USAID
FROM THE AMERICAN PEOPLE

All Children Reading–Philippines

Characteristics of Select Philippine Mother Tongue Languages Used in Basic Education Teaching and Learning

A Reference Document

Submission Date: February 14, 2021

AID-OAA-TO-16-00017 Number: REQ-ASIA-16-00017
Task Order Contracting Officer's Representative: Mitch Kirby

Submitted by: **RTI International**
3040 East Cornwallis Road
Research Triangle Park, NC 27709-0155
Telephone: (919) 541-6000

Authored by: **Jason Lobel (University of Hawai'i at Mānoa) and Sarah Pouezevara (RTI International)**

Table of Contents

List of Tables

Acronyms

AM	amplitude modulation
EGRA	Early Grade Reading Assessment
FM	frequency modulation
GUMIL	Ilokano Writers Association of the Philippines (*Gunglo dagiti Mannurat nga Ilokano iti Filipinas*)
KWF	Commission on the Filipino Language (*Komisyon sa Wikang Filipino*)
L2	second language
L3	third language
MT	mother tongue
MTB-MLE	Mother Tongue–Based Multilingual Education
USAID	United States Agency for International Development

1 Preface

1.1 About This Document

This reference document is a companion to the Language Complexity Study conducted by RTI International under the All Children Reading–Philippines project in 2020. The study was a secondary analysis of Early Grade Reading Assessment (EGRA) data at multiple time points, looking specifically at the effect of language complexity on reading acquisition of second and third languages (L2 and L3, respectively). The study used the 2013 and 2019 National EGRA data sets to analyze performance according to categories of language complexity based on syllabic complexity, orthographic depth, and other related items. The study was guided by the methodology described by Brunette et al. (2019), who studied the effects of language complexity on reading outcomes in Uganda. The languages selected for analysis were those among the officially supported mother tongue (MT) languages of instruction in the Philippines. Analysts assessed whether the level of complexity was in any way predictive of the average increase or decrease in L2 and L3 scores across schools as measured in 2019 (provided there were schools that reported using MT in 2019, which was not the case for three of the languages: Chavacano, Ivatan, and Sambal.

Dr. Jason Lobel of the University of Hawai'i at Mānoa, a linguist with extensive background in Philippine and North Bornean languages and a consultant to RTI, established the language complexity categories. The central part of this current document (see Section 2) is the Dr. Lobel's work, which describes the features of each of the languages studied and how he used these features to classify each language into a category of complexity. He then used this complexity value as the outcome variable in regression analyses of the EGRA scores.

This current report is also intended to serve as a reference guide for individuals developing literacy instruction materials or early literacy assessments to help identify elements of the languages that may be particularly difficult to teach or that have an impact on developing reading instruction materials for young learners. Understanding the characteristics of languages, such as deep (opaque) versus shallow (transparent) orthographies, can be helpful when designing more effective literacy instruction.

RTI editors have edited this report for readability and style, but this document remains targeted toward technical audiences working in literacy development and assessment. Therefore, this report includes a considerable amount of technical terminology that is not always explained for readers. RTI has added a simple glossary near the end of this report to establish the meaning of specific technical terms. RTI has also added a brief literature review to explain the factors that influence reading acquisition. However, the remainder of the report is the work of the lead author (Dr. Lobel), and the accuracy of the linguistic characteristics of each language and the basis for their classification by complexity is therefore his sole responsibility. Similarly, language names and their spellings follow academic and historical conventions, according to the author, rather than current official government spellings within the Philippines itself, although to the extent possible, all commonly occurring variations of each language name have been recorded. The language names and spellings used in the headings or text of this document are the choice of the author, are used for reference only, and do not constitute any preference on the part of RTI or the United States Agency for International Development (USAID).

1.2 The Relationship of Complexity to Learning to Read

For linguists, complexity is a quantifiable property of a language, associated with its phonetic and morphological inventory, such that languages with more phonological and/or morphological distinctions and interconnections between language units are considered to be more complex (Biber et al., 2011; Bulté & Housen, 2014; Dahl, 2004; Miestamo, 2003). Miestamo (2003) notes the following: "Whether and how the different aspects of complexity

defined this way contribute to difficulty for different types of language users is a very important question, but a separate one, to be addressed with the help of psycholinguists studying language processing and acquisition." Therefore, whereas "absolute" complexity refers to an objective property of the language, "relative" complexity relates to the cost or difficulty those properties pose for the language users. This is precisely the type of complexity in which this study is interested—do the properties of a language of first literacy have an impact on how difficult or easy it will be to acquire L2 and L3 and literacies?

There are several ways in which we hypothesize that language complexity may influence literacy acquisition: orthographical (spelling), phonological (pronunciation), morphological (units of meaning), and sociolinguistic (exposure to literacy day-to-day). The most basic theory of reading acquisition states that for alphabetic languages, children must learn to associate sounds, or phonemes, with the corresponding written symbol—the grapheme. Learning letter-sound correspondence and the ability to manipulate sounds in words has been shown to be a strong predictor of later reading achievement (Dubeck & Gove, 2015). Some orthographies will be easier to learn than others (Borleffs et al., 2019). There is considerable evidence that the rules of transparent orthographies are acquired more easily than opaque ones (Borleffs et al., 2019). Some languages are highly transparent, meaning that there is a consistent one-to-one relationship between letters and sounds. As summarized in the EGRA Toolkit (RTI International, 2016): "A child learning to read in a consistent, transparent orthography of a language with relatively low phoneme inventory, simple syllable structures, and short average word lengths will be at an advantage for mastering the letter–sound mappings and decoding skills more rapidly than a child learning to read in a language with an opaque orthography, many irregularities, many phonemes, complex syllable structures, and long average word lengths."

