic, Logosyllabic, Syllabic

Table 5. Relevant variables for each dataset.

3. RESULTS

3.1. TEMPORAL AND GEOGRAPHIC DISTRIBUTION

Figure 8 presents charts summarizing the Classic distribution of the GP variable, showing an increase of *ti* relative to *tä* over time. Figure 8a shows the raw frequencies across all regions. Figure 8b highlights the mean and median values for the two variants in dated texts. Figure 8c provides the combined relative cumulative frequencies for the two variants in all regions, while Figure 8d excludes the Northern region, which experienced a kind of reversal pattern, with conservative *tä* recovering and eventually outcompeting innovative *ti*, as described in more detail below. The overall pattern, an S-Curve, shows the innovative variant reaching a proportion of 60% of the total cases of the GP variable by the end of the Classic period, having overtaken conservative *tä* during the first half of the eighth century. As a linguistic variable characterizable as change-in-progress, it may have been associated with social factors, and if so, it may have functioned as a sociolinguistic marker.

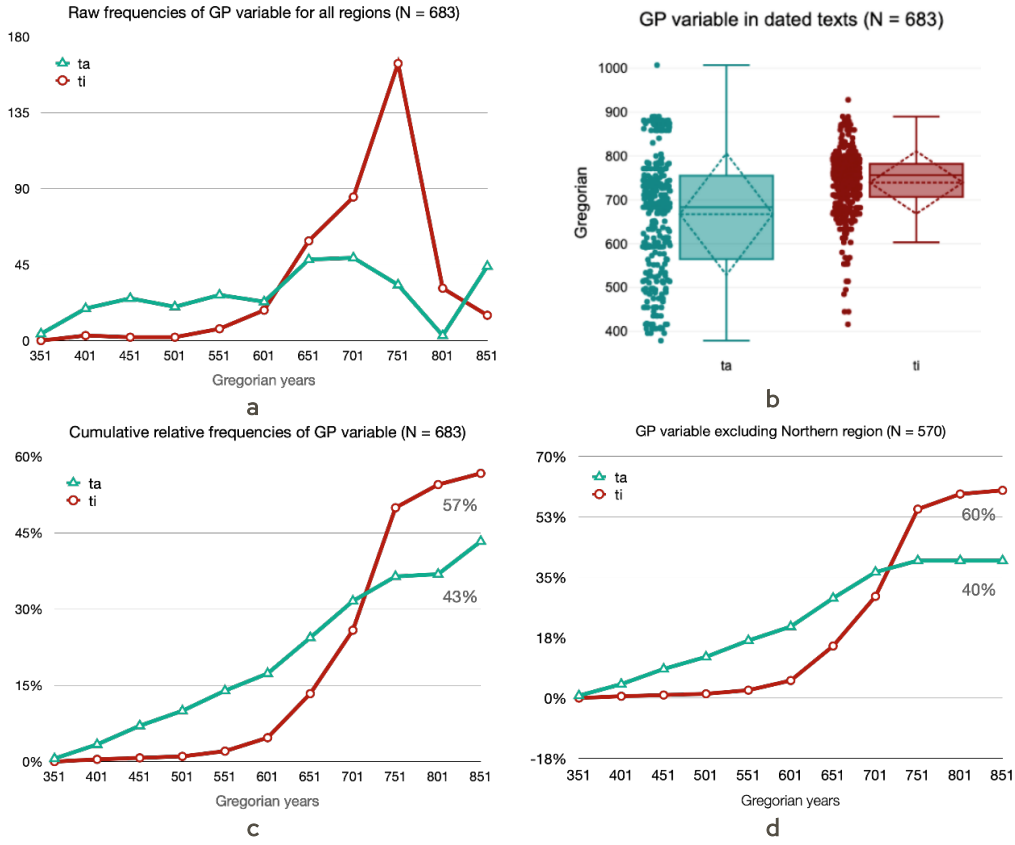

Figure 8. Frequencies of the two major GP variants in dated texts. a) Raw frequencies reflecting overall text production during the Classic period, including generalized decline and cessation after CE 751. b) Boxplot of GP variants with mean (dashed horizontal lines), median (solid horizontal lines), and standard deviation (dashed diamonds) values. Prepared with DATAtab (DATAtab Team, 2025). c) Combined relative cumulative frequencies for all regions. d) Combined relative cumulative frequencies of Preposition variable minus Northern texts. a) and c)–d) prepared with Apple Numbers.

As is to be expected, matters are more complex when the data are broken down by region, as in Figure 9. First, Figure 9a shows the overlapping distributions of innovative *ti* for all seven regions, which appears to show three major groupings. The first grouping (Figure 9b), with the highest proportions of innovative *ti* at or close to the Completed phase, appears to show a Southern > Pasion > Usumacinta gradation, though this is not the last word on the matter. Next is the second grouping (Figure 9c), consisting of the Eastern and Central regions, ending at the Nearing-Completion and Mid-range phases, respectively. And last (Figure 9d), with the lowest rates of innovative *ti*, are the Northern and West regions, concluding in the Mid-range and New and Vigorous phases, respectively.

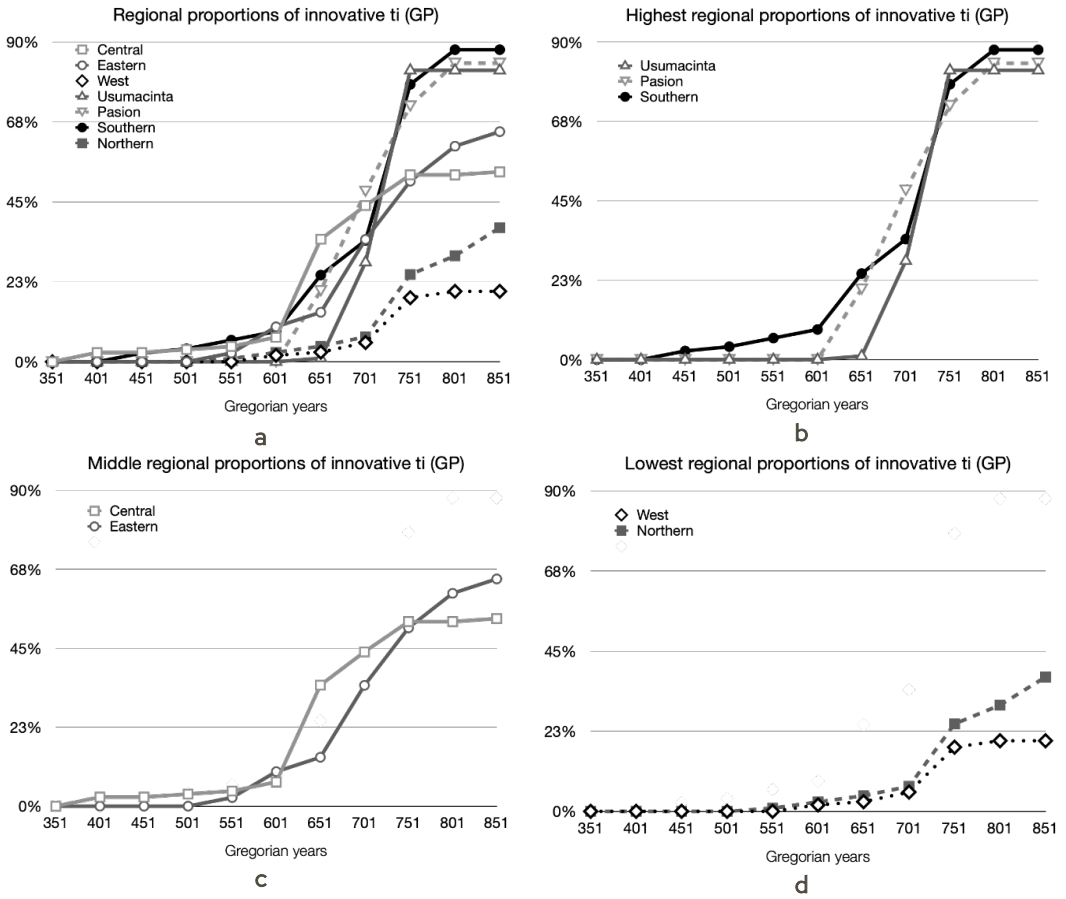

Figure 9. Combined relative cumulative frequencies for the innovative variant *ti*. a) Overlapping combined relative cumulative frequencies for all regions. b) Southern, Pasion, and Usumacinta regions. c) Eastern and Central regions. d) Northern and Western regions. All charts prepared with Apple Numbers.

In almost all major regions, conservative *tä* shows lower mean temporal values (Gregorian dates) than the innovative *ti* (Table 6). Even in the West region (Palenque, Tortuguero, Tonina), where a preference for *tä* has been noted (Macri, 1988, 1991, 2021; Carter, 2009, p. 6–8, 17–21; Kelly, 2022, p. 101–107), and supported here (80.5% overall), innovative *ti* was increasing in use relative to conservative *tä*, though at a much lower rate compared to the rest of the lowlands. The only exception is the Northern region (or at least parts thereof), as already noted by Macri (2021) and Kelly (2022). There the reverse is true: the mean values of conservative *tä* were higher than those of innovative *ti*. This matter is revisited in Section (4.4).

Region	Mean *tä*	Mean *ti*	N *tä*	% *tä*	N *ti*	% *ti*
West	688.87	754.77	91	**80.5%**	22	19.5%
Central	544.59	684.56	198	**65.1%**	106	34.9%
Northern	836.51	786.68	73	**59.8%**	49	40.2%
Pasion	580.6	742.57	14	37.8%	23	**62.2%**
Eastern	590.95	741.58	73	32.3%	153	**67.7%**
Usumacinta	631.5	752.6	19	18.4%	84	**81.6%**
Southern	529	732.42	13	15.3%	72	**84.7%**

Table 6. Descriptive statistics of GP variable for each region (mean Gregorian dates, frequency, proportions). The proportions of the GP variable are for all texts (N = 990), both dated (N = 683) and undated (N = 391). The mean dates are based on dated examples only. Unprovenienced texts (N = 84) were excluded.

Intraregional relative proportions of the GP variables (Table 6) are considered under the assumption that, with a change-in-progress, diffusion across space recapitulates diffusion within a social group, as stated earlier. It is also possible to consider the Mean Gregorian values of the variable for each region. Taking both kinds of information into account, a cluster hierarchical analysis was carried out, resulting in a Cluster Dendrogram (Figure 10) showing an increase in the proportion of innovative *ti* toward the left. This analysis differs somewhat from the charts obtained from the data from dated texts only (cf. Figure 9), placing not only the Usumacinta region above the Pasion in rates of innovative *ti*, but also the Eastern region. This is a matter that cannot be resolved at this time, so that a clear difference between these three regions is not proposed here. The dendrogram suggests a divide between two main clusters, the first grouping the Northern and West regions, and the second and more innovative including the Central, Pasion, Eastern, Usumacinta, and Southern regions, with the last one showing the highest proportion of innovative *ti* relative to conservative *tä*. I conclude that there was a general gradation from Southern → Pasion/Eastern/Usumacinta → Central in the *popular* spread of innovative *ti*, in spite of the fact that its earliest attestations come from the Central region site of Tikal.

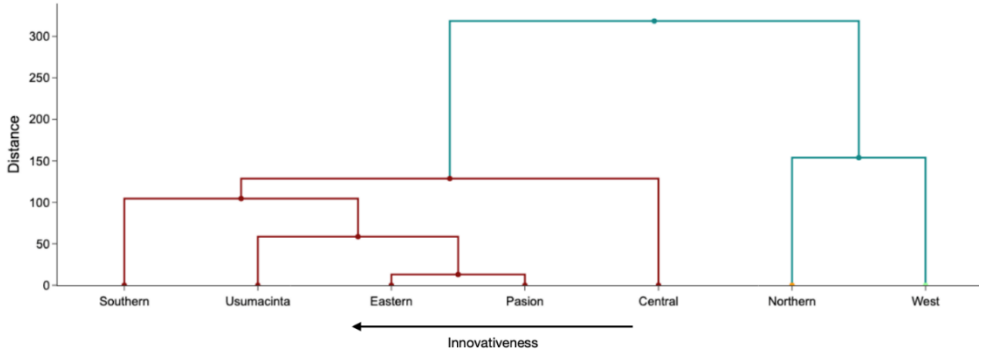

Figure 10. Cluster dendrogram of regional similarity based on intraregional proportions and Mean Gregorian values of the GP Variable. Prepared with DATAtab (DATAtab Team, 2025).

