1 Introduction

This special issue of Language and Linguistic Compass is dedicated to a three-part survey of the grammatical structure of Mayan languages. The articles in this collection focus on core, well-studied aspects of Mayan linguistics: phonology (Bennett), syntax (Coon), and semantics (Henderson). The perspective taken is broadly typological, but also informed by theoretical issues in formal linguistics. Our goal is to share the richness of Mayan languages with the broader linguistic community, in the hopes of encouraging future research on these fascinating languages.

This introduction includes an overview of the Mayan language family and its internal structure (§2), as well as a discussion of the kinds of publications that we’ve drawn on in compiling the target articles (§3). We close with an outline of the glossing and transcription conventions that we’ve adopted for this special issue (§4).

2 Genetic classification and language names

2.1 The region and speakers

Mayan languages are spoken by over six million people across Mexico, Guatemala, Belize, and Honduras, as well as in various diaspora communities in the United States and Canada. The Maya area is divided into two geographical regions, the Highland and Lowland areas, with the boundary falling roughly along the chain of mountains that runs through central Guatemala (Figure 1; see Law 2014 for discussion).

Guatemala has the highest density of Mayan speakers of any country today. Official census estimates suggest that about 40% of Guatemalans speak a Mayan language natively; this almost certainly...
underestimates the actual percentage of native Mayan speakers in the country, which is likely 50-60% or higher (Fischer and Brown 1996a; England 2003). In some areas the percentage of native Mayan speakers is closer to 90-100%, as in parts of the Q’eqchi’-speaking region in the department of Alta Verapaz (Richards

Many (perhaps most) speakers of Mayan languages in Guatemala are bilingual in Spanish, but monolingualism in Mayan languages is widespread, especially in rural areas. Multilingualism in two or more Mayan languages is also attested, particularly in municipalities which lie on a border between distinct linguistic areas (e.g. both K’iche’ and Uspantek are spoken as home languages in the municipality of Uspantán; K’iche’ serves as a *lingua franca* in this area for speakers of Uspantek, Ixil and Q’eqchi’).

Outside of Guatemala, Mexico has the largest population of Mayan speakers in Mesoamerica. Based on census results reported in Law (2013) we can estimate that as of 2010 there were at least 2.5 million speakers of Mayan languages in Mexico, or just over 2% of the total population (the number of Mayan speakers in Belize is comparatively quite small, probably under 30,000 in total, and in Honduras there are only a handful of speakers of Ch’orti’; Lewis et al. 2015). These speakers are primarily concentrated in Southern Mexico and the Yucatán Peninsula, though Huastec is spoken further north in the San Luis Potosí and Veracruz regions of Central Mexico. As in Guatemala, bilingualism with Spanish is common, especially among younger generations, though many monolingual speakers exist as well.

### 2.2 Genetic classification

The Mayan family consists of about thirty languages, usually grouped into five or six major sub-groups. From top to bottom in the figure in (1), these branches are Yucatecan, Huastecan, Ch’olan–Tseltalan (a.k.a. Greater Tseltalan), Q’anjob’alan, Mamean, and K’ichean (Campbell and Kaufman 1985; England and Zavala 2013). The Ch’olan–Tseltalan and Q’anjob’alan branches together make up the Western Branch, while the Mamean and K’ichean branches make up the Eastern Branch. Most of these languages (about 21) are spoken in Guatemala. Some controversy over sub-family divisions exists, as discussed in Campbell and Kaufman 1985; Law 2014; Campbell to appear. In (1) we provide a reasonably well-accepted set of genetic subgroupings within the family. Genetic classification is particularly difficult in the case of Tojolab’al, which has been substantially shaped by areal contact between Mayan languages belonging to different sub-branches of the family (e.g. Law 2014).1

1It bears mentioning that even firmly-classified Mayan languages may show extensive traces of areal influence from more distantly-related Mayan languages. Uspantek and Sipakapense, for instance, are both K’ichean-branch languages that have been profoundly affected by contact with nearby Mamean languages (Campbell 1977, Barrett 1999, Law 2013, 2014; Campbell to appear).
It should be noted that three languages included in (1)—Achi, Akatek, and Chalchitek—are sometimes characterized as dialects of K’iche’, Q’anjob’al, and Awakek respectively, rather than as independent languages. See Kaufman (1976); England (2003); Richards (2003); Law (2013); Campbell (to appear) for additional details.
2.3 Language health

The health of Mayan languages varies quite widely. There are currently at least 7 Mayan languages with over 400,000 speakers each (K’iche’, Q’eqchi’, Kaqchikel, Mam, Tseltal, Tsotsil, and Yucatec), but also at least 6 languages with fewer than 5000 speakers each (Sakapultek, Uspantek, Tektitek, Mocho’, Itzaj, and

![Distribution of Mayan Languages](image)

**Figure 2:** Distribution of Mayan languages. Based on Map 4 in Law 2014; used with permission.
Lacandon) (Richards 2003; Law 2013). Two languages, Itzaj and Mocho’, have fewer than 100 remaining speakers and are quite clearly moribund (Hofling 2000; Palosaari 2011). Ch’olti’ and Chicomuceltec are now extinct, fairly recently in the latter case (Campbell and Canger 1978, Law 2014).