English is an example of a highly non-transparent, or opaque, language in which sounds may be represented by multiple letters (the phoneme /k/ may be spelled with the letters c, k, ck, ch, and qu); conversely, the letter may represent different phonemes (the letter c may represent the phoneme /k/ in one word and /s/ in another) (RTI International, 2016). In some languages, the letter-sound correspondence changes depending on the letter's position within the word or due to adjacent words (the bird versus the ant); its meaning ("read" in present tense versus past tense), or derivative forms (heal versus health) (Borleffs et al., 2019). Sounds can also map onto more than one letter. McEachern (2013) notes that "sounds [ŋ], [ɡ̟], [ʃ], and [ʤ] exist in many Philippine languages, but the orthography generally employs digraphs to spell them: ng, ts/ti/ty/tiy, sy/si/siy, and dy/di/diy, respectively. Moreover, most languages have fewer consonant and vowel phonemes than English, and word-initial consonant clusters such as /str/, /pr/, /sk/, /sp/ and /br/, are rare and found mostly in loan words. Consonant clusters such as /ks/ and /sp/, are, however, found in word-medial position. Finally, but not exhaustively, one sound can be represented by more than one letter, as in the case of English phonemes "th" [ð] and "ph" [f]. Words containing these digraphs could pose a challenge to beginning readers who may try to decode them by sounding out each grapheme separately, without realizing that the digraphs should be read as one sound. This is where the influence of strong, language-specific instruction strategies, and cross-language bridging strategies, is critical.

Learning letter-sound correspondence is highly dependent on one's mastery of the oral language and the phonology of the language, which is acquired in one's home language well before literacy learning begins (Milledge & Blythe, 2019). Reading is the process of mapping phonological and semantic knowledge onto the orthographic forms of words (Frost, 1998). Learning letter-sound correspondence can be made more complex by a given language's rules for *graphemic parsing* (determining whether a sound is represented by a letter or a letter cluster), (Storkel, 2001). *Phonotactic regularity*, or the likelihood that a sequence of sounds will occur, is posited to have an impact on oral language development because common sound sequences are learned more quickly (Storkel, 2001). Consequently, knowing common sound sequences can help children identify new or unfamiliar words and add new words to their lexicon more quickly (because they are based on known phonological patterns). According to Coady & Aslin (2004), "The more similar a new word is to other

words already in the lexicon, the more readily it will be learned. Presumably, then, learning will be facilitated for those words that contain the more frequent sounds and sound combinations." Storkel (2001) also noted that older children have been found to be more sensitive to larger units of sound (such as diphones). This sensitivity may have implications for instructional sequencing of literacy tasks particularly for non-native languages for which the phonology is not learned from birth.

Other characteristics of a language that may influence how difficult or easy it is to learn include *morphological complexity* and *word length*. Morphemes are the smallest unit of meaning of a word, for example, word roots, prefixes, and suffixes. Being able to read a word root will facilitate reading words based on that same root (Elbro & Arnbak, 1996); thus, morphological awareness is also correlated with word reading ability in some languages and continues to be an important skill for reading development in upper grades after basic decoding has been acquired (Borleffs et al., 2019). Brunette et al. (2019) noted that the average length of words may affect reading outcomes and is an important factor in interpreting the traditional measurement of oral reading fluency in the form of correct words per minute. Agglutination (when affixes attach to root words) in a language can affect word length (Abadzi, 2012), and all Philippine languages are agglutinative. However, both small and large amounts of morphology can attach to a word root; therefore, this characteristic alone does not necessarily imply longer word length, particularly for early reading materials. Agglutination has more of an effect in higher level written materials (Bible translations, newspaper articles), whereas the most common scenario in readings for the lower primary grades would be a two-syllable verbal root, and one or two affixes consisting of one syllable each. As a result, in readings for the lower primary grades, most verbs are going to be three or four syllables long, whether in a morphologically, less complex language such as Cebuano or a morphologically more complex language such as Bikol or Waray-Waray.

Sociolinguistic factors that put some languages at a disadvantage in terms of effective teaching and learning in the public school system include the level and quality of support for teaching (e.g., books, teacher training, leveled readers) and the extent to which the language is used in print outside of school. When multiple languages exist side by side in a given community and are used interchangeably by the population, it may be difficult for a child or parent to identify the child's "first" language. Some languages are spoken but are rarely written, and some languages lack a tradition of print media; therefore, orthographic forms are inconsistent. Dialectal variation among languages may become problematic when geographic variations of a language evolve to use different words, not just variation in the pronunciation of words. For example, there is basically zero variation in pronunciation among Ilonggo and Bikol dialects, but there are words that are completely different. Having an entirely different word does not affect learning to read the orthography, but it creates a challenging situation for teachers, curriculum and materials developers, and for those who develop educational assessments or any other materials that would otherwise be mass produced and distributed. These materials would have to ensure that the words that appear in texts and tests would, as much as possible, be words that are understood in all the dialects of the language.