3.2. LOGISTIC REGRESSION ANALYSES

Now it is time to consider the results of the Logistic Regression analyses. Three sets of tests were carried out: 1) the first considers the entire time span with the more detailed Generic Preposition

Dataset; 2) the second considers the entire time span with the more restricted Accession Statements Dataset; and 3) the third breaks down the time span into three S-Curve phases (Incipient, New and Vigorous, Mid-range) with the Generic Preposition Dataset.

3.2.1. ENTIRE TIME SPAN: GP DATASET

Table 7 summarizes the Logistic Regression results for the GP variable based on the Generic Prepositions Dataset, with the goal of determining whether the following independent factors influenced its distribution: Text Type (media), Region, Time (Gregorian), and Linguistic Context. Since two regions show a high incidence of the conservative variant, and one of them even shows a late surge in it, the summary includes the significant positive correlations for both variants, not just the innovative one. Also, the Central and Eastern regions were used as reference categories for the Region independent variable, as they are the most frequently attested in the dataset.

Dependent Variant	Significant relationships	Coeff. B	Standard error	z	*p*-value	Odds Ratio	95% conf. interv.
tä	Region West (Central)	2.84	0.37	7.69	<.001	17.06	8.28 - 35.16
	Region Northern (Central)	3.02	0.4	7.52	<.001	20.52	9.33 - 45.11
	Region West (Eastern)	2.64	0.35	7.47	<.001	14.01	7 - 28.01
	Region Northern (Eastern)	2.82	0.37	7.63	<.001	16.85	8.16 - 34.79
	Gregorian	-0.01	0	10.02	<.001	0.99	0.98 - 0.99
	Context Broad Oblique Case (Calendar/Temporal)	2.18	0.69	3.15	.002	8.82	2.28 - 34.19
	Context Broad Locative (Complementizer)	0.92	0.37	2.47	.013	2.5	1.21 - 5.17
	Context Broad Calendar/Temporal (Complementizer)	0.64	0.27	2.35	.019	1.89	1.11 - 3.23
	Context Broad Oblique Case (Complementizer)	2.18	0.69	3.15	.002	8.82	2.28 - 34.19
ti	Region Southern (Central)	1.53	0.47	3.23	.001	4.61	1.83 - 11.63
	Region Southern (Eastern)	1.72	0.47	3.64	<.001	5.61	2.22 - 14.21
	Gregorian	0.01	0	10.02	<.001	1.01	1.01 - 1.02
	Context Broad Complementizer (Calendar/Temporal)	0.64	0.27	2.35	.019	1.89	1.11 - 3.23

Table 7. Summary of statistically significant results of LR analysis for GP linguistic variable (N = 143) in Generic Prepositions Dataset. Independent variables: Text Type (Portable vs. Monumental), Region, Time (Gregorian), Broad Linguistic Context. Reference categories, whenever there are more than two options, are shown between parentheses.

These results can be further summarized as in Table 8. Overall, innovative *ti* is associated with the Southern region, with the Complementizer linguistic function, and increases over time (Gregorian), while conservative *tä* is associated very strongly with the West and Northern regions, with all but the Complementizer linguistic function, and generall decreases over time. The spread of *ti* was a change-in-progress, essentially completed in some regions (Southern, Pasion, Eastern) long before the cessation of the inscriptional record. Text type according to media (portable vs. monumental) does not appear to exert any influence in this dataset.

Variant	Region	Time (Gregorian)	Text Type (Media)	Linguistic Factor
tä	West and Northern	Negative	–	Locative, Calendar/Temporal, Oblique Case
ti	Southern	Positive	–	Complementizer

Table 8. Independent factors favoring the distribution of each GP variants.

The question of linguistic contexts is worth discussing further. Macri (1991, 2021) has proposed that the syntactic function of the prepositional phrase headed by the GP variable, especially the complementizer cases, could have had a significant influence on the choice of variant, but did not test this idea quantitatively. She also suggested a phonological conditioning (Macri, 2021, p. 8–9), though the present dataset has not been coded to test this idea yet. This paper defined a Broad Linguistic Context independent variable consisting of five categories: Complementizer, Locative, Oblique Case, Calendar/Temporal, With/For. The results from the general Logistic Regression analysis support Macri's (1991, 2021) proposal that the Complementizer function was crucial to the spread of innovative *ti*. This issue is revisited below, in connection with the third Logistic Regression test.

3.2.2. ENTIRE TIME SPAN: ACCESSION STATEMENTS DATASET

This test uses the Accession Statements Dataset in order to focus on the following independent factors: Text Type (portable vs. monumental), Rank Size, and Interactions. This dataset is composed of records of accession to office, so that an association with social factors and their proxies might be more likely than in texts lacking such content. Table 17 summarizes the significant positive correlations for both GP variants when only Text Type, Rank Size, and Interactions are considered as independent factors.

Dependent Variant	Significant relationships	Coeff. B	Standard error	z	*p*-value	Odds Ratio	95% conf. interv.
tä	Rank Size Middle (Bottom)	2.08	0.92	2.25	.024	7.98	1.31 - 48.75
	Interactions Nondipl (None)	1.12	0.42	2.68	.007	3.05	1.35 - 6.92
	Interactions Nondipl (Dipl)	2.66	0.72	3,69	<.001	14.29	3.48 – 58.64
ti	Rank Size Bottom (Middle)	2.08	0.92	2.25	.024	7.98	1.31 - 48.75
	Rank Size Bottom (Top)	1.36	0.83	1.65	.099	3.9	0.77 - 19.66
	Interactions Dipl (None)	1.54	0.68	2.28	.023	4.68	1.24 - 17.66
	Interactions Dipl (Nondipl)	2.66	0.72	3.69	<.001	14.29	3.48 - 58.64

Table 9. Summary of statistically significant results of LR analysis for GP linguistic variable (N = 157) in Accession Dataset, excluding Ø variant (19 cases), to test for significance of Rank Size, Media, and Interactions. Independent variables: Rank Size, Media, Interactions. Reference categories, whenever there are more than two options, are shown between parentheses.

Table 10 provides a simplified summary. Overall, innovative *ti* is associated with Bottom site Rank Size ("villages"), and with Diplomatic interaction statements, while conservative *tä* is associated Middle site Rank Size ("smaller cities or large towns") and with Non-diplomatic interaction statements. Perhaps smaller sites ("villages") were more likely to resort to diplomatic strategies, presumably because of their more limited resources (smaller armies) than larger sites ("smaller cities or large towns"), which likely had more political strategies at their disposal, including larger armies

that allowed for more military options. As with the overall Generic Prepositions Dataset, text type according to media (portable vs. monumental) does not appear to exert any influence in this dataset. This could suggest that innovative *ti* became associated with less exclusionary or confrontational political discourse, perhaps a useful type of discourse during times (*e.g.* Late Classic) of seemingly increasing rates of warfare. These associations suggest that the GP variable, characterizable as a change-in-progress, was in fact a sociolinguistic marker.

Variants	Rank Size	Text Type (Media)	Interactions
tä	Middle	–	Non-diplomatic
ti	Bottom	–	Diplomatic

Table 10. Social factor proxies and GP variants in Accession Statements Dataset.

3.2.3. BREAKDOWN BY S-CURVE PHASES

This test consists of applying a Logistic Regression analysis to each phase of the S-Curve, using the Generic Preposition Dataset, generally following the example by Nevalainen and Raumolin-Brunberg (2003, p. 193–198), to investigate which factors were more important during each phase. Labov's classification of S-Curve phases is applied: Incipient (below 15%), New and Vigorous 15–35%), and Mid-range (36–65%). **Figure 11** shows the temporal distribution of innovative *ti* divided into the respective phases.

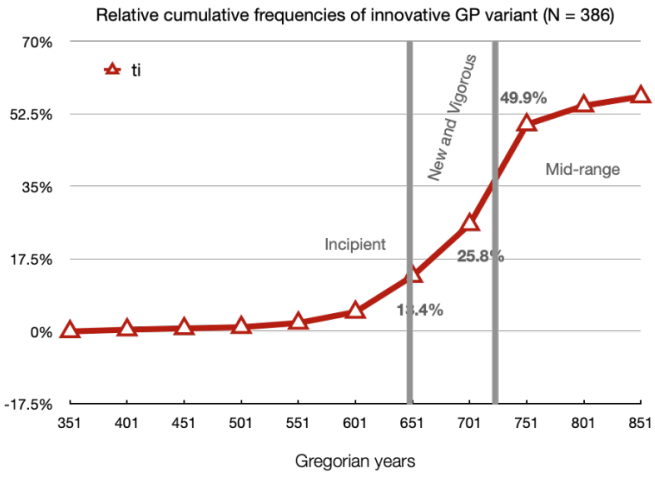

Figure 11. S-Curve phases for innovative GP variant *ti*. Prepared with Apple Numbers.

Table 11 presents the significant positive correlations for innovative *ti* for each phase derived from a Logistic Regression analysis, while Table 12 does the same for conservative *tä*. Starting with Table 11, it can be shown, first, that linguistic factors were the most important, and second, that innovative *ti* is 9.07 times more likely to occur in Locative and 4.23 times more in Complementizer functions than in Calendrical/Temporal functions. During the New and Vigorous phase, the only

independent variable that proved influential was Text Type, with innovative *ti* being 7.32 times more likely to occur in monumental texts than portable ones. And during the Mid-range phase, innovative *ti* became more likely to occur in the Eastern and Usumacinta regions (compared to Central), and it reversed its association with monumental texts, becoming 2.53 times more likely to occur in portable texts than monumental ones. Several points should be highlighted: linguistic factors were more important initially (Incipient phase), after which style/register (Text Type proxy) and region became more important; also, while in general the Southern region can be argued to be the focus of the spread of innovative *ti*, it is only strongly correlated with it during the Incipient phase.

Dependent Variant	Significant relationships	Coeff. B	Standard error	z	*p*-value	Odds Ratio	95% conf. interv.
Incipient	Region Central (West)	2.22	1.15	1.94	.052	9.25	0.98 - 87.39
	Region Southern (West)	2.51	1.14	2.2	.028	12.35	1.31 - 115.98
	Gregorian	0.02	0.01	3.35	.001	1.02	1.01 - 1.03
	Context Broad Locative (Calendar/Temporal)	2.2	0.92	2.4	.016	9.07	1.5 - 54.86
	Context Broad Complementizer (Calendar/Temporal)	1.44	0.62	2.34	.019	4.23	1.26 - 14.21
New and Vigorous	Type Monument	1.99	0.99	2.01	.044	7.32	1.05 - 50.89
Mid-range	Type Portable	0.93	0.44	2.1	.035	2.53	1.07 - 6.03
	Region Eastern (Central)	3.06	0.83	3.69	<.001	21.25	4.18 - 107.91
	Region Usumacinta (Central)	2.37	0.58	4.11	<.001	10.72	3.46 - 33.23

Table 11. Summary of statistically significant results of LR analysis for innovative *ti* GP variant in Generic Prepositions Dataset, to test for significance of Media, Region, Time (Gregorian), and Broad Linguistic Context. Reference categories, whenever there are more than two options, are shown between parentheses.

Table 12 presents the summary of the Logistic Regression analysis results for the conservative *tä* variant. During the Incipient phase, and only during this phase, linguistic factors are prominent: Calendar/Temporal and With/For functions were strongly correlated. It was also strongly associated with the West region (9.25 times more likely than the Central region), and negatively associated with the Southern region. During the New and Vigorous phase it was strongly associated with only one factor: portable texts. And during the Mid-range phase, it became associated with Monumental texts, with the Northern region (whether Central or Eastern region was used as the reference category), and with the West and Central regions (when the Eastern region was used as the reference category). The results are discussed further in Section (3.2.4).

Dependent Variant	Significant relationships	Coeff. B	Standard error	z	*p*-value	Odds Ratio	95% conf. interv.
Incipient	Region West (Central)	2.22	1.15	1.94	.052	9.25	0.98 - 87.39
	Region West (Southern)	2.51	1.14	2.2	.028	12.35	1.31 - 115.98
	Region Southern (Eastern)	-1.82	0.88	2.06	.039	0.16	0.03 - 0.91
	Region Southern (West)	-2.51	1.14	2.2	.028	0.08	0.01 - 0.76
	Gregorian	-0.02	0.01	3.03	.002	0.98	0.97 - 0.99
	Context Broad Calendar/Temporal (Complementizer)	1.66	0.6	2.76	.006	5.24	1.62 - 17.01
	Context Broad With (Complementizer)	2.48	1.23	2.03	.043	11.96	1.08 - 132.06
New and Vigorous	Type Portable	1.99	0.99	2.01	.044	7.32	1.05 - 50.89
Mid-range	Type Monument	0.89	0.44	2.01	.045	2.44	1.02 - 5.82
	Region West (Eastern)	3.74	0.81	4.6	<.001	42.29	8.58 - 208.5
	Region Central (Eastern)	3.04	0.83	3.65	<.001	20.89	4.09 - 106.62
	Region Northern (Central)	0.98	0.5	1.96	.05	2.67	1 - 7.14
	Region Northern (Eastern)	4.02	0.78	5.14	<.001	55.83	12.05 - 258.63

Table 12. Summary of statistically significant results of LR analysis for conservative *tä* GP variant in Generic Prepositions Dataset, to test for significance of Media, Region, Time (Gregorian), and Broad Linguistic Context. Reference categories, whenever there are more than two options, are shown between parentheses.