While the health of some Mayan languages appears to be secure, at least for the near future, a pattern of gradual language shift toward Spanish can be observed even in language communities with large numbers of speakers (see e.g. Heinze 2004; Espantzay Serech 2006; French 2010 and references there for Kaqchikel, and England 2002, 2003; Richards 2003; Barrett 2008 for Guatemalan Mayan languages more generally). Literacy rates in Mayan languages are generally quite low (Richards 2003), though literacy has improved somewhat with the increased availability of Mayan language education and Mayan language publications (see England 2003; Brody 2004 and the contributions to Fischer and Brown 1996b for details). Still, as with overall language health, the presence and effectiveness of these materials and publications varies greatly. Historically, education for Maya people has been motivated by the goal of assimilation into the Spanish speaking mainstream (King 1994; Fischer and Brown 1996a). Only recently (1987 for Guatemala and 2010 for Mexico) have official alphabets been developed for the sole purpose of written production in Mayan languages (Brody 2004; Cruz Gómez 2010).

2.4 Language names

Throughout this volume we adopt the naming conventions recommended for Mayan languages by the Academia de Lenguas Mayas de Guatemala (ALMG, the state institution responsible for supporting Mayan languages in Guatemala) and by the Instituto Nacional de Lenguas Indígenas (INALI, the state institution responsible for supporting indigenous languages in Mexico) (see also Mateo-Toledo 2003). Some widely-known alternative language names are given in parentheses in (1). This list of alternative names is not exhaustive: to give just two examples, a common autonym for Uspantek is Tz’umun Kaab’ ‘sweet hummingbird’, and Lak Ty’añ ‘our (inclusive) speech’ is used by Ch’ol speakers to refer to their language. However Uspantek and Ch’ol, respectively, are the officially recognized names for these language, and the ones used in most publications.

A few words on the spelling of Mayan language names: alongside now-standard spellings such as

---

2These recommendations can be found at the ALMG website (http://www.almg.org.gt/index.php/comunidades/) and in INALI (2009) (available online at http://www.inali.gob.mx/pdf/CLIN_completo.pdf). We diverge from the recommendations of INALI in using Yucatec rather than Maya, and in using Mocho’ rather than Quto’k. We also omit the accent mark on Lacandón.
*K’iche’*, one also encounters antiquated Hispanicized spellings like *Quiché*; these spellings are no longer in use, and should be avoided (see again Mateo-Toledo 2003). There are also some common spelling variants involving the interchange of *c~k* (e.g. *Yucatec~Yukatek*), *ts~tz* (e.g. *Tsotsil~Tzotzil*), vowel length (e.g. *K’iche’~K’ichee’*), accent marks (e.g. *Mocho’~Mochó*), and the presence or absence of a final *-o* in names that end in *k/e* (e.g. *Uspantek~Uspanteko*; these vowels correspond to inflectional gender suffixes in Spanish). A few points of variation are limited to individual languages or sub-families (e.g. *Huastec~Wastek, Ch’ol(an)~Chol(an)*; the latter is a point of variation among different Ch’ol dialects). Again, we recommend using the names in (1), allowing for the fact that authors may differ as to whether they employ an o-final form for languages like *Uspanteko, Tekiteko, Yucateco*, etc.

### 3 Reference grammars and other resources

The three overview articles in this volume draw on a wealth of existing work on Mayan languages, which make up the most well-documented of Mesoamerican language families. Formal work on these languages began in the 1970s, and more recently, Mayan linguistics has benefited immensely from the many contributions of native-speaker linguists. England and Zavala (2013) compile a comprehensive and recent annotated bibliography, to which we refer the reader for further information. Our attention here is focused on grammatical characteristics of the contemporary languages; Law (2013, 2014) and Campbell (to appear) provide a thorough overview of Mayan historical linguistics.