1.3 Languages in the Philippines (RTI)

The Philippines is considered to be one of the 10 most linguistically diverse countries in the world, and the second most diverse in Southeast Asia (Lewis et al., 2013). The Philippines is a country of approximately 180 languages, ranging in population from the single digits to more than 20 million. Out of these 180 languages, eight are traditionally considered "major" languages, defined in the latter portion of the twentieth century as having more than a million speakers each. The eight languages are as follows: Cebuano, Central Bikol, Ilokano, Ilonggo/Hiligaynon, Kapampangan, Pangasinan, Tagalog, and Waray-Waray. These languages, as well as several other regional languages such as Chavacano, Maguindanaon, Maranao, and Tausug, and were adopted as languages of instruction under the Mother

Tongue–Based Multilingual Education (MTB-MLE) curriculum launched in 2012. Although all Philippine languages, with the exception of Chavacano, belong to the Austronesian language family, their orthographies were adopted from the Spanish language family, with subsequent alterations, and their vocabularies have also been influenced by periods of colonial influence. In particular, 377 years of Spanish colonial influence left behind many proper nouns and place names, and some borrowed words that require the use of letters that would not otherwise be part of the native phonology.

In 1937, in an effort to build a sovereign nation, even while the country was still under American occupation, the Philippine government declared (Executive Order 134, 1937) establishment of one national language, based on Tagalog a Philippine language that was native to Manila, the national capital, and surrounding provinces. The national language has gone through different phases of acceptance, and is now known as Filipino, as defined in the 1987 Philippine Constitution as a language that should be enriched from other local and non-local languages. In reality, despite efforts to include non-Tagalog words of Philippine origin, the structure and lexicon remain almost identical to Tagalog (Gonzalez, 1998; McFarland, 1994). As such, Filipinos use the two words interchangeably, or vacillate between using Tagalog to refer to the language as spoken in its native region, and Filipino in the context of its wider use across the country (McEachern, 2013). This distinction is how the Philippine Department of Education (DepEd) distinguishes between implementing Tagalog as a MT in areas where no other lingua franca is present, but Filipino as the official language of instruction throughout the nation. The Commission on the Filipino Language (*Komisyon sa Wikang Filipino* [KWF]) now controls decisions about orthography development and standardization for all Philippine languages used as a medium of instruction.

This current report provides brief overviews of each of the previously mentioned languages in terms of the complexity of their phonological, orthographical, and morphological systems, as well as a rapid scan of sociolinguistic issues, including the availability of print and audiovisual materials in each language. The information is based on a combination of academic reference sources and personal exposure to the language families. The languages are described separately, but there are important overlaps across all of the languages that result from the influence of foreign loan words, including the use of “foreign” letters (i.e., those used exclusively in words, especially proper nouns, of foreign origin); the lack of marking of glottal stops and word stress; and the agglutinative “Philippine-type” morphology that contributes to the substantial length of nouns and verbs.

2 Language Descriptions

2.1 Tagalog

One of only two Philippine languages with more than 10 million native speakers, Tagalog, which now has an estimated 20 million native speakers (Eberhard et al., 2020)[1], was first declared as the national language of the Philippines by President Manuel L. Quezon in 1937 (Quezon, 1937). Tagalog was subsequently reaffirmed as the national language in 1959 under the name “Pilipino,” and, in 1973, as “Filipino,” a name that was retained in the Constitution of 1986 after the fall of the Ferdinand Marcos dictatorship (Gonzalez, 2007). As the Philippine language with the longest history of use in formal education, its orthography has undergone a number of changes. The Spanish-based spelling system that first replaced the native *Baybayin* orthography that was in use 500 years ago gave way to a 20-letter alphabet named *abakada* (after its first four letters). *Abakada*, in turn, yielded a more complicated alphabet under the Pilipino years, which attempted to incorporate the foreign (i.e., Spanish and English) letters and digraphs used in personal names and place names. The current official orthography known as the “Modern Filipino Alphabet” has been in use

[1] Eberhard et al. (2020) is also the source for all other population statistics in subsequent sections of this report.

since 1987, including in Department of Education instructional materials. It is a 28-value alphabet that incorporates all 26 of the English letters, plus the Spanish ñ and the longstanding Philippine digraph ng (which, as in English, spells the velar nasal /ŋ/), as presented in **Table 1**.

Table 1. The National (or Modern Filipino) Alphabet

a	h	ñ	t
b	i	ng	u
c	j	o	v
d	k	p	w
e	l	q	x
f	m	r	y
g	n	s	z

Note that although the letters <c, f, j, ñ, q, v, x, z> are not necessary in representing the phonemes native to any of the Philippine MTB-MLE languages,[2] their use in spelling the languages cannot be avoided because of personal names and place names that are either of foreign origin or are spelled as if they were. Some examples include Bulacan, Carlos, Cavite, Cebu, Cotabato, Davao, Juan, Luzon, Mindanao, Nueva Ecija, Nueva Viscaya, Rizal, Soccsksargen, Visayas, and Zamboanga. As such, they are considered part of the spelling system of Philippine languages that children will need to learn to recognize and process.

The native phonology of Tagalog consists of 16 consonants and three vowels (**Table 2**). With the exception of recent loan words or proper nouns that retain their original and often opaque spellings, Tagalog orthography is highly transparent.