3.2.4. SYNTHESIS OF RESULTS

Table 13 presents a simplified summary of the results, although caution should be observed, since these results are a synthesis of different analyses using datasets of different sizes designed for different purposes. Note that the linguistic factors play a major role during the Incipient phase, and also that it is both the Complementizer and Locative functions that play an important role in the spread of innovative *ti* at this point, not just the Complementizer function proposed by Macri (1991, 2021). Also during the Incipient phase, conservative *tä* shows a clear association with the West region, while innovative *ti* shows a clear association with the Southern region, as it does also for the entire temporal range of interest (cf. Table 8). Note too that Text Type, used as a proxy for style/register (*i.e.* "formal/official" vs. "informal/unofficial"), plays an important role during the New and Vigorous phase, and also that during the Mid-range phase the associations with text or media types are reversed. Surprisingly, while the breakdown by phase reveals strong associations with Text Type, the results for the entire temporal range do not (cf. Table 8).

Variants	Incipient Phase	New and Vigorous Phase	Mid-range Phase
tä	West region (Central or Eastern or Southern as reference), Calendar/Temporal Contexts, With/For Contexts	Portable texts (15.6% of dated texts during this phase)	Monumental texts (82.1% of dated texts during this phase), Northern region (72.9% of its texts monumental during this phase), West region (74.4% of its texts monumental during this phase)
ti	Southern region (West as reference), Time (Gregorian), Locative Contexts, Complementizer Contexts	Monumental texts (84.4%)	Portable texts (17.9%), Eastern region (15.7% of its texts were portable during this phase)

Table 13. Independent factors influencing GP variants in Generic Prepositions Dataset according to S-Curve phases.

The associations with Text Type during the New and Vigorous and Mid-range phases can be accounted for as follows. During the New and Vigorous phase, the majority of dated texts were monumental (84.4% of the total for the phase) and innovative *ti* accounted for the majority of instances of the GP variable in dated texts (57.8%). It seems likely that scribes perceived *ti* as more formal and official sounding, and that this perception promoted its rapid increase in usage during this time; if so, the change-in-progress could be proposed to be a change from above.

During the Mid-range phase, two regions that were exhibiting the highest rates of conservative *tä*, the Northern region and the West region, were characterized primarily by monumental texts, 72.9% for the former, 74.4% for the latter. During this phase the primary development was the retention of *tä* in the West (Late Classic), where it clearly must have been regarded as more appropriate for the written register in general, and the takeover of *tä* in the Northern region (Late and Terminal Classic), possibly, as argued below, instigated by immigrants from the West region following the decline and collapse in their homelands. Also, at the same time that the Northern region —with its strong preference for monumental texts— was increasing its rate of use of conservative *tä*, the inscriptional records in the remaining regions (including the West region) were declining and coming to close, accounting for the overrepresentation of the combination of Northern region texts, monumental texts, and conservative *tä*.

Thus, it is the New and Vigorous phase, when all regions were actively engaged in the inscriptional record, that should be regarded as more representative, and it is during that phase when innovative *ti* surpasses conservative *tä*. Innovative *ti*'s association with monumental texts, presumably more "formal" and "official," at this time suggests that change in this sociolinguistic marker was a change from above —from above awareness, perhaps even prescribed by scribes, at least at some sites (where *ti* replaced *tä* completely). This implies that West region scribes *actively* resisted innovative *ti*, almost certainly an above-awareness process, perhaps even, from their perspective, a sociolinguistic stereotype.

3.2. NORTHERN REGION

As already hinted at, the most interesting results reported here may pertain to the Northern region. There, the GP variable shows the reverse pattern of the rest of the Maya lowlands, a crossover effect, as seen in **Figure 12**. The variant *ti*, an innovation from the perspective of the Ch'olan languages (Proto-Mayan **tya* > Proto-Ch'olan-Tzeltalan **ta* > Proto-Ch'olan **tä* ~ **ti*), is actually attested in the Northern region prior to the conservative variant *tä* (**Figure 12a**). This makes sense if, as is often assumed —and as there is substantial evidence to suppose— scribes in that region were primarily Yucatecan or Ch'olan/Yucatecan speakers, as Proto-Yucatecan can only be reconstructed with **ti?*.

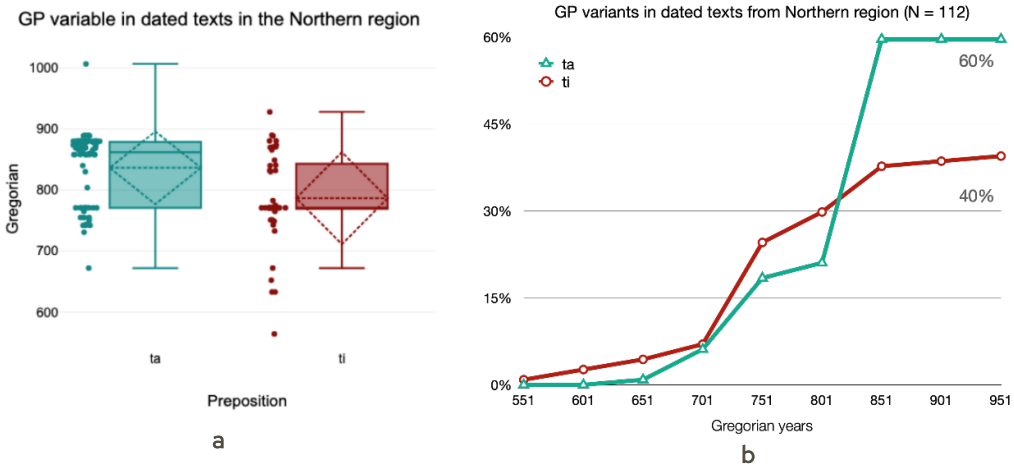

Figure 12. Distribution of GP variable in Northern region. a) Box Plot. Prepared with DATAtab (DATAtab Team, 2025). b) Combined relative cumulative frequencies. Prepared with Apple Numbers.

Despite the general supposition, which is well-founded, there is much evidence that the Northern region scribes, Ch'olan/Yucatecan bilinguals or Yucatecan monolinguals, generally wrote using the Ch'olan-based matrix with occasional embedding of exclusive Yucatecan traits (*e.g.* Justeson; Fox, 1989; Lacadena; Wichmann, 2000, 2002, 2005). In this region, during the eighth century, instances of conservative *tä* began to increase, even approximating the frequency of *ti*; then, during the second half of the ninth century, conservative *tä* overtook *ti*, very suddenly and definitively (Figure 12b). This is a reversal of what transpires in most of the Maya lowlands, where *ti* increases over time at the expense of *tä*. That said, in the West region, especially, and the Central region, to a much lesser extent, this process was slower than in other regions.

It is as though the Northern region scribes all of a sudden began to identify with Ch'olan speakers from the West or Central region (or both), and thus, this reversal could be understood as a Labov-hypercorrection (Chambers; Trudgill, 2004, p. 82).

But there is an alternative. The second half of the ninth century corresponds to the extremely rapid process of cessation of inscriptions (Ebert *et al.*, 2014), assumed to be indicative of the collapse of royal dynasties throughout the southern Maya lowlands, associated with depopulation of sites, and migration to other areas (Dahlin; Quizar; Dahlin, 1987; Martin; Grube, 2008, p. 227–229). Some parts of the Northern region were not affected by these events and processes, and in fact, it is in that region where a few sites, like Chichen Itza, gain more power and flourish. This is when the dramatic increase of conservative *tä* began in the Northern region (Figure 13), during the period of collapse of royal dynasties and scribal institutions from the southern Maya lowlands, including the regions with high incidence of use of *tä*, especially the West region, and to a lesser extent the Central region. Rather than an act of identity on the part of Yucatecan or Ch'olan/Yucatecan bilingual scribes in the Northern region, perhaps the unusual rise of *tä* reflects a different process: the immigration of scribes from the West or Central region especially, arriving at sites throughout the

Northern region, and either influencing the scribal practices of local scribes, or assuming control of such practices—and perhaps of the politics of the region in general. If so, the unusual reversal would not have been a case of Labov-hypercorrection, but an imposition by a foreign elite group, a linguistic superstratum.

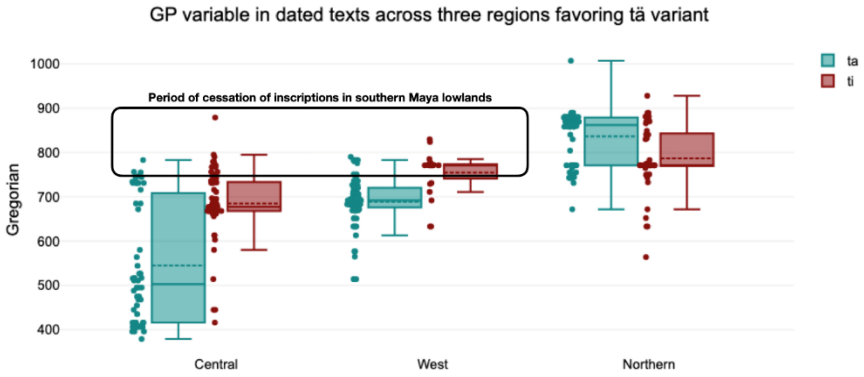

Figure 13. Box Plot indicating distribution of GP variable in Central, West, and Northern regions with reference to period of cessation of inscriptions across the southern Maya lowlands (including the Central and West regions). Prepared with DATAtab (DATAtab Team, 2025).

Before proceeding further, it is necessary to determine whether the West or Central region is a more likely source of influence on the Northern region. While the overall Classic distribution of the GP variable in the Central region was of 65.1% conservative *tä* to 34.9% innovative *ti* (cf. Table 6), during the Late and Terminal Classic periods, specifically, the tables had turned, reflecting 23% conservative *tä* to 77% innovative *ti*. In contrast, during the Late and Terminal Classic periods the West region maintained a higher proportion of conservative *tä*, 79%, to innovative *ti*, 21%. This suggests that the West region was the more likely source of influence on the Northern region during this time (60.3% *tä*, 39.7% *ti*).

There is additional linguistic evidence to support the idea of a strong influence from Ch'olan scribes: the unusual reversal of the GP variable was not an isolated phenomenon. During the second half of the ninth century, a significant increase in phonographic spellings of terms reflecting Ch'olan phonological innovations is observed in that region (Figure 14).

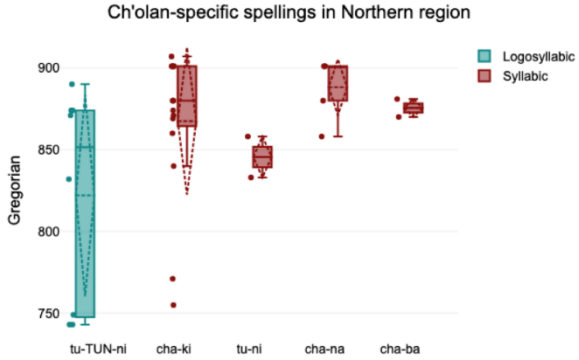

Figure 14. Box Plot of distribution of four terms spelled phonographically and exhibiting Ch'olan-Tzeltalan and Ch'olan phonological innovations: *tun* 'stone', *chahk* 'rain/thunder', *chan* 'sky', and *chab'* 'land'.