#### 3.1 Anthologies

The late 1970s and early 1980s saw an explosion of work on Mayan linguistics, including McClaran (1976), England (1978), as well as issues of the short-lived *Journal of Mayan Linguistics* (three volumes from 1979–1981). England and Elliott (1990) is an important collection of descriptive and analytical articles from the tail-end of this period. More recent compilations include the proceedings of a special session on Mayan linguistics at the annual meeting of SSILA (Avelino 2011), as well as proceedings of the first two meetings of *Form and Analysis in Mayan Linguistics* (FAMLi; Shklovsky et al. 2011; Clemens et al. 2012). A forthcoming volume, edited by Judith Aissen, Nora England, and Roberto Zavala, will contain articles on many topics in Mayan grammar as well as sketches of representative languages in the family (Aissen et al. forthcoming).
### 3.2 OKMA and CIESAS

Among the world’s endangered and minority languages, Mayan languages are remarkable in the amount and range of scholarship produced by native-speaker linguists. As documented in greater detail in England (2007), Maya-speaking linguists began to emerge in the early 1970s as part of the *Proyecto Lingüístico Francisco Marroquín* (PLFM) in Guatemala, and then as part of OKMA (*Oxlajuuj Keej Maya’ Ajtz’iib’*) in the late 1980s, under the leadership of Terrence Kaufman and Nora England. Since its inception in 1991, the Mexican Centro de Investigaciones y Estudios Superiores en Antropología Social (CIESAS) has graduated nearly 100 MA students from its Indoamerican Linguistics program. The program is designed specifically for native speakers of languages of Mesoamerica and roughly one quarter of the students have been speakers of Mayan languages.

The overview articles that follow have relied heavily on these works, but the benefits of scholarship by native speakers reaches far beyond linguistic documentation and analysis. As England notes:

“It is striking that in addition to producing excellent basic documentation and advanced linguistic analyses of their languages they have also sought ways to apply linguistics to education and language policy. For Mayas, linguistics and such applications are not separate endeavors, but rather part of a common goal of maintaining and increasing their political and cultural space in society” (England 2007, 1).

Maya-speaking linguists have gone on to make contributions to language policy and pedagogy, as well as to PhD programs in linguistics locally and abroad.