Table 2. The Phoneme System of Tagalog

	Consonants				**Vowels**			
	Bilabial	**Dental, Palatal**	**Velar**	**Glottal**		**Front**	**Central**	**Back**
Stop	p	t	k	ʔ	High	i		u
	b	d	g		Mid			
Nasal	m	n	ŋ		Low		a	
Lateral		l						
Flap		r						
Fricative		s		h				
Glide	w	y						

In addition to these native phonemes, some Tagalog speakers differentiate /e/ from /i/, and /o/ from /u/. This differentiation is often considered to be the result of the introduction of Spanish loan words (because Spanish has a native vowel system of /a e i o u/). However,

[2] Note, however, that a phoneme /ʤ/ (i.e., the "j" sound) does exist in Cebuano dialects found throughout most of Bohol Island, as well as southwestern Leyte Island, and that in Maranao, orthographies commonly used by native speakers over the past half century, the letter "z" is used as an alternate spelling for the heavy fricative otherwise spelled as "sh."

speakers of many rural dialects of Tagalog still clearly lack a distinction between /e/ and /i/ and between /o/ and /u/, as in the observable variation between [pero], [peru], [piro], and [piru] when pronouncing the word *pero* (Spanish 'but').

Three things should be noted about Tagalog orthography. First, the **glottal stop (/ʔ/)**—although fully phonemic in all positions—is only written consistently between consonant-final prefixes and orthographically vowel-initial root words, where it is written as a hyphen. The glottal stop is unwritten in word-final position, even though it is involved in such minimal pairs as *baga* /bága/ 'embers' and *baga* /bágaʔ/ 'lungs.' The glottal stop does not occur preconsonantally in Tagalog, but does occur in postconsonantal position in some rural dialects of Tagalog (e.g., *bang-áw* /baŋ.ʔaw/ 'crazy, rabid'; *gab-i* /gab.ʔi/ 'night'; *ngay-on* /ŋay.ʔun/ 'today, now'). Second, **stress** is unmarked in spite of being phonemic (meaning that the meaning of a word can change depending on the syllable on which the stress is placed), e.g., *bukas* /búkas/ 'tomorrow' versus *bukas* /bukás/ 'open,' two words which, despite sharing the same spelling, are synchronically unrelated to one another. A system of diacritics exists that marks both word stress and word-final glottal stop, yet is only rarely used in publications and is never used in handwriting.

A third orthographical characteristic inherent in the Tagalog orthography, as well as those of most other Philippine languages with any significant time depth of printed work, is a holdover from the Spanish era: using all **five of the vowel graphemes** of Spanish (and English) to write what were originally—and still largely are—three-vowel languages. The spelling of Spanish borrowings usually retained the spellings of the vowels in the Spanish originals, and these forms were juxtaposed over the native lexicon, in which the vowels were spelled largely according to their position in the word. For example, a /u/ in a non-final syllable was spelled as <u>, but as <o> in a word-final syllable (e.g., *tugtog* 'play (an instrument)', *ulo* 'head,' *buhok* 'hair'), although this is not always the case, such as in *totoo* 'true.'

Four more phonemes have been introduced into many dialects of Tagalog and some other Philippine languages because of Spanish (and later, English) borrowings, and later reinforced by phonological processes affecting words in the native strata. These phonemes are written as digraphs or trigraphs, reducing the otherwise transparent, one-to-one correspondence of the orthography. These four phonemes are as follows:

- The "j" sound (/ʤ/), which is spelled alternately as <dy>, <ds>, or <diy>
- The "ch" sound (/ʧ/), which is spelled alternately as <ty>, <ts>, or <tiy>
- The Spanish "ñ" sound (/ɲ/), which is spelled <ny> or <niy>
- The "sh" sound (/ʃ/), which is spelled <sy> or <siy>.

Common Spanish loan words with these sounds include *medyas* /me.ʤas/ ~ /med.yas/ 'socks,' *kotse* /ko.ʧe/ ~ /kut.si/ 'car,' *pinya* /pin.ya/ ~ /pi.ɲa/ 'pineapple,' and *siyudad* ~ *syudad* /ʃu.dad/ ~ /syu.dad/ 'city.'[3] Processes such as palatalization have also resulted into these four sounds encroaching on the native strata of many Philippine languages, including the following:

- Tagalog, where rural Tagalog *siya* [sí.ya] 'he/him, she/her' has become *sya* [ʃa] in Manila Tagalog
- Rural Tagalog *níya* [ní.ya] 'he/his, she/her' has become *nya* [ña] in Manila Tagalog
- Rural Tagalog *tiyagâ* [ti.ya.gáʔ] 'work hard' has become *tyagâ* [ʧa.gáʔ] in Manila Tagalog
- Rural Tagalog *diyan* [dí.yan] 'there' has become *dyan* [ʤan] in Manila Tagalog.

[3] Note, however, that not all Tagalog speakers use these sounds because others pronounce the words as /mid.yas/ 'socks,' /kut.si/ 'car,' and /si.yu.dad/ 'city.'

Notwithstanding the previously mentioned issues, each letter in the Tagalog orthography is transparent regarding its pronunciation (except in place names and personal names, and where recent English loan words have been borrowed unchanged), and even sequences of two written vowels unambiguously indicate two vowels separated by a phonemic glottal stop. Unless appearing as a member of a digraph, no letter has more than one pronunciation, and digraphs are limited to the previously mentioned <ng>, <ny>, <sy>, <ts>, <ty>, <ds>, and <dy>. Note that in Tagalog and in other Philippine languages, the use of letters representing foreign sounds does not mean that those letters are pronounced as in the foreign source languages. On the contrary, most place names and personal names have traditionally been adapted to the phonology of the language into which they have been borrowed.