Following the differentiation of Ch'olan-Tzeltalan into distinct Ch'olan and Tzeltalan groups, Ch'olan and Yucatecan interacted especially closely (Lowland Mayan contact sphere), and Ch'olan acquired a number of morphological and syntactic traits through such contact (Hopkins, 1984, 1985; Kaufman; Norman, 1984; Kaufman, 2015), while Yucatecan borrowed a large number of ritually and politically charged etyma from Ch'olan (Justeson *et al.*, 1985). Some of these loans from Ch'olan into Yucatecan are reconstructible to Proto-Yucatecan, as with items #1 and #5 in Table 14, even when at least one Yucatecan variety preserved the native reflex in a specialized context (*e.g.* Colonial Yucatec <cauac> '19th day of ritual calendar' for #1) or with a shifted meaning (*e.g.* Contemporary Yucatec *tòon* 'testicles' for #2). These two etyma experienced a sound change that can be attributed to Ch'olan-Tzeltalan, the **k(') > ch(')* shift (cf. Kaufman; Norman, 1984; Law *et al.*, 2014; Mora-Marín, 2022d), and a sound change attributable specifically to Ch'olan, the **oo > *uu > u* shift (Kaufman; Norman, 1984; Mora-Marín; Frazier, 2021). Other similarly interesting etyma from Table 14 are also attested in Northern region texts, pointing to Ch'olan innovations, but not with the same frequency as #1 and #5.

Proto-Mayan	Ch'olan(-Tzeltalan) *k > ch (tʃ) shift	Ch'olan *oo > *uu > u	Ch'olan loan from Yucatecan	Native Yucatecan reflexes	Yucatecan loans from Ch'olan
1. *kahoq 'rain/thunder'	*chahuk > Proto-Ch'olan *chahuk ~ *chahk			Colonial Yucatec <cauac> '19th ritual day name (calendar)', thus Proto-Yucatecan *ka(h)wak	Proto-Yucatecan *chahk ~ *cháak
2. *kaab' ~ *kab' 'land'	Proto-Ch'olan *chab'		*kab'	Proto-Yucatecan *kab'	
3. *kaan 'snake'	*chaan > Proto-Ch'olan *chan			Proto-Yucatecan *kaan ~ *kàan	Colonial Yucatec <chicchan> 'boa constrictor'
4. *kaʔŋ 'sky'	*chaʔŋ > *chaan > Proto-Ch'olan *chan			Proto-Yucatecan *kaʔn ~ *káʔn	Colonial Yucatec <lahun chan> 'Ten Sky (deity name)'
5. *tooŋ 'stone'	*tooŋ > *toon	*tuun > Proto-Ch'olan *tun		Yucatec *tòon* 'testicles', thus Proto-Yucatecan *toon ~ *tòon	Proto-Yucatecan *tuun(ich) ~ *tùun(ich) 'stone'

Table 14. A few of the etyma that experienced sound changes in Ch'olan-Tzeltalan and Ch'olan but not Yucatecan.

These terms could be spelled logographically, logosyllabically, or syllabically. In general, the most frequent spelling type for each one was logographic, such as **CHAK** for *chahuk/chahk* and **TUN** for *tun*, or a logosyllabic spelling, such as **CHAK-ki** and **TUN-ni**, respectively, that did not disambiguate the linguistic origin of the etymon: **CHAK-ki** does not disambiguate the initial consonant, the one that experienced the Ch'olan-Tzeltalan **k > ch (tʃ)* shift, and **TUN-ni** does not disambiguate the vowel, the one that experienced the Ch'olan **oo > *uu > u* shift. In the Northern region, however, syllabic spellings such as **cha-ki** for *chahk* and **tu-ni** for *tun*, as well as logosyllabic spellings such as **tu-TUN-ni** became more common (cf. **Figure 14**): these spellings clearly

disambiguate the consonant of *chahk* and the vowel of *tun*, pointing to their Ch'olan pronunciations. The question is whether such phonographically explicit Ch'olan spellings were significantly overrepresented in the Northern region compared to the rest of the Maya lowlands. **Table 15** shows a comparison of phonographically explicit (logosyllabic, syllabic) and non-explicit (only logographic) spellings of these two terms, contrasting the Northern region and the rest of the Maya lowlands.

Ch'olan term	Northern	Rest	Total
Explicit *chahk*	17	46	63
Non-explicit *chahk*	13	361	374
Explicit *tuun*	14	30	44
Non-explicit *tuun*	179	971	1150
Total	223	1408	1631

Table 15. Totals per term collected by means of the MHD (Looper; Macri, 1991–2025).

Table 16 collapses the two etyma into two categories: explicit spellings versus non-explicit spellings, in the process setting up a contingency table for the application of a Chi-Square Test for Independence. This test examines whether the number of phonetically explicit spellings of these Ch'olan terms in the Northern region is comparable to (not statistically different from) the Rest of the Mayan region. The results reject the null hypothesis, showing a statistically significant difference, more specifically an *overrepresentation* of phonetically explicit spellings for *chahk* and *tuun* in the Northern region compared to the rest of the Maya region, by about twice the expected number.

Frequencies	Northern region	Rest	Row totals
Explicit Spellings	**31**	*76*	107
Expected	14.62968	92.37032	107.00000
Non-explicit Spellings	*192*	**1,332**	1,524
Expected	208.37032	1,315.62968	1,524.00000
Column totals	223	1,408	1,631
Expected	223.00000	1,408.00000	1,631.00000
Chi-square test (df=1)	Value	p-value	
Pearson Chi-Square	22.70912	0.00000188	
Yates continuity corrected Chi-square	21.34309	0.00000384	

Table 16. Chi-Square Test of Independence. Overrepresented values shown in italics, underrepresented values are underlined. Prepared using StatPlus for Mac.

The following interpretations are offered: 1) Northern region scribes, likely Ch'olan/Yucatecan bilinguals, were highlighting their knowledge of Ch'olan variants due to their prestige; 2) Northern region scribes, whether Ch'olan/Yucatecan bilinguals or Yucatecan monolinguals, were strongly influenced by West region scribes, perhaps after prolonged trade and political interactions over time; 3) West region scribes, escaping the political collapse of the southern lowlands during the Late and Terminal Classic periods, immigrated into the Northern region, and became strongly influential on the

local scribal traditions due to their prestige; or 4) West region lords, along with their scribes, escaped the political collapse of the southern lowlands and invaded the Northern region, taking over the political and scribal institutions at such sites. In any case, the Ch'olan scribes influencing the scribal record from the Northern region during this time were likely predecessors of Colonial Acalan and contemporary Yokot'an, who preserved a preference for *tä* over *ti* (Acalan) or *tä* exclusively (Yokot'an). These four scenarios may all be applicable, depending on the specific part of Northern region or the specific time. Whatever the specific scenario, it is also possible that the influential Ch'olan scribes could be identified with the ethnohistorical "Putun Maya" merchants and warriors who spread their influence during the Terminal Classic period (cf. Sharer, 1994, p. 348–349, 382–383).

One last observation and interpretation is offered here. As noted above, in the earliest Northern region examples of the GP variable, scribes were already using **ti**, likely to represent the Yucatecan cognate *ti?. Previously, in Section (3.2.3), it was suggested that it was unlikely that Ch'olan speakers would have heard a Yucatecan shape /ti?/ as [ti], since Ch'olan allows /CV?/ shapes, but that some sort of influence from Yucatecan was nonetheless likely implicated in the origin of the innovative variant *ti* reconstructible to Proto-Ch'olan *tä ~ *ti. Perhaps the source was not the spoken Yucatecan form *ti?, but its spelling, **ti**, in what would amount to an innovation arising through spelling pronunciation. This grammatical particle, then, could be added to the list of grammatical morphemes that entered Ch'olan from a likely Yucatecan substrate (Hopkins, 1984, 1985).

3.3. THE CENTRAL REGION

Carter (2009, p. 6–8, 17–21) supported prior work suggesting that *ta* (*i.e. tä*) was the earlier and more widespread variant, that it was retained or preferred at western sites like Palenque, Tortuguero, and Tonina well into the Late Classic period, but that *ti* generally took over throughout the lowlands, starting its rapid spread "during the eleventh k'atun" (after CE 633). Carter (2009, p. 20–21) concludes that

> Throughout most of the Lowlands, *ta* and *ti* coexisted in courtly and vernacular languages alike, although *ta* was evidently considered more correct, at least for use in elite inscriptions, during the Early Classic. The situation was reversed in the Late Classic, possibly due to the increased cultural prestige of a scribal school or tradition connected to the Kan dynasty based at Calakmul and, earlier, at Dzibanche (Martin, 2005). This tradition's influence may have derived from Calakmul's political and military successes, but it was not limited to Calakmul's subject polities or even to sites with which it maintained friendly relations. At most sites, *ti* became a marker of formal discourse, but not of Calakmul identity or affiliation. At Tonina, Palenque, and Tortuguero, by contrast, local vernacular languages may have militated against the adoption of *ti* as a preposition.

Carter's claim that "*ta* and *ti* coexisted in courtly and vernacular languages alike" merits revision: conservative *tä* preceded innovative *ti*, which appears initially at the beginning of the fifth century, a few centuries after the earliest instances of *tä*. His suggestion that "the increased cultural prestige" of the scribes from the Snake Kingdom (*e.g.* Dzibanche, Calakmul, La Corona) may have played a

role in the spread of *ti* is an interesting possibility worth testing. This suggestion is underlain by an assumption: that the Snake Kingdom's rivals, foremost among them Tikal, did not favor *ti*, a site located also within the Central region, just like the major Snake Kingdom sites. **Figure 15** provides charts of the distribution of the GP variable in dated texts from the Snake Kingdom (Calakmul, La Corona), the Southern region (Copan, Quirigua), and Tikal.

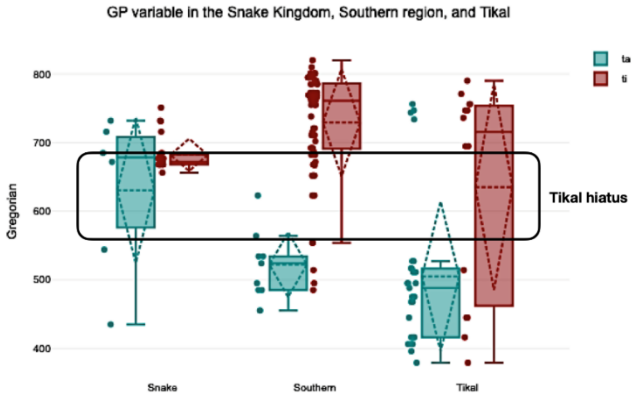

Figure 15. Box Plot comparing distribution of GP variable in the Snake Kingdom, Southern region, and Tikal. Black rectangle indicates the duration of the Tikal hiatus (ce 557–692).

Unfortunately, it is not a straightforward task to test the idea that the Snake Kingdom popularized the spread of innovative *ti* in the southern Maya lowlands. For one, there is the problem of the already mentioned Tikal hiatus (Moholy-Nagy, 2003, 2016), a period of CE 557–692, or 135 years, during which no dated inscriptions have been documented at the site of Tikal. This hiatus period happens to take up the majority of the range of 188 years worth of texts from the site of La Corona (ce 544–732). Not only that, but given that the range of dated texts with cases of the GP variable at Tikal is 411 years, and that following the hiatus most cases of the GP variable at Tikal, 69.23%, are cases of the innovative *ti*, as noted by Macri (2021, p. 4, 12), it would be expected that, had a significant number of texts from such survived at Tikal, many more instances of *ti* would be contained in them. Thus, the comparison, given the current circumstances, would not be fair.

Consequently, I have only included in my comparison texts from the post-hiatus Late Classic period for the respective sites and regions. Table 17 provides the basic statistics for the GP variable for Tikal, the Snake Kingdom, and the Southern region following Tikal's hiatus (*i.e.* after CE 692). Note that, following the hiatus, the proportion of innovative *ti* to conservative *tä* is very similar for Tikal and the Snake Kingdom: 53% to 47.1% in the former, 56% to 44% in the latter. Note too that neither the Snake Kingdom nor Tikal comes close to the Southern region's distribution, 96% for *ti* to 4% for *tä*. Note, finally, that the Snake Kingdom follows both Tikal and the Southern region in its earlier uses of innovative *ti* (cf. Figure 13). While it is undeniable the Snake Kingdom (Calakmul and La Corona) picked up innovative *ti* at the beginning of the seventh century at a very rapid rate, there

is no reason to suppose that it was its main popularizer, nor that it even showed a significantly greater use than its main rival, Tikal.

	tä		ti		Total
Location	n	% within Preposition	n	% within Preposition	n
Snake	4	44.44%	5	55.56%	9
Southern	2	3.7%	52	96.3%	54
Tikal	8	47.06%	9	52.94%	17
Total	14		66		80

Table 17. Post-hiatus distribution of GP variable for locations of interest.

It should be noted that Kelly (2022, p. 239–243), though she did not study the GP variable in this regard, found only limited evidence suggestive of a major role of the Snake Kingdom in the spread of other orthographic and linguistic innovations.