### 3.3 Reference Grammars

The table in (2) is an incomplete compilation of resources on the grammars of contemporary Mayan languages. These include full reference grammars, pedagogical grammars, and—for languages which do not have a full grammar—some works with relatively comprehensive grammatical sketches. We include these here without detailed commentary, though see England and Zavala 2013 for an annotated bibliography, as well as references cited throughout the three chapters to follow.
<table>
<thead>
<tr>
<th>Yucatecan</th>
<th>Yucatec</th>
<th>Blair (1964); Blair and Vermont Salas (1965); Straight (1976); Bricker et al. (1998)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Itzaj</td>
<td>Hofling (2000); Schumann Galváz (1971, 2000); Zocal Chayax et al. (2001)</td>
<td></td>
</tr>
<tr>
<td>Lacandon</td>
<td>Bruce (1968); Berqvist (2008)</td>
<td></td>
</tr>
<tr>
<td>Mopan</td>
<td>Comunidad Lingüística Mopan (2001); Hofling (2011)</td>
<td></td>
</tr>
<tr>
<td>Huastecan</td>
<td>Huastec</td>
<td>Edmonson (1988); Constable (1989)</td>
</tr>
<tr>
<td>Ch’olan–Tseltalan</td>
<td>Ch’ol</td>
<td>Schumann (1973); Attinasi (1973); Warkentin and Scott (1980); Vázquez Álvarez (2011)</td>
</tr>
<tr>
<td></td>
<td>Chontal</td>
<td>Knowles (1984); Pérez González (1985)</td>
</tr>
<tr>
<td></td>
<td>Ch’orti’</td>
<td>Pérez Martínez (1994)</td>
</tr>
<tr>
<td></td>
<td>Tsotsil</td>
<td>Sarles (1966); Haviland (1981); Aissen (1987)</td>
</tr>
<tr>
<td>Q’anjob’alan</td>
<td>Tojolab’al</td>
<td>Furbee-Losee (1976); Curiel Ramírez del Prado (2007)</td>
</tr>
<tr>
<td></td>
<td>Chuj</td>
<td>Hopkins (1967); Domingo Pascual (2007); Buenrostro (2013)</td>
</tr>
<tr>
<td></td>
<td>Q’anjob’al</td>
<td>Raymundo González et al. (2000); Comunidad Lingüística Q’anjob’al (2005)</td>
</tr>
<tr>
<td></td>
<td>Akatek</td>
<td>Zavala (1992)</td>
</tr>
<tr>
<td></td>
<td>Popti’</td>
<td>Day (1973); Craig (1977); Comunidad Lingüística Jakalteka (2001)</td>
</tr>
<tr>
<td></td>
<td>Mocho’</td>
<td>Palosaari (2011)</td>
</tr>
<tr>
<td>Mamean</td>
<td>Mam</td>
<td>England (1983); Pérez Vail and Jiménez (1997); Pérez et al. (2000); Rojas Ramirez et al. (2002)</td>
</tr>
<tr>
<td></td>
<td>Tektitek</td>
<td>Pérez Vail (2007)</td>
</tr>
<tr>
<td></td>
<td>Awakatek</td>
<td>Comunidad Lingüística Awakateka (2001); Delgado Rojas and Mateo Toledo (2006)</td>
</tr>
<tr>
<td></td>
<td>Ixil</td>
<td>Ayres (1991); Poma et al. (1996)</td>
</tr>
<tr>
<td>K’ichean</td>
<td>Kaqchikel</td>
<td>García Matzar and Rodríguez Guaján (1997); García Matzar et al. (1999); Majzul et al. (2000); Brown et al. (2006)</td>
</tr>
<tr>
<td></td>
<td>Tz’utujil</td>
<td>Dayley (1985); García Ixmatá (1997)</td>
</tr>
<tr>
<td></td>
<td>K’iche’</td>
<td>Mondloch (1981); Larsen (1988); López Ixcoy (1997); Par Sapón and Can Pixabaj (2000)</td>
</tr>
<tr>
<td></td>
<td>Achi</td>
<td>Antonio Lopez and Sis Iboy (1998)</td>
</tr>
<tr>
<td></td>
<td>Sakapultektek</td>
<td>DuBois (1981); Comunidad Lingüística Sakapultekteka (2001); Mó Iséém (2007)</td>
</tr>
<tr>
<td></td>
<td>Sipakapense</td>
<td>Barrett (1999); Comunidad Lingüística Sipakapense (2005)</td>
</tr>
<tr>
<td></td>
<td>Poqomam</td>
<td>Smith-Stark (1983); Santos Nicolás and Benito Pérez (1998)</td>
</tr>
<tr>
<td></td>
<td>Poqomch’i’</td>
<td>Malchic Nicolás et al. (2000)</td>
</tr>
<tr>
<td></td>
<td>Q’eqchi’</td>
<td>Tzul and Cacao (2002); Caz Cho (2007)</td>
</tr>
<tr>
<td></td>
<td>Uspantek</td>
<td>Can Pixabaj (2007)</td>
</tr>
</tbody>
</table>
4 A note on transcription

4.1 Mayan orthographies

In the modern era Mayan languages have been written in a range of different orthographic systems, both standardized and ad hoc (see England 1996; Kaufman 2003 for some historical discussion). (These systems are all based on the Latin alphabet; for an introduction to Mayan hieroglyphics see Coe and Van Stone 2005.) In this volume we follow the orthographic norms developed for Mayan languages by ALMG and INALI. In collaboration with local linguistic communities ALMG has proposed standard orthographies for all 22 officially-recognized Mayan languages in Guatemala. INALI has developed standard orthographies for 8 Mayan languages spoken in Mexico: Yucatec, Ch’ol, Tseltal, Tsotsil, Tojolab’al, Mocho’, Mam, and Tektitek (all except Yucatec are spoken in the Mexican state of Chiapas; see http://www.inali.gob.mx/es/difusion/publicaciones.html). To the best of our knowledge there are no official orthographies for Itzaj, Lacandon, Mopan, Chontal, Huastec, or the extinct Chicomuceltec and Ch’olti’. For these languages we have generally followed the orthographic conventions used by the authors of the primary source materials we draw on.

There are two prominent differences between the standard orthographies employed in Guatemala and those used in Mexico. First, the affricates /ts ts’/ are written as ⟨tz tz’⟩ in Guatemala, but as ⟨ts ts’⟩ in Mexico. Second, the glottalized /b/ is always written as ⟨b’⟩ for Guatemalan Mayan languages, but most Mexican orthographies omit the apostrophe ⟨b⟩, reflecting phonetic differences in the realization of this segment across languages (see Bennett this volume).

In writing this volume it has often been necessary to re-transcribe data from earlier sources so that it conforms to these modern orthographic standards. One frequent substitution was the replacement of ⟨ʔ⟩ with standard ⟨‘⟩ for the glottal stop (IPA /ʔ/), which is also no longer represented in word-initial position (see Kaufman 2003 for discussion). Other substitutions included replacing ⟨ñ⟩ with standard ⟨nh⟩ (IPA /ɲ/) and ⟨ʌ⟩ with standard ⟨ã⟩ (IPA /ã/ or /i/, depending on the language).