Similar to other Central Philippine languages, there are few restrictions regarding consonant clusters in Tagalog, in which there are more than 100 permissible combinations of consonants that may occur in root words. Among the few restrictions are the prohibition of /h/ or /ʔ/ before another consonant. Likewise, unlike Ilokano and Tausug, geminate clusters ("long consonants" or sequences of two identical consonants) do not occur in Tagalog.[4] Independent vowel length (i.e., vowel length that is unrelated to word stress) is also not a feature of Manila Tagalog, including the standardized version taught in schools as Filipino, although it is a feature of the vast majority of other Tagalog dialects (e.g., rural Tagalog *nākáin* [naː.ká.ʔin] 'eat(s), is eating' versus *nakáin* [na.ká.ʔin] 'accidentally got eaten, was able to be eaten.'

Syllable structure in the native stratum is CV(C), because underlyingly, all orthographically vowel-initial words actually begin with a glottal stop, and consonant clusters did not occur syllable internally. Underlyingly, all orthographically vowel-final words actually end in /h/, so the structure of a two-syllable word can be represented as CV(C)CVC. However, this structure is more of a theoretical abstraction because the influx of Spanish and, more recently, English borrowings has caused most native speakers of Philippine languages to grow accustomed to pronouncing words of other syllabic structures (e.g., Spanish borrowings such as *problema* 'problem' and *trabaho* 'work' with initial consonant clusters, and *istrikto* 'strict' with word-medial clusters of three consonants). Phonological processes have also restructured words in the native lexicon to other syllabic structures. For example, in rural Tagalog *buwan* [bú.wan] is 'moon,' but the word in Manila Tagalog is *bwan* [bwan]; in rural Tagalog *tuwi* [tú.wiʔ] is 'each, every (time),' but the word in Manila Tagalog is *twing* [twiŋ]. Repetition can then result in words of effective CCVC, such as the CCVC structure in Manila Tagalog (e.g., *bwan-bwan* /bwanbwan/ 'monthly'). Restructured English borrowings with word-final consonant clusters have long been a part of colloquial Tagalog (e.g., *pers* 'first,' *prend* 'friend,' *drayb* 'drive,' and *layk* 'like') and even classic Tagalog slang such as *ermats* 'mother' and *erpats* 'father.' Currently, the ubiquity of television, cell phones, and the Internet in areas covering the vast majority of the Philippine population has resulted in children and teenagers being constantly exposed to English, and therefore accustomed to pronouncing common English words regardless of their syllabic structure, ultimately incorporating them into spoken Tagalog and into most other Philippine languages. Note finally that similar to most other Philippine and Philippine-type languages, the vast majority of root words in the native stratum are disyllabic, with only a small percentage consisting of three or more syllables, and although many grammatical function words are monosyllabic, words in the general vocabulary could not traditionally consist of fewer than two syllables.

Tagalog has very little in the way of phonotactic or morphophonemic alternation (i.e., situations in which the sound of a letter changes based on derivative forms of the word). The only regular morphophonemic alternations in Tagalog occur with prefixes *maN-*, *naN-*, *paN-*, and *pinaN-*, in which, as in other Malayo-Polynesian languages, the final nasal phoneme, symbolized as "N," assimilates as /m/, /n/, or /ŋ/, depending on the initial phoneme of the root to which it is prefixed. Other alternations are limited to irregular forms, in which an /n/ is

[4] A small number of rural northern Tagalog dialects have, however, borrowed words with geminate consonants from neighboring languages.

inserted in the suffixed forms of verbs such as *kúhà* ‘get’ (e.g., *kukunin* ‘will get’ instead of expected ***kukuhain*) and *túto* ‘learn’ (e.g., *natunan* ‘learned’ as opposed to expected ***natutuhan*). These are irregular forms, however, and are not the result of any regular phonological or morphophonemic rule.

In terms of morphology, Tagalog is of average complexity vis-à-vis other Philippine and Philippine-type languages, somewhat less complex than Bikol or Waray-Waray, and somewhat more complex than Cebuano or Ilonggo/Hiligaynon. Similar to all Philippine-type languages, Tagalog is an agglutinative language, whose verb system can mark six focuses/voices: Actor Focus, Object Focus, Location Focus, Secondary Object Focus, Instrument Focus, and Reason Focus. Tagalog’s verb system can also mark singular and plural numbers (although plural forms are rarely used in Manila Tagalog or its textbook “Filipino” form). Also, the following six modes can often co-occur: basic, abilitative/accidental, causative, reciprocal, social, and reflexive. However, there are neither morphological imperative forms[5] nor subjunctive forms in Manila Tagalog, unlike some of the other MTB-MLE languages discussed in this report. As previously mentioned, the forms of the affixes are highly regular, and are therefore easily readable and recognizable for students learning to read.

Out of all of the languages native to the Philippines, Tagalog has, by far, been used in the largest number of publications over the past century, and it is the only language for which non-religious books other than dictionaries can be found at nationwide chain bookstores. A considerable number of small (usually 8-page or 16-page) national newspapers are also available in Tagalog, as is one magazine, the weekly *Liwayway*, which has now been in publication for nearly a century. A number of Tagalog dictionaries are available; the highest quality are the English-Tagalog Dictionary (English, 1977) and Tagalog-English Dictionary (English, 1986). Furthermore, before the onset of the current MTB-MLE policy, Tagalog was also the only Philippine language in which textbooks were available. Similar to the other seven languages traditionally counted as major languages in the Philippines (i.e., Bikol, Cebuano, Ilokano, Ilonggo/Hiligaynon, Kapampangan, Pangasinan, and Waray-Waray), the Catholic Mass has long been widely available in Tagalog, as have a considerable amount of religious books, including more than one translation of the entire Bible. Although most of the regional Philippine languages have at least minimal news broadcasts on television, Tagalog, by nature of being both the language of the national capital and the basis of the Philippine national language, is virtually the only Philippine language with a large amount of televised material. Some examples of televised material include overdubbed versions of foreign cartoons and soap operas, as well as nationally promoted music. However, a very small amount of local music is also being produced in some of the regional languages. Finally, similar to the other major regional languages, Tagalog is also widely used in radio broadcasts on both amplitude modulation (AM) and frequency modulation (FM) radio.