4. DISCUSSION

4.1. SPREAD AND HISTORICAL LINGUISTIC ASSESSMENT

It is important to emphasize that the earliest occurrences of innovative variants do not necessarily tell us something about their popular spread. The earliest examples of innovative *ti* can be traced to Tikal, in the Central region, and yet it is clear that it was the Southern region, and very likely the ancient city of Copan, specifically, that led this Classic-period change-in-progress. It is the large-scale distributional patterns, geographic and temporal, that must be assessed in order to establish where, when, how, and why an innovation spread in the past.

Figure 16 illustrates the proposed spatial transmissions of the variants studied in this paper. The first map (Figure 16a) shows the spread of *ti* from the Southern region westward (Pasion > Usumacinta) and northward (Eastern > Central). The Southern region's lead in the spread of innovative *ti* can be said to be consistent with the Eastern Ch'olan varieties, Ch'olti' and Ch'orti', attested from the area in and around Copan and Quirigua, though it should be noted that while Ch'olti' attests exclusively to *ti*, Ch'orti' attests to both *ta* and *ti*. The Northern region already had *ti* (perhaps *ti?*), exclusively at an early point (cf. Figure 12), likely reflecting the local Yucatecan varieties. As several authors have suspected, it is likely that Ch'olan *ti* was the result of influence from Yucatecan **ti?*, whether through direct borrowing, but omitting the /ʔ/ to avoid homophony with Ch'olan **ti?* 'mouth; speech', as suggested here, or through spelling pronunciation given the Yucatecan scribes' use of **ti** to spell Yucatecan **ti?*, as also suggested here.

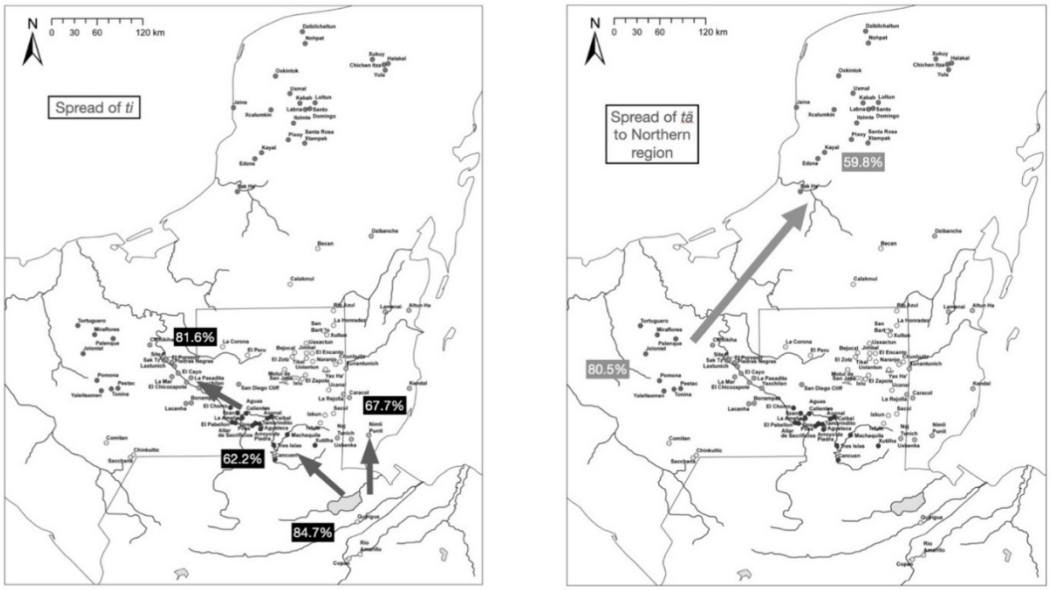

Figure 16. Spatial spread of GP variable. a) Innovative GP variant *ti*. b) Conservative GP variant *tä* spreading to the Northern region either from the West region (more likely) or the Central region (less likely) or both during period of collapse of southern lowland dynastic polities.

Now, as far as the conservative *tä* variant is concerned (Figure 16b), the West region preserved it, avoiding adopting and popularizing the innovative *ti* for most of the Classic period. This is consistent with the GP variable as attested in Acalán, with <ta> in high frequency and <ti> in very low frequency. Contemporary Yokot'an varieties only seem to attest to *tä*. The evidence from Ch'ol, where *ti* is dominant and *tyä* (*i.e.* /tä/) is preserved only in a few idiomatic expressions, suggests that perhaps Western Ch'olan (Ch'ol, Acalan, Yokot'an) was differentiated somewhat during the time when innovative *ti* was spreading across the lowlands. But overall, it seems entirely plausible that the real-time spread of innovative *ti* could correspond to its spread within an undifferentiated or weakly differentiated Ch'olan speech community.

What is most interesting about this variant, is its sudden increase in the Northern region in the Late and Terminal Classic periods, coinciding with the timing of termination of dated inscriptions in the southern lowlands (cf. Ebert *et al.*, 2014). It was proposed, in Section (4.4), that perhaps scribes from the West region migrated to the Northern region, accounting for the sudden rise and unexpected dominance of *tä* in that region, as well as the unusually high frequency of explicit spellings exhibiting exclusive Ch'olan phonological changes. Whether this was a political takeover of Northern polities by elites from the West, with the imposition of a Ch'olan linguistic superstratum, or the result of prestigious Ch'olan scribes having a great deal of influence on the scribal practices of the Northern region, is unclear. Thus, rather than a reversal, or a case of Yucatecan having had *tä* early on, like Ch'olan, and then changing to *ti*, as suggested by Kelly (2022, p. 101–107), what is

documented in the Northern region is more likely external linguistic influence. The Ch'olan speakers responsible may correspond to the influential "Putun Maya" from the ethnohistorical and archaeological literature.

4.2. LINGUISTIC FACTORS

The GP variable was tested for evidence of linguistic embedding and the results suggested, following an original suggestion by Macri (1991, 2021), that the Complementizer and Locative functions were the most influential in the spread of innovative *ti*, with the former showing an earlier temporal association than the latter. The phase-by-phase Logistic Regression test suggested that linguistic context was a determining factor during the Incipient phase, after which non-linguistic factors appear to be more important.

4.3. SOCIAL FACTORS

As far as the investigation of proxies for social factors (style/register, population density, political strategies) on the spread of the GP variable, Table 18, a combination of Tables 8 and 10 above, offers some ideas. Generally speaking, innovative *ti* likely spread in part due to its association with one or more social factors reflected via the proxies employed here.

Variant	Rank Size	Media	Interpolity Interactions
tä	Middle	Portable texts (New and Vigorous), Monumental texts (Mid-range)	Non-diplomatic
ti	Bottom	Monumental texts (New and Vigorous), Portable texts (Mid-range)	Diplomatic

Table 18. Possible social factors influencing the distribution of the morphological variables.

Conservative *tä* shows a correlation with both Middle sites (*e.g.* Cancuen, Kabah, Pomona Tabasco, Tortuguero, Uxul) and Non-diplomatic political strategies, while *ti* shows a correlation with both Bottom sites (*e.g.* Bejucal, Comalcalco, Chinikiha, Dzibanche, La Mar, La Sufricaya) and Diplomatic strategies. Perhaps the larger and more densely populated the site, the more likely (or capable) it is to engage in warfare; and perhaps the smaller and less densely populated the site, the more likely (or motivated) it is to engage in diplomatic strategies.

Text Type did not prove to be a strong factor during the Incipient phase. During the New and Vigorous phase, innovative *ti* was 7.32 times more likely to appear in monumental texts than portable texts, and thus may have been perceived as more formal/official. This association flipped, though to a lesser degree (2.44 Odds Ratio), in the Mid-range phase, probably due to a variety of factors that remain to be carefully elucidated.

5. CONCLUSIONS AND FUTURE RESEARCH

This paper set out to test a quantitative approach to the study of the historical sociolinguistics of Epigraphic Mayan texts during the Classic period, employing datasets prepared by means of the Maya Hieroglyphic Database (MHD) by Looper and Macri (1991–2025). It focused on the distribution, temporal and geographic, of a morphological variable that has received some attention in epigraphic literature: the Generic Preposition (GP) (*tä* ~ *ti*). It also studied the linguistic and social factors that influenced the distribution of these variables, as well as to what extent the epigraphic and historical linguistic lines of evidence can be reconciled.

More generally, the paper shows that it is not only possible to apply a historical sociolinguistic framework to EMY texts, but also, that the EMY corpus, with its precise chronological information, can be used to explore real-time process of change, and also, that while the social profiles of ancient Mayan scribes remain largely obscure, it is possible to utilize proxies to approximate the social factors influencing the distribution of linguistic variables. Some of the choices for analyzing temporal and regional patterns made here will need to be improved in future iterations. Also, having explored broad, regional patterns in this paper, a future paper should focus on the trajectories that characterized individual sites, and even zoom in further, paying attention to specific political events that occurred at key moments during the evolution of the GP variable.

In the end, the paper utilized a dual approach: it employed categories derived from the variationist study of (primarily) contemporary Western societies, the basis for the application of the Uniformitarian Principle, and categories derived from patterns in the data (cf. Lauersdorf, 2018). The results suggest that ancient Mayan scribes, at the very least, were no less sensitive to sociocultural forces in their speech and writing than their ancient Roman counterparts (Joseph and Wallace, 2011) or contemporary New Yorkers (Labov 1966) for that matter.

There is a long list of tasks awaiting scholars interested in the historical sociolinguistics of ancient Mayan society and their contemporary descendants: 1) the need for a systematic consideration of graphic, graphemic, orthographic, and linguistic variables; 2) a comprehensive characterization of linguistic variation at various scales (*i.e.* phonological vs. morphological vs. syntactic vs. discourse-pragmatic); 3) further problematization of the notion of "conservative" vs. "standard" vs. "prestige" written languages as applied to Epigraphic Mayan; 4) the utilization of comprehensive databases such as the Maya Hieroglyphic Database; 5) further problematization of the creation and utilization of analytical proxies for social factors; 6) a more detailed reexamination of explicit data on the social and linguistic identities of scribes, along the lines of Montgomery (1995) and Houston (2016); 7) a representative and tagged linguistic corpus of Epigraphic Mayan, designed with individual texts as the basic organizational unit; and 8) the development of a systematic variationist sociolinguistics of contemporary Mayan communities.

Finally, Epigraphic Mayan is undergoing a revitalization in the hands of indigenous Mayans from different linguistic backgrounds (e.g. Matsumoto 2015; Paz Joj 2021); a sociolinguistic study of this process could very well offer insights into the past.

ACKNOWLEDGMENTS

I am very grateful to the participants at several of the Meetings and Incubators organized by the North American Research Network in Historical Sociolinguistics (NARNiHS), which I have participated in, and off, since 2019, and especially Mark Lauersdorf, Joe Salmons, and Israel Sanz-Sánchez for their encouragement and feedback. Thanks are also due to the folks who attended Quarantinis and Glyphs #56 on March 15th, 2024, for their feedback on an early version of this research, including Victoria Bricker, Nicholas Hopkins, John Justeson, Barbara MacLeod, and Martha Macri for their detailed feedback. This research would not have been possible without the generosity of Martha Macri and Matthew Looper, who have shared different versions of the Maya Hieroglyphic Database with me over the years, prior to its publication online. Finally, I am indebted to the reviewers and the journal editors, especially Joe Salmons and Mary Kate Kelly, whose feedback gave me the chance to drastically improve the final product.

ADDITIONAL INFORMATION

CONFLICT OF INTEREST

There are no competing or conflicting interests to report.

STATEMENT OF DATA AVAILABILITY

All three datasets (Appendices 1–3) analyzed in this paper are publicly available at the following URLs as CSV files:
https://davidmm.web.unc.edu/2025/08/07/appendices-for-historical-sociolinguistics-paper-on-the-generic-preposition-of-mayan-writing/
https://cdr.lib.unc.edu/concern/data_sets/sq87c915d
Appendix 1. Accession Statements Dataset.
Appendix 2. Generic Preposition Dataset.
Appendix 3. Northern Region Ch'olan Terms Dataset.

AI USAGE STATEMENT

There was no use of Artificial Intelligence in this work.

REVIEW AND AUTHORS' REPLY

Review: https://doi.org/10.25189/2675-4916.2025.V6.N1.ID794.R

Author's Reply: https://doi.org/10.25189/2675-4916.2025.V6.N1.ID794.A

REFERENCES

AISSEN, Judith; ENGLAND, Nora C.; ZAVALA MALDONADO, Roberto (eds.). *The Mayan Languages*. London and New York: Routledge/Taylor and Francis Group, 2017.

BABOSHKIN, Maxim. Gramática descriptiva de la lengua de las inscripciones jeroglíficas mayas. PhD Thesis, Universidad Nacional Autónoma de México, 2022.