---

3The orthographic norms developed by INALI for Tektitek refer to that language as Kakchikel. This nomenclature reflects an unfortunate historical misunderstanding rather than any special affinity with the Mayan language Kaqchikel spoken in Guatemala (Kaufman 1969).
4.2 Abbreviations used in glosses

The abbreviations used in glosses throughout the articles in this volume are provided in (3) below. In most cases, we follow the Leipzig Glossing Rules, with some Mayan-family specific additions and exceptions. Notably, we follow Mayan tradition in using the labels “Set A” (ergative/possessive) and “Set B” (absolutive) for the two series of person markers found across the family (Ch’orti’ also has a third, known as “Set C”). Though more opaque to non-specialists, these labels avoid committing to any particular analysis, for example about whether a verb form has been nominalized and contains a possessor or a true verbal subject; this becomes especially useful in cases of split ergativity (Coon, this volume).

We make frequent use of the abbreviations -ITV, -TV, and DTV for “intransitive verb”, “transitive verb”, and “derived transitive verb” status suffixes, respectively. We depart from Leipzig Glossing Rules in using -NML rather than -NMLZ for nominal suffixes, as the nominal suffixes in question are not always category changing (i.e. not nominal-izers). Some other glossing conventions—for example, whether null morphemes are represented—are discussed as they arise in this volume.

(3) Glossing Conventions

| 1, 2, 3 1ST, 2ND, 3RD person | EXH exhortative | OPT optative |
| A set A (ERGATIVE, GENITIVE) | EXT existential | PASS passive |
| ACAUS anticausative | EVID evidential | PLRC pluractional |
| AF Agent Focus | FOC focus | PRF perfect |
| AFF affective | IMP imperative | PFV perfective |
| ANTIP antipassive | IMM immediate | POS positional suffix |
| APPL applicative | IPFV imperfective | POT potential |
| ASP aspect | INCH inchoative | PL plural |
| ATT attenuator | INC incompletive | PREP preposition |
| B set B (ABSOLUTIVE) | INCL inclusive | PROG progressive |
| CAUS causative | INDF indefinite | PRON pronoun |
| CL clitic | IDF ideophone | PROSP prospective |
| CLF classifier | INF infinitive | PRES present |
| COMP complementizer | INS instrumental | Q question particle |
| COMPL completive | INTS intensifier | REL relative clause suffix |
5 A disclaimer

Finally, it is important to recognize what we have not been able to do in these three articles. First, as the titles suggest, we do not cover all subfields of linguistics. In particular, the large sociolinguistic and historical linguistic literature on Mayan languages is not addressed in any detail. This is perhaps not ideal given that much is known about the historical development of Mayan languages, as well as the sociolinguistic contexts in which they have been spoken, both pre-colonially as part of the Mesoamerican cultural area and post-colonially where the influence of European languages like Spanish and English has become dominant. For references and overviews see Law 2013, 2014; Campbell to appear.

Perhaps more controversially, even within those subfields we have addressed, we have not been able to fully represent all theoretical perspectives on the data presented. Instead, we have taken a broadly generative stance. This is primarily due to the backgrounds of the authors and is not meant to suggest that this is the only way to successfully analyze the grammars of Mayan languages. That said, none of the articles are particularly theoretical. Instead, we have tried to front the empirical phenomena themselves, and so non-generative works have been cited for raw data and empirical generalizations.

Finally, while the articles are meant to address the family as a whole, not all Mayan languages have been equally represented. This is partially due to authors’ backgrounds—none of the authors have done original research in more than two of the sub-groups of Mayan languages. More than that, though, the literature on Mayan languages is itself uneven. For instance, there has been much more research done on the semantics
of Yucatec than any other Mayan language, and so our semantics article almost certainly has a bias toward that language. With that caveat in mind, we have at least tried to balance our articles with respect to their coverage of the major sub-groups of Mayan languages.

References


Brody, Michal. 2004. *The fixed word, the moving tongue: Variation in written Yucatec Maya and the meandering evolution toward unified norms*. Doctoral Dissertation, University of Texas Austin, Austin, TX.


England, Nora. 2007. The influence of Mayan-speaking linguists on the state of Mayan linguistics. In


Vázquez Álvarez, Juan J. 2011. *A grammar of Chol, a Mayan language*. Doctoral Dissertation, University of Texas Austin, Austin, TX.