2.2 Cebuano (Also Known as Binisaya, Bisaya, and Sinugbuanong Binisaya)

Cebuano is the only Philippine language whose number of native speakers (approximately 16 million) approaches that of Tagalog. Cebuano is also the only language to rival its geographical scope in the Philippines, spoken as a native language in the central portion of the Visayan Islands and throughout most of central and eastern Mindanao.

Cebuano and the other Bisayan languages included in this report (i.e., Ilonggo/Hiligaynon, Tausug, and Waray-Waray) are largely similar to their close relative, Tagalog (**Table 2**).[6] In most of its dialects, the phoneme system of Cebuano is identical to that of Tagalog. However, as shown in **Table 3**, an additional consonant, “j” /ʤ/, is in some dialects of Cebuano (primarily those spoken throughout Bohol Island and portions of southern Leyte

[5] Some rural Tagalog dialects have one or more imperative forms, as did literary Tagalog in past centuries.
[6] The core Central Philippine subgroup consists of Tagalog; Mamanwa, and the various Bikol, Bisayan, and the Mansakan languages.

Island), and an additional vowel /ə/[7] is found in a more limited portion of Bohol and Leyte Islands. Although the "j" sound is sometimes found written in informal writing by speakers of these dialects, the "schwa"[8] /ə/ is never written as a distinct vowel in spite of being phonemic and is instead misunderstood by non-linguists as a "tense" pronunciation of the vowel /u/.

Table 3. The Phoneme System of Some Cebuano Dialects

	Consonants				Vowels			
	Bilabial	**Dental, Palatal**	**Velar**	**Glottal**		**Front**	**Central**	**Back**
Stop	p	t	k	ʔ	High	i	(ə)	u
	b	d	g		Mid			
Nasal	m	n	ŋ		Low		a	
Lateral		l						
Flap		r						
Affricate		(ʤ)						
Fricative		s		h				
Glide	w	y						

With the exception of unofficial writings in the "j" dialects of Cebuano, Cebuano is, on the surface, orthographically identical to Tagalog, with one exception. Unlike Standard/Manila Tagalog, in which morpheme-internal post-consonantal glottal stop is only usually found in reduplicated monosyllables (e.g., *ap-ap* /ʔapʔap/ ~ *an-an* /ʔanʔan/ 'white skin blotches'), post-consonantal glottal stops are an integral part of all Cebuano dialects and are usually indicated consistently with a hyphen (e.g., *bag-o* /bagʔu/ 'new,' *pus-on* /pusʔun/ 'lower abdomen').

The Cebuano dialect of Cebu City and northern Cebu Island generally drops intervocalic /l/ ("L") except where it borders on /i/; thus, traditional Cebuano *wala* [wa.láʔ] 'none' becomes *wa* [waːʔ], *sulod* [su.lúd] 'interior' becomes *sud* [suːd], and *balay* [ba.láy] 'house' becomes *bay* [baːy]. Traditionally, the literary standard was the L-retaining form. However, as a disconnect developed between modern Cebuanos and traditional Cebuano literature, the L-retaining dialects fell out of favor, and the written form of Cebuano, when it appears in writing, usually vacillates between writing the "L," not writing the "L," and a combination of the two. However, current Department of Education materials appear to follow the more conservative "L-retaining" literary standard. Orthographical decisions are also difficult because, although the "L-retaining" dialects are the more historically correct, the "L-dropping" dialects have become more prestigious by virtue of being the speech of the Cebuano diaspora's capital, Cebu City. Thus, for this otherwise highly transparent language, if the more traditional L-retaining form of the word is used in school materials, then it would likely cause confusion for children who, in most areas, would not immediately recognize the meaning of words, such as the previously mentioned *sulod* and *balay*, if the <l> is pronounced and would need to be explicitly taught to read words with this silent "L."

Similar to Tagalog, the orthography used for Cebuano includes many letters (<c, f, j, ñ, q, v, x, z>) and digraphs (<ll, gu, qu>) from the Spanish/English alphabets that are not otherwise used in the native vocabulary. Some examples include letters such as <e>, <o>, and <c>,

[7] This vowel is more correctly represented with a "barred i" symbol, /ɨ/, but it will be written in this report as /ə/ because of the difficulty of distinguishing the "ɨ" symbol from a regular "i" in most fonts.
[8] The term "schwa" is used in Philippine-centric literature as a shorthand way of referring to what is actually an unrounded, tense high central vowel represented by / ɨ / in the International Phonetic Alphabet.

which are found in the name of the language itself and in the major Cebuano-speaking areas of Cebu and Bohol, in various local place names (e.g., Boljoon, Carcar, Carmen, Catmon, Cebu City, Dalaguete, Mactan, Mandaue, Minglanilla, San Fernando); and in well-known family names (e.g., Quezon). Although not strictly necessary in spelling words from the native stratum of the Cebuano language, most native Cebuano speakers are exposed to words such as these on a daily basis, and, as such, the letters used in spelling them have become part of their orthographic repertoire.