BAILEY, Guy; WIKLE, Tom; TILLERY, Jan; SAND, Lori. Some patterns of linguistic diffusion. *Language Variation and Change*, v. 3, p. 359–390, 1993.

BLANTON, Richard E.; FEINMAN, Gary M.; KOWALEWSKI, Stephen A.; PEREGRINE, Peter N. A dual-processual theory for the evolution of Mesoamerican civilization. *Current Anthropology*, v. 37, p. 1–14, 1996.

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

BRICKER, Victoria R. The Last Gasp of Maya Hieroglyphic Writing in the Books of Chilam Balam of Chumayel and Chan Kan. *In*: HANKS, William F.; RICE, Don S. (eds.). *Word and Image in Maya Culture. Explorations in Language, Writing, and Representation*. Salt Lake City: University of Utah Press, 1989. p. 39-50.

BRICKER, Victoria R. Mayan. *In*: WOODARD, Roger D. (ed.). *The Cambridge Encyclopedia of the World's Ancient Languages*. Cambridge University Press, 2004. p. 1041–1070.

BRICKER, Victoria R.; ORIE, Olanike O. Schwa in the Modern Yucatecan Languages and Orthographic Evidence of Its Presence in Colonial Yucatecan Maya, Colonial Chontal, and Precolumbian Maya Hieroglyphic Texts. *International Journal of American Linguistics*, v. 80, p. 175–207, 2014.

BRITAIN, David. Innovation Diffusion in Sociohistorical Linguistics. *In*: HERNÁNDEZ-CAMPOY, Juan Manuel; CONDE-SILVESTRE, Juan Camilo (eds.). The Handbook of Historical Sociolinguistics. Blackwell Publishing Ltd., 2012. p. 451–464.

BROWN, Clifford T.; WITSCHEY, Walter R. T. The Geographic Analysis of Maya Settlement and Polity. Proceedings of the 2001 Pacific Neighborhood Consortium and Electronic Cultural Atlas Initiative Conference. Taipei: Academica Sinica, 2001.

BROWN, Clifford T.; WITSCHEY, Walter R. T. The Electronic Atlas of Ancient Maya Sites. Paper presented at the Symposium on Current Applications of Remote Sensing and GIS in North America and Mesoamerican Archaeology, 67th Annual Meeting of the Society for American Archaeology, Denver, CO, Friday, March 22, 2002.

CAMPBELL, Lyle. On Sound Change and Challenges to Regularity. *In*: DURIE, Mark; ROSS, Malcolm (eds.). *The Comparative Method Reviewed: regularity and irregularity in language change*. Oxford: Oxford University Press, 1996. p. 72–89.

CARTER, Nicholas P. Paleographic trends and linguistic processes in Classic Ch'olti'an: a spatiotemporal distributional analysis. Master's thesis, Department of Anthropology, Brown University, Providence, 2009.

CHAMBERS, J. K. Patterns of Variation including Change. *In*: CHAMBERS, J. K.; SCHILLING, Natalie (eds.). *The Handbook of Language Variation and Change*. Second Edition. Chichester, West Sussex: Wiley-Blackwell, 2013. p. 297–324.

CHAMBERS, J. K.; TRUDGILL, Peter. *Dialectology*. Second Edition. Cambridge University Press, 2004.

CHASE, Diane Z.; CHASE, Alren F. Archaeological Perspectives on Classic Maya Social Organization from Caracol, Belize. *Ancient Mesoamerica*, v. 15, p. 139–147, 2004.

COE, Michael D. *Old gods and young heroes: The Pearlman Collection of Maya ceramics. The Israel Museum*. The Maremont Pavilion of Ethnic Arts, Jerusalem, Spring 1982, 1982.

DAHLIN, Bruce H.; QUIZAR, Robin; DAHLIN, Andrea. Linguistic Divergence and the Collapse of Preclassic Civilization in Southern Mesoamerica. *American Antiquity*, v. 52, p. 367–382, 1987.

DATAtab Team. *DATAtab: Online Statistics Calculator*. DATAtab e.U. Graz, Austria, 2025. https://datatab.net.

EBERT, Claire E; PRUFER M., Keith M.; MACRI, Martha J.; WINTERHALDER, Bruce; KENNETT, Douglas. Terminal Long Count Dates and the Disintegration of Classic Period Maya Polities. *Ancient Mesoamerica*, v. 25, p. 337–356, 2014.

FABRICIUS, Anne H. Ongoing change in modern RP: Evidence for the disappearing stigma of t-glottaling. *English World-Wide*, v. 23, p. 115–136, 2002.

FOX, James; JUSTESON, John. Hieroglyphic Evidence for the Languages of the Classic Maya. Unpublished manuscript used with permission of author, 1982.

GRUBE, Nikolai. The Orthographic Distinction between Velar and Glottal Spirants in Maya Hieroglyphic Writing. *In*: WICHMANN, Søren (ed.). *The Linguistics of the Maya Script*. Salt Lake City: University of Utah Press, 2004. p. 61–82.

GRUBE, Nikolai; MARTIN, Simon. Política clásica maya dentro de una tradición mesoamericana: Un modelo epigráfico de organización política 'hegemonica'. *In*: TREJO, Silvia (ed.). *Modelos de entidades políticas mayas. Primer Seminario de Mesas Redondas de Palenque*. Instituto Nacional de Antropología e Historia, Mexico City, 1998. p. 131–146.

HAUSMAN, Heidi. Mayan Sites/Ruinas Mayas, 2013. https://adaptwest.databasin.org/datasets/3aa6b24c882144d6a4197bd277ae753d/.

HOFLING, Charles Andrew. The Morphosyntactic Basis of Discourse Structure in Glyphic Text in the Dresden Codex. *In*: HANKS, William F.; RICE, Don S. (eds.). *Word and Image in Maya Culture. Explorations in Language, Writing, and Representation*. Salt Lake City: University of Utah Press, 1989. p. 51–71.

HOFLING, Charles Andrew. Comparative Maya (Yucatec, Lacandon, Itzaj, and Mopan Maya). *In*: AISSEN, Judith; ENGLAND, Nora C.; ZAVALA MALDONADO, Roberto (eds.). *The Mayan Languages*. London and New York: Routledge/Taylor and Francis Group, 2017. p. 685-759.

HOPKINS, Nicholas A. La influencia del yucatecano sobre el cholano y su contexto histórico. *In*: *Investigaciones recientes en el área maya; XVII Mesa Redonda; 21-27 junio 1981*. San Cristóbal de Las Casas, Chiapas: Sociedad Mexicana de Antropología, 1984. p. 191–207.

HOPKINS, Nicholas A. On the History of the Chol Language. In BENSON, Elizabeth P. (ed.). *Fourth Palenque Round Table, 1980, Vol. VI*. San Francisco: Center for Pre-Columbian Art Research, 1985. p. 1–5.

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

HOPKINS, Nicholas A. Decipherment and the Relation between Mayan Languages and Maya Writing. *In*: MACRI, Martha J.; FORD, Anabel (eds.). *The Language of Maya Hieroglyphs*. San Francisco: Pre-Columbian Art Research Institute, 1997. p. 77–88.

HOPKINS, Nicholas A.; JOSSERAND, J. Kathryn. Issues of Glyphic Decipherment. Paper presented at the 17th Annual University Museum Maya Weekend, Maya Epigraphy– Progress and Prospects, University of Pennsylvania, Philadelphia, 1999.

HOUSTON, Stephen D. *Hieroglyphs and History at Dos Pilas: Dynastic Politics of the Classic Maya*. Austin: University of Texas Press, 1993.

HOUSTON, Stephen D. Crafting credit: authorship among Classic Maya painters and sculptors. *In*: COSTON, Cathy Lynne (ed.). *Making Value, Making Meaning: Techné in the Pre-Columbian World*. Washington, D.C.: Dumbarton Oaks, 2016. p. 391–431.

HOUSTON, Stephen D.; ROBERTSON, John; STUART, David S. The Language of Classic Maya Inscriptions. *Current Anthropology*, v. 41, p. 321–356, 2000.

HOUSTON, Stephen D. Quality and Quantity in Glyphic Nouns and Adjectives. Research Reports on Ancient Maya Writing 47. Washington D.C.: Center for Maya Research, 2001.

HOUSTON, Stephen D.; STUART, David S. On Maya Hieroglyphic Literacy. *Current Anthropology* 33:589–593, 1992.

HOUSTON, Stephen D.; STUART, David S.; ROBERTSON, John. Disharmony in Maya Hieroglyphic Writing: Linguistic Change and Continuity in Classic Society. *In*: CIUDAD RUIZ, Andrés; FERNÁNDEZ MARQUÍNEZ, María Yolanda; GARCÍA CAMPILLO, José Miguel; IGLESIAS PONCE DE LEÓN, María Josefa; LACADENA GARCÍA-GALLO, Alfonso; SANZ CASTRO, Luis Tomás (eds.). *Anatomía de una civilización: Aproximaciones interdisciplinarias a la cultura maya*. Madrid: Sociedad Española de Estudios Mayas, 1998. p. 275–296.

HOUSTON, Stephen D.; STUART, David S.; ROBERTSON, John. Disharmony in Maya Hieroglyphic Writing: Linguistic Change and Continuity in Classic Society. *In*: WICHMANN, Søren (ed.). *The Linguistics of the Maya Script*. Salt Lake City: University of Utah Press, 2004. p. 83–99.

HRUBY, Zachary X. Evidence for linguistic conservatism in the hieroglyphic script of the Central Petén. *Mayab*, v. 15, p. 49–59, 2002.

JOSEPH, Brian D.; WALLACE, Rex. Socially Determined Variation in Ancient Rome. *Language Variation and Change*, v. 4, p. 105–119, 1992.

JOSSERAND, J. Kathryn. Languages of the Preclassic period along the Pacific Coastal Plains of Southeastern Mesoamerica. *In*: LOVE, Michael; KAPLAN, Jonathan (eds.). *The Southern Maya in the Late Preclassic: The Rise and Fall of an Early Mesoamerican Civilization*. University Press of Colorado, Boulder, 2011. p. 141–174.

JOSSERAND, Kathryn; HOPKINS, Nicholas A. Classic Maya Social Interaction and Linguistic Practice: Evidence from Hieroglyphic Inscriptions and Mayan Languages. *In*: BLOS, Vera Tiesler; COBOS, Rafael; ROBERTSON, Merle Greene (eds.). *La organización social entre los mayas; memoria de la Tercera Mesa Redonda de Palenque, Vol. II*. México, D.F.: Instituto Nacional de Antropología e Historia, and Mérida: Universidad Autónoma de Yucatán, 2002. p. 355–72.

JUSTESON, John S. Hieroglyphic Evidence for Lowland Mayan Linguistic History. *International Journal of American Linguistics*, v. 51, p. 469–471, 1985.

JUSTESON, John S. The Representational Conventions of Mayan Hieroglyphic Writing. *In*: HANKS, William F.; RICE, Don S. (eds.). *Word and Image in Maya Culture. Explorations in Language, Writing, and Representation*. Salt Lake City: University of Utah Press, 1989. p. 25–38.

JUSTESON, John S.; FOX, James A. Hieroglyphic evidence for the languages of the Lowland Maya. Unpublished manuscript in possession of author, 1989.

JUSTESON, John S.; NORMAN, William M.; CAMPBELL, Lyle; KAUFMAN, Terrence. *The Foreign Impact on Lowland Mayan Language and Script*. Middle American Research Institute, Publication 53. New Orleans: Tulane University, 1985.

KAUFMAN, Terrence S. *El proto-tzeltal-tzotzil: fonología comparada y diccionario reconstruido*. Centro de Estudios Mayas, Cuaderno 5. Mexico: UNAM, 1972.

KAUFMAN, Terrence S. Archaeological and linguistic correlations in Mayaland and associated areas of Mesoamerica. *World Archaeology*, v. 8, p. 101–118, 1976.

KAUFMAN, Terrence S. Mayan Comparative Studies. Unpublished manuscript, 2015. https://www.researchgate.net/publication/340925779_Mayan_Comparative_Studies_Kaufman

KAUFMAN, Terrence S. Aspects of the Lexicon of Proto-Mayan and Its Earliest Descendants. *In*: AISSEN, Judith; ENGLAND, Nora C.; ZAVALA MALDONADO, Roberto (eds.). *The Mayan Languages*. London and New York: Routledge/Taylor and Francis Group, 2017. p. 62-111.

KAUFMAN, Terrence; JUSTESON, John. The History of the Word for Cacao in Ancient Mesoamerica. *Ancient Mesoamerica*, v. 18, p. 193–237, 2007.