The three characteristics that distinguish Cebuano orthography from being fully transparent are also the same as for Tagalog and most other MTB-MLE languages: the lack of stress marking, the lack of indication of word-final glottal stop, and the use of five vowel graphemes for a three-vowel phoneme language. Cebuano, similar to Tagalog, has little phonotactic or morphophonemic alternation, with the exception of the behavior of prefixes *maN-, naN-, and paN-*, as previously mentioned in Section 2.1 of this report.

Although still possessing a fully productive, agglutinative, Philippine-type morphological system, modern Cebuano has slightly less morphological complexity than perhaps any other MTB-MLE language with the possible exception of Tausug. On one hand, unlike Tagalog, Cebuano has both a morphologically distinct imperative form, and a simple subjunctive form (marked in the same way as the imperative form). On the other hand, the aspect system of Cebuano otherwise marks only a past/non-past contrast, unlike most of the other MTB-MLE languages that mark three-way if not four-way tense/aspect contrasts. The systems of focus and mode in Cebuano are highly similar to those of Tagalog, and the morphology of Cebuano is likewise highly regular.

Similar to the other seven languages historically labeled as major Philippine languages, the Catholic Mass has long been available in Cebuano, as have many prayer books and more than one Bible translation. A number of dictionaries is also available (e.g., Wolff, 1972). A few small (8-page or 16-page) newspapers are available in Cebuano, at least one of which also has regional variants targeted toward areas outside the Central Visayas Region (primarily in Mindanao), and a weekly magazine, *Bisaya*, has been in print since the 1930s. A limited amount of Cebuano television programming is available, primarily in the form of news broadcasts. Cebuano is also used on the radio in Cebuano-speaking areas (primarily on AM stations), but only a small amount of local Cebuano music is evident. Few Cebuano books of any type are currently in circulation.

2.3 Ilonggo (Also Known as Hiligaynon)

Ilonggo[9] is another widespread language, spoken by approximately six million native speakers throughout most of the Western Visayas Region and in specific portions of west-central Mindanao, especially around Cotabato City. Historically, there was substantial confusion over the status of Kinaray-a, another language spoken in the Western Visayas Region, vis-à-vis Ilonggo, and some locals still speak of “Hiligaynon” as having two types: one being Ilonggo, and the other being Kinaray-a. However, Kinaray-a is clearly a separate language not closely related to Ilonggo, and it is also now one of the official MTB-MLE languages supported by the Department of Education curriculum. Another false distinction has been made between Ilonggo/Hiligaynon and Capiznon, even though the latter is clearly a dialect of the former language (and even residents of Capiz Province most often refer to their language as “Ilonggo” and not “Capiznon”), differing in only a relatively small amount of vocabulary from the Ilonggo spoken elsewhere on Panay and Negros Islands. In general, there is much less dialectal variation in the Ilonggo language than in other MTB-MLE languages such as Bikol, Cebuano, Kinaray-a, Maguindanaon, Tagalog, or Waray-Waray.

[9] While the majority of speakers of this language refer to it as either “Ilonggo” or “Hilonggo” (or derivations such as “Inilonggo” or “Hinilonggo”), a source at the Department of Education notes that DepEd uses Hiligaynon as the term for the language, and Ilonggo as the term for the people.

The phonology of Ilonggo/Hiligaynon is identical to that of Tagalog, with only three phonemic vowels in the native stratum, to which even Spanish loans have been adapted. Ilonggo/Hiligaynon is also orthographically identical to Tagalog, including the use of foreign letters <c, f, j, ñ, qu, v, z> and the vowel graphemes <e> and <o>, which do not represent distinct phonemes in Ilonggo/Hiligaynon, in local place names and personal names. Some examples include Bacolod /bakulud/,` Cadiz /kadis/, Capiz /kapis/, Iloilo /ʔiluʔilu/, Murcia /murʃa/, Negros /nigrus/ ~ /negros/, Valladolid /balyadulid/, and Victorias /bikturyas/. As such, these letters are part of the orthography that a person must learn to become fully literate.

The orthographical issues noted for Tagalog and Cebuano also apply to Ilonggo/Hiligaynon, but as with Cebuano, the post-consonantal glottal stop is consistently written with a hyphen (e.g., *bag-o* 'new,' *bug-at* 'heavy,' *gab-i* 'night,' *pus-on* 'lower abdomen'). Similar to Tagalog and Cebuano, there is virtually no phonotactic or morphophonemic alternation in Ilonggo/Hiligaynon other than the previously mentioned *maN-/naN-/paN-* prefixes. The morphological complexity of the language is intermediate between that of Tagalog and Cebuano, with all three languages having similar systems of focus and mode. The tense/aspect system of modern spoken Ilonggo/Hiligaynon is somewhat inconsistent, however, in that a future tense is marked in the Actor Focus, but not generally used in any of the non-Actor Focuses. Similar to Cebuano, Ilonggo/Hiligaynon has both a morphologically distinct imperative form and a single subjunctive form. However, unlike Tagalog and Cebuano, the Actor Focus <*um*> and *mag-* paradigms have merged in Ilonggo/Hiligaynon, thereby reducing the number of basic Actor Focus paradigms from three (i.e., <*um*>*, mag-*, and *maN-*) to two (i.e., *mag-* and *maN-*).