KAUFMAN, Terrence; JUSTESON, John. The Epi-Olmec Language and its Neighbors. *In*: ARNOLD, Philip J., III; POOL, Christopher A. (eds.). *Classic Period Cultural Currents in Southern and Central Veracruz*. Washington, D.C.: Dumbarton Oaks Research Library and Collection, 2008. p. 55–83.

KAUFMAN, Terrence; NORMAN, William. An outline of Proto-Cholan phonology, morphology, and vocabulary. *In*: JUSTESON, John S.; CAMPBELL, Lyle (eds.). Phoneticism in Maya Hieroglyphic Writing. Institute for Mesoamerican Studies Publication No. 9. Albany: State University of New York, 1984. p. 77–166.

KELLY, Mary Kate. Speech Carved in Stone: Language Variation Among the Ancient Lowland Maya. PhD Dissertation, Department of Anthropology, Tulane University, 2022.

KERSWILL, Paul. Social Dialectology. *In*: AMMON, Ulrich (ed.). *Sociolinguistics: an international handbook of the science of language and society, Volume 1*. Berlin: Walter de Gruyter, 2004. p. 22–33.

KNOWLES, Susan M. A Descriptive Grammar of Chontal Maya (San Carlos Dialect). Ph.D. Thesis, Department of Anthropology, Tulane University, 1984.

KNUDSEN, Andrew Tate. The Role of Glyph Blocks in Punctuating Mayan Texts. MA thesis, Linguistics Department, University of North Carolina, 2023.

LABOV, William. Some principles of linguistic methodology. *Language in Society*, v. 1, p. 97–120, 1972.

LABOV, William. *The Social Stratification of English in New York City*. Washington, D.C.: Center for Applied Linguistics. Cambridge: Cambridge University Press, 1966.

LABOV, William. Building on Empirical Foundations. In Perspectives on Historical Linguistics: Papers from a conference held at the meeting of the Language Theory Division, Modern Language Association, San Francisco, 27–30 December 1979, edited by Winfred P. Lehmann, and Yakov Malkiel, p. 17–92. John Benjamins Publishing Company, 1982.

LABOV, William. Principles of Linguistic Change. Volume 1: Internal Factors. Oxford, UK & Cambridge, USA: Blackwell, 1994.

LABOV, William. Principles of Linguistic Change. Volume 2: Social Factors. Oxford, UK & Cambridge, USA: Blackwell, 2001.

LACADENA GARCÍA-GALLO, Alfonso; WICHMANN, Søren. The dynamics of language in the Western Lowland Maya region. Paper presented at the 2000 Chacmool Conference, Calgary, November 9-11, 2000.

LACADENA GARCÍA-GALLO, Alfonso; WICHMANN, Søren. The distribution of Lowland Maya languages in the Classic period. *In*: BLOS, Vera Tiesler; COBOS, Rafael; ROBERTSON, Merle Greene (eds.). *La organización social entre los mayas; memoria de la Tercera Mesa Redonda de Palenque, Vol. II*. México, D.F.: Instituto Nacional de Antropología e Historia, and Mérida: Universidad Autónoma de Yucatán, 2002. p. 275–319.

LACADENA GARCÍA-GALLO, Alfonso; WICHMANN, Søren. On the representation of the glottal stop in Maya writing. *In*: WICHMANN, Søren (ed.). *The Linguistics of the Maya Script*. Salt Lake City: University of Utah Press, 2004. p. 100–164.

LACADENA GARCÍA-GALLO, Alfonso; WICHMANN, Søren. The dynamics of language in the Western Lowland Maya region. *In*: WATERS-RIST, Andrea; CLUNEY, Christine; MCNAMEE, Calla; STEINBRENNER, Larry (eds.). *Art for Archaeology's Sake:*

Material Culture and Style Across the Disciplines, Proceedings of the 33rd Annual Chacmool Conference (2000). Calgary: The University of Calgary, 2005. p. 32–48.

LAUERSDORF, Mark Richard. Historical (Standard) Language Development and the Writing of Historical Identities: A Plaidoyer for a Data-Driven Approach to the Investigation of the Sociolinguistic History of (Not Only) Slovak. *In*: DICKEY, Stephen M.; LAUERSDORF, Mark Richard (eds.). *V zeleni drželi zeleni breg: Studies in Honor of Marc L. Greenberg*. Bloomington: Slavica Publishers, 2018. p. 199–218.

LAW, Danny. Pronominal borrowing among the Maya. *Diachronica*, v. 26, n. 2, p. 214–252, 2009.

LAW, Danny; ROBERTSON, John; HOUSTON, Stephen; ZENDER, Marc; STUART, David. Areal Shifts in Classic Mayan Phonology. *Ancient Mesoamerica*, v. 25, p. 357–366, 2014.

LAW, Danny; STUART, David. Classic Mayan: An overview of language in ancient hieroglyphic script. *In*: AISSEN, Judith; ENGLAND, Nora C.; ZAVALA MALDONADO, Roberto (eds.). *The Mayan Languages*. London and New York: Routledge/Taylor and Francis Group, 2017. p. 128–174.

LOOPER, Matthew G.; MACRI, Martha J. 1991–2025. *Maya Hieroglyphic Database*. Department of Art and Art History, California State University, Chico, 2017. http://mayadatabase.org.

LOOPER, Matthew G.; MACRI, Martha J. MHD Reference Materials 5: The Historical Development of the Maya Script: Preliminary Results. *Glyph Dwellers Report* 75, 2022.

LOOPER, Matthew G.; MACRI, Martha J.; POLYUKHOVYCH, Yuriy; VAIL, Gabrielle. MHD Reference Materials 1: Preliminary Revised Glyph Catalog. *Glyph Dwellers Report* 71, 2022.

LOOPER, Matthew; SCHOLNICK, Jonathan B.; MUNSON, Jessica L.; POLYUKHOVYCH, Yuriy; MACRI, Martha J. Measuring grapheme innovation in Classic Maya Writing. Poster presented at the Society for American Archaeology Meetings, San Francisco, CA. April 17, 2015.

MACRI, Martha J. A Descriptive Grammar of Palenque Mayan. PhD dissertation, UC-Berkely, Linguistics, 1988.

MACRI, Martha J. Prepositions and complementizers in the Classic Period inscriptions. *In*: ROBERTSON, Merle Greene; FIELDS, Virginia M. (eds.). *Sixth Palenque Round Table, 1986*. Norman, Oklahoma: University of Oklahoma Press, 1991. p. 266–272.

MACRI, Martha J. Prepositions ti and ta in Classic Maya Monument Texts. *Glyph Dwellers Report* 69, 2021.

MACRI, Martha J.; LOOPER, Matthew. *The New Catalog of Maya Hieroglyphs. Volume One: The Classic Period Inscriptions*. Norman: University of Oklahoma Press, 2003.

MACRI, Martha J.; VAIL, Gabrielle. *The New Catalog of Maya Hieroglyphs. Volume Two: The Codical Texts*. Norman: University of Oklahoma Press, 2009.

MARCUS, Joyce. Territorial Organization of the Lowland Classic Maya. *Science*, v. 180, p. 911–916, 1973.

MARCUS, Joyce. *Emblem and State in the Classic Maya Lowlands: An Epigraphic Approach to Territorial Organization*. Dumbarton Oaks, Washington, D.C, 1976.

MARTIN, Simon. Of snakes and bats: shifting identities at Calakmul. *The PARI Journal*, v. 6, n. 2, p. 5–15, 2005.

MARTIN, Simon. *Ancient Maya Politics: A Political Anthropology of the Classic Period 150–900 CE*. Cambridge University Press, 2020.

MARTIN, Simon; GRUBE, Nikolai. Evidence for Macro-Political Organization among Classic Maya Lowland States, 1994a. http://www.mesoweb.com/articles/Martin/Macro-Politics.pdf.

MARTIN, Simon; GRUBE, Nikolai. Classic Maya Politics within a Mesoamerican Tradition: An Epigraphic Model of 'Hegemonic' Political Organization. Paper presented at the Primer Seminario de las Mesas Redondas de Palenque, Palenque, México, 29 September–1 October, 1994b.

MARTIN, Simon; GRUBE, Nikolai. Maya Superstates. *Archaeology*, v. 48, n. 6, p. 41–46, 1995.

MARTIN, Simon; GRUBE, Nikolai. *Chronicle of the Maya Kings and Queens: Deciphering the Dynasties of the Ancient Maya*. Thames and Hudson, London, 2000.

MARTIN, Simon; GRUBE, Nikolai. *Chronicle of the Maya Kings and Queens: Deciphering the Dynasties of the Ancient Maya*. Second edition. Thames and Hudson, London, 2008.

MARTIN, Simon; HOUSTON, Stephen D.; ZENDER, Marc. Sculptors and Subjects: Notes on the Incised Text of Calakmul Stela 51. *Maya Decipherment: Ideas on Ancient Maya Writing and Iconography*. Posted January 7, 2015. https://decipherment.wordpress.com/2015/01/07/sculptorsandsubjectsnotesontheincisedtextofcalakmulstela51/.

MATHEWS, Peter; JUSTESON, John S. Patterns of sign substitution in Mayan hieroglyphic writing: the "Affix Culster." *In*: JUSTESON, John S.; CAMPBELL, Lyle (eds.). Phoneticism in Maya Hieroglyphic Writing. Institute for Mesoamerican Studies Publication No. 9. Albany: State University of New York, 1984. p. 185–232.

MATSUMOTO, Mallory E. La estela de Iximche' en el contexto de la revitalización lingüística y la recuperación jeroglífica en las comunidades mayas de Guatemala. *Estudios de Cultura Maya*, v. XLV, p. 225–258, 2015.

MATSUMOTO, Mallory E. Sharing Script: Development and Transmission of Hieroglyphic Practice among Classic Maya Scribes. PhD dissertation, Brown University, 2021.

MOHOLY-NAGY, Hattula. The Hiatus at Tikal, Guatemala. *Ancient Mesoamerica*, v. 14, p. 77–83, 2003.

MOHOLY-NAGY, Hattula. Set in Stone: Hiatuses and Dynastic Politics at Tikal, Guatemala. *Ancient Mesoamerica*, v. 27, p. 255–266, 2016.

MONTGOMERY, John. Sculptors of the Realm: Classic Maya Artist's Signatures and Sculptural Style during the Reign of Piedras Negras Ruler 7. MA thesis, Department of Art and Art History, University of New Mexico, Albuquerque, 1995.

MORA-MARÍN, David F. The Grammar, Orthography, and Social Context of Late Preclassic Mayan Texts. Unpublished Ph.D. Thesis. University at Albany, Albany, New York, 2001.

MORA-MARÍN, David F. Pre-Ch'olan as the Standard Language of Classic Lowland Mayan Texts. Paper presented at the 2002-2003 SSILA Meeting, Atlanta, Georgia, January, 2003.

MORA-MARÍN, David F. The Preferred Argument Structure of Classic Lowland Mayan Texts. *In*: WICHMANN, Søren (ed.). *The Linguistics of the Maya Script*. Salt Lake City: University of Utah Press, 2004. p. 339–364.

MORA-MARÍN, David F. Affixation Conventionalization Hypothesis: Explanation of Mayan Spelling Practices. Paper presented at the 43th Conference on American Indian Languages in Oakland, 2005.

MORA-MARÍN, David F. Full Phonetic Complementation, Semantic Classifiers, and Semantic Determinatives in Ancient Mayan Hieroglyphic Writing. *Ancient Mesoamerica*, v. 19, p. 195–213, 2008.

MORA-MARÍN, David F. A Test and Falsification of the "Classic Ch'olti'an" Hypothesis: A Study of Three Proto-Ch'olan Markers. *International Journal of American Linguistics*, v. 75, n. 2, p. 115–157, 2009a.

MORA-MARÍN, David F. Reconstruction of Proto-Ch'olan Independent Pronouns: Grammaticalization and Evidence for Sociolinguistic Variation. *Transactions of the Philological Society*, v. 107, n. 1, p. 98–129, 2009b.

MORA-MARÍN, David F. Consonant Deletion, Obligatory Synharmony, Typical Suffixing: An Explanation of Spelling Practices in Mayan Writing. *Written Language and Literacy*, v. 13, p. 118–179, 2010.

MORA-MARÍN, David F. A Historical Sociolinguistic Study of Two Morphological Markers in Classic Lowland Mayan Inscriptions. Paper presented at the Society for American Archaeology Conference Session on Research Utilizing the Maya Hieroglyphic Data Base, Friday, April 1st, 2011, San Jose, California, 2011.

MORA-MARÍN, David F. An Historical Sociolinguistic Approach to Classic Mayan Writing: A study of Two Morphological Innovations, *-(a)wan* 'intransitivizer of positionals' and *-(V)lel* 'abstractivizer of nouns'. Paper presented at the American Anthropological Association Session on Historical Sociolinguistics of the Maya Lowlands, organized by Marc Zender, Saturday, December 2nd, 2017.

MORA-MARÍN, David F. Framing the Historical Sociolinguistics of the Maya Lowlands (Southeastern Mexico, Guatemala, Belize, Honduras) during the Classic Period (ca. 200-900 CE). Presented at the The North American Research Network in Historical Sociolinguistics 2010 Research Incubator at KFLC: The Languages, Literatures, and Cultures Conference, April 11–13, 2019, University of Kentucky, Lexington, 2019.

MORA-MARÍN, David F. Prerequisites for A Historical Sociolinguistics of Ancient Mayan Hieroglyphic Texts. Presented at the The North American Research Network in Historical Sociolinguistics 2020 Research Incubator at KFLC: The Languages, Literatures, and Cultures Conference, April 16–18, 2020, University of Kentucky, Lexington, 2020.

MORA-MARÍN, David F. Historical Sociolinguistics of Mayan Writing: Graphic Designs of Syllabogram *ʔu* and Parameters of Variation. Presented at The North American Research Network in Historical Sociolinguistics 2021 Research Incubator at KFLC: The Languages, Literatures, and Cultures Conference, April 22–24, 2021, 2021a.

MORA-MARÍN, David F. Historical Sociolinguistics of Mayan Writing: Graphic and Graphemic Variation and Change. Presented at the Third Annual Meeting North American Research Network in Historical Sociolinguistics A Sister Society of the Linguistic Society of America (LSA) 8–11 January 2021, 2021b.

MORA-MARÍN, David F. A Study of T670 and Two Lexical Determinatives in Mayan Writing. *Glyph Dwellers Report*, v. 76, p. 1–23, 2022a.

MORA-MARÍN, David F. 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, Jueves 28 y viernes 29 de julio, 2022, Universidad del Valle y Museo Popol Vuh, 2022b.

MORA-MARÍN, David F. Graphic, Graphemic, Orthographic, and Linguistic Variables: A Close look at Accession Statements in Classic Mayan Texts. Paper presented at The North American Research Network in Historical Sociolinguistics 2022 Research Incubator at KFLC: The Languages, Literatures, and Cultures Conference, April 21, 2022, 2022c.

MORA-MARÍN, David F. Evidence, New and Old, Against the Late *k(') > ch(')* "Areal Shift" Hypothesis. *In*: SILVA, Wilson; LEE, Nala; CHACON, Thiago. *Festschrift for Lyle Campbell*. Edinburgh University Press, 2022d p. 130-163.

MORA-MARÍN, David F. Putting It All Together: Sociopolitical Interactions and the Distribution of Linguistic Variables in Mayan Writing. Paper presented at the Fifth Annual Meeting of The North American Research Network in Historical Sociolinguistics, A Sister Society of the Linguistic Society of America (LSA), January 7, 2023, 2023a.

MORA-MARÍN, David F. Evidence for lexical and phonetic determinatives in Mayan writing: The case of T713. Ancient Mesoamerica, First View. Published online 27 February 2023, 2023b.

MORA-MARÍN, David F. Evolution of Mayan Spelling Practices from the Late Preclassic Through the Late Classic Periods. Presented at the Coffee, Clever T-Shirts, and Papers in Honor of John S. Justeson Symposium at the Society for American Archaeology, Friday, March 31st, 2023, 2023c.

MORA-MARÍN, David F. A Historical Sociolinguistic Study of Four Linguistic Variables in Epigraphic Mayan Texts. Paper presented at the Society for the Study of Indigenous Languages of the Americas, Sunday, January 26th, 2025, 2025a.

MORA-MARÍN, David F. Zooming In, Zooming Out: The Spread of Linguistic Variables in Epigraphic Mayan Seen from Multiple Perspectives. Paper presented at the The North American Research Network in Historical Sociolinguistics (NARNiHS) Research Incubator, Friday, April 25, 2025, 2025b.

MORA-MARÍN, David F. A Historical Sociolinguistic Study of Two Linguistic Variables in Epigraphic Mayan During the Classic Period. Accepted for publication in *Journal of Historical Sociolinguistics*, n.d.

MORA-MARÍN, David; FRAZIER, Melissa. The Historical Reconstruction of Greater Tzeltalan (Mayan) Vowel Assimilation and Vowel Raising Patterns. *Transactions of the Philological Society*, v. 119, p. 182–240, 2021.

MORA-MARÍN, David; HOPKINS, Nicholas; JOSSERAND, Kathryn. The Linguistic Affiliation of Classic Lowland Mayan Writing and the Historical Sociolinguistic Geography of the Mayan Lowlands. Paper presented in the American Anthropological Association session on "The Ch'orti' Area: Past and Present on the Southeastern Maya Periphery," organized by Brent E. Metz and Cameron L. McNeil, 2005.

MORA-MARÍN, David; HOPKINS, Nicholas; JOSSERAND, Kathryn. The Linguistic Affiliation of Classic Lowland Mayan Writing and the Historical Sociolinguistic Geography of the Mayan Lowlands. *In*: METZ, Brent E.; MCNEIL, Cameron L.; HULL, Kerry (eds.). *The Ch'orti' Area: Past and Present on the Southeastern Maya Periphery*. University Press of Florida, 2009. p. 15–28.

MUNSON, Jessica; AMATI, Viviana; COLLARD, Mark; MACRI, Martha J. Classic Maya Bloodletting and the Cultural Evolution of Religious Rituals: Quantifying Patterns of Variation in Hieroglyphic Texts. *PLoS ONE*, v. 9, n. 9, p. e107982, 2014. doi:10.1371/journal.pone.0107982.

MUNSON, Jessica; LOOPER, Matthew; SCHOLNICK, Jonathan. Ritual networks and the structure of moral communities in Classic Maya society. *Journal of Anthropological Archaeology*, v. 74, 2024. https://doi.org/10.1016/J.JAA.2024.101584.

MUNSON, Jessica; MACRI, Martha. Sociopolitical network interactions: A case study of the Classic Maya. *Journal of Anthropological Archaeology*, v. 28, p. 424–438, 2009.

MUNSON, Jessica; SCHOLNICK, Jonathan; LOOPER, Matthew; POLYUKHOVYCH, Yuriy; MACRI, Martha. Ritual diversity and divergence of Classic Maya dynastic traditions: A lexical perspective on within-group cultural variation. *Latin American Antiquity*, v. 27, n. 1, p. 74–95, 2016.

NEVALAINEN, T.; RAUMOLIN-BRUNBERG, Helena. *Historical Sociolinguistics: Language Change in Tudor and Stuart England*. London: Longman Pearson Education, 2003.

NEVALAINEN, T.; RAUMOLIN-BRUNBERG, Helena. *In*: HERNÁNDEZ-CAMPOY, Juan Manuel; CONDE-SILVESTRE, Juan Camilo (eds.). The Handbook of Historical Sociolinguistics. Blackwell Publishing Ltd., 2012. p. 22–40.

PAZ JOJ, Walter Amilcar. Los mayas de hoy: reavivando el sistema de escritura antigua. *The Mayanist*, v. 2, n. 2, p. 41–62, 2021.

RANKIN, Robert. The Comparative Method. *In*: JOSEPH, Brian D.; JANDA, Richard D. (eds.). *The Handbook of Historical Linguistics*. Blackwell Publishing, 2005. p. 183–212.

ROBERGE, Paul R. Language History and Historical Sociolinguistics. *In*: AMMON, Ulrich (ed.). *Sociolinguistics: an international handbook of the science of language and society, Volume 3*. Berlin: Walter de Gruyter, 2006. p. 2307–2311.

ROBERTSON, John S. A Ch'olti'an Explanation for Ch'orti'an Grammar: A Postlude to the Language of the Classic Maya. Mayab 11:5–11, 1998.

ROMAINE, Suzanne. Socio-historical Linguistics. Its Status and Methodology. Cambridge University Press, 1982.

ROMAINE, Suzanne. Historical sociolinguistics: Problems and methodology. *In*: AMMON, Ulrich; DITTMAR, Norbert; MATTHEIER, Klaus J. (eds.). *Sociolinguistics: An International Handbook of the Science of Language and Society*. Berlin: Walter de Gruyter, 1988. p. 1452–69.

ROMAINE, Suzanne. Historical Sociolinguistics. *In*: AMMON, Ulrich; DITTMAR, Norbert; MATTHEIER, Klaus J.; TRUDGILL, Peter (eds.). *Sociolinguistics: An International Handbook of the Science of Language and Society, Volume 2*. Berlin: Walter de Gruyter, 2005. p. 1696–1703.

SCHNEIDER, Edgar W. Investigating Variation and Change in Written Documents. *In*: CHAMBERS, J. K.; TRUDGILL, Peter; SCHILLING-ESTES, Natalie (eds.). *Handbook of Language Variation and Change*. Blackwell Publishing, 2004. p. 67–96.

SHARER, Robert J. *The Ancient Maya*. Fifth Edition. Stanford University Press, 1995.

SMAILUS, Ortwin. *El Maya-Chontal de Acalán: Análisis lingüiístico de un documento de los años 1610–12*. Centro de Estudios Mayas, Cuadernos 9. Mexico: UNAM, 1975.

STUART, David. Ten Phonetic Syllables. *Research Reports on Ancient Maya Writing* 14. Washington, D.C.: Center for Maya Research, 1987.

STUART, David. Hieroglyphs on Maya Vessels. *In*: KERR, Justin (ed.). *The Maya Vase Book, A Corpus of Rollout Photographs of Maya Vases, vol. 1*. Kerr Associates, New York, 1989. p. 149–160.

TATE, Carolyn. *Ah Ts'ib*: Scribal Hands and Sculpture Workshops at Yaxchilán. *In*: FIELDS, Virginia; ROBERTSON, Merle Greene (eds.). *Seventh Palenque Round Table, 1989*. San Francisco, The Pre-Columbian Art Research Institute, 1994. p. 95–104.

THOMPSON, J. Eric S. *A Catalogue of Maya Hieroglyphics*. Norman: University of Oklahoma Press, 1962.

TRUDGILL, Peter. Linguistic change and diffusion: description and explanation in sociolinguistic dialect geography. *Language in Society* v. 3, p. 215–246, 1974.

VAN STONE, Mark. Identifying Individual Hands in the Monuments of *K'inich Ahkal Mo' Naab* of Palenque. Report submitted to FAMSI, 2000. http://www.famsi.org/reports/99027/99027VanStone01.pdf

VAN STONE, Mark. Aj-Ts'ib, Aj-Uxul, Itz'aat, and Aj-K'uhu'n: Classic Maya Schools of Carvers and Calligraphers in Palenque Ater the Reign of Kan-Bahlam. PhD dissertation, Department of Art and Art History, University of Texas, Austin, 2005.

WALD, Robert F. The Languages of the Dresden Codex: Legacy of the Classic Maya. *In*: WICHMANN, Søren (ed.). *The Linguistics of the Maya Script*. Salt Lake City: University of Utah Press, 2004. p. 27–58.

WALKDEN, George. The many faces of uniformitarianism in linguistics. *Glossa: a journal of general linguistics*, v. 4, n. 1, 52, p. 1–17, 2019.

WEINREICH, Uriel; LABOV, William; HERZOG, Marvin. Empirical foundations for a theory of language change. *In*: LEHMANN, Winifred P.; MALKIEL, Yakov (eds.). *Directions for Historical Linguistics*. Austin: University of Texas Press, 1968, p. 95–189.

WICHMANN, Søren. Mayan Historical Linguistics and Epigraphy: A New Synthesis. *Annual Review of Anthropology* 35:279–294, 2006.

WINTER, Wegner. Sociolinguistics and dead languages. In: JAHR, Ernst Hakon (ed.). *Language Change: Advances in Historical Sociolinguistics*. Berlin, New York: De Gruyter Mouton, 1999. p. 67–84.

WITSCHEY, Walter R. T.; BROWN, Clifford T. Electronic Atlas of Ancient Maya Sites, 2010. http://MayaGIS.smv.org.

WITSCHEY, Walter R. T.; BROWN, Clifford T. *Atlas of the Ancient Maya World*. Leiden, Boston: Brill, 2025.

WOLFRAM, Walt; SCHILLING-ESTES, Natalie. Dialectology and Linguistic Diffusion. *In*: JOSEPH, Brian D.; JANDA, Richard D. (eds.). *The Handbook of Historical Linguistics*. Blackwell Publishing, 2005. p. 713–135.