Two Bible translations and numerous prayer books are available in the Ilonggo/Hiligaynon language, as is the Catholic Mass. One weekly magazine, *Hiligaynon*, is available, as are a small number of short (eight-page) newspapers, sold primarily in Iloilo City and Bacolod City. Although Ilonggo/Hiligaynon is used on the radio (primarily AM stations) and a small number of television news broadcasts, no non-religious books or professional quality dictionaries are commercially available in the language. The only Ilonggo/Hiligaynon music being produced appears to be limited to availability on the Internet..

2.4 Waray-Waray or Waray

The Waray-Waray (or simply "Waray") language is spoken by approximately 2.5 million native speakers throughout most of Samar Island, the northern and eastern portions of Leyte Island, and the eastern portion of Biliran Island. The dialect of the regional capital, Tacloban City, is the variety used in official written materials, including those used in the public school system for teaching and learning. However, a number of quite diverse, and generally more conservative, dialects are found throughout Samar Island, and significant dialectal variation is also found in portions of Leyte and Biliran Islands.

The phonology of the Tacloban dialect of Waray-Waray is identical to that of Tagalog (**Table 2**), but the dialects of eastern and northeastern Samar retain a fourth vowel, the "schwa" /ə/,[10] and therefore have the phoneme inventory listed in **Table 4**. Some Waray-Waray dialects also have a phonemic contrast between the flap <r> and the trill <rr>, the latter of which is only found in Spanish loan words.

[10] To be technically correct, this fourth vowel is a high central tense vowel /ɨ/, as mentioned in Section 2.2 of this report.

Table 4. The Phoneme System of Some Waray-Waray and Northern Samarenyo Dialects

	Consonants				Vowels			
	Bilabial	**Dental, Palatal**	**Velar**	**Glottal**		**Front**	**Central**	**Back**
Stop	p	t	k	ʔ	High	i	(ə)	u
	b	d	g		Mid			
Nasal	m	n	ŋ		Low		a	
Lateral		l						
Flap		r						
Trill		(rr)						
Fricative		s		h				
Glide	w	y						

The orthography of Waray-Waray is identical to that of Cebuano and Ilonggo, and is virtually the same as Tagalog but with the more consistent writing of postconsonantal glottal stop with a hyphen (e.g., *bag-o* /bagʔu/ 'new', *gab-i* /gabʔi/ 'night', *matab-ang* /matabʔaŋ/ 'lacking salt, lacking flavor'). The issues previously mentioned regarding the lack of representation of word stress and word-final glottal stop, on the one hand, and the overspecification of the three-vowel system, on the other hand, also apply to Waray-Waray. Note also similar to the situation with the schwa-retaining dialects of Cebuano, the schwa is not written as a distinct vowel in the schwa-retaining dialects of Waray-Waray and Northern Samarenyo, and is instead merged in the orthography with the phoneme /u/, written as <u> and <o> depending on the position in a word.

As with Tagalog, Cebuano, and Ilonggo, written Waray-Waray by default contains extra letters and digraphs that are not strictly necessary in spelling native words (<c, f, j, ñ, q, v, x, z, ch, ll, e, o>) that appear in widely recognized proper names. Some examples of proper names include Allen, Arteche, Calbayog, Can-avid, Catbalogan, General MacArthur, Giporlos, Guiuan, Jaro, Jiabong, La Paz, Lavezares, Leyte, Llorente, Mercedes, Pambujan, Quinapondan Salcedo, San Jorge, San Juanico, San Miguel, San Roque, Santa Fe, Tacloban, Taft, and Villareal. As such, similar to all other major Philippine languages, written Waray-Waray always includes both letters that are not necessary for writing native Waray-Waray words, and letters which, in foreign languages, represent sounds that are not found in Waray-Waray.

Similar to Cebuano, Ilonggo, and Tagalog, there is little phonotactic or morphophonemic alternation Waray-Waray, other than the previously mentioned assimilation involving the prefixes *maN-/naN-/paN-* (see Section 2.1 of this report). Although having the same system of focus and mode found in Tagalog, Cebuano, and Ilonggo, Waray-Waray is considerably more complex morphologically than the three previously mentioned languages, both in terms of the maximum number of affixes that can usually appear on a root word and in terms of having both two morphologically distinct imperative forms and two subjunctive forms (a past subjunctive and a future subjunctive). Marking of plurality on Waray-Waray verbs is also much more regular and common than in Cebuano, Ilonggo, or Tagalog. Also, similar to Tagalog, Waray-Waray marks a four-way contrast in its tense/aspect system. The morphology of Waray-Waray is highly regular, as is the case with Cebuano, Ilonggo, and Tagalog.

While the Bible and various prayer books are available in Waray-Waray, as is the Catholic Mass, there is considerably less non-religious print literature available for Waray-Waray than

for Tagalog, Cebuano, or Ilonggo. No books, magazines or newspapers are available in the Waray language, so its usage in the media is largely limited to AM radio and a small amount of television news broadcasts.

2.5 Central Bikol (Also Known as Bikol Naga)

Commonly misunderstood as referring to a single language, the name "Bikol" actually refers most properly to an entire subgroup of languages, although it is also used by locals in Sorsogon Province in the Bikol Region[11] to refer to two languages, Northern Sorsoganon and Southern Sorsoganon, which belong instead to the Bisayan subgroup of languages yet are called "Bikol" by their speakers by virtue of being spoken in the Bikol Region. The language most commonly referred to as "Bikol" without any qualifier is the Central Bikol language, which is most widely known locally as "Bikol Naga" because it is the language of Naga City, which is has long been considered to be the cultural and religious center of the Bikol Region. However, it is important to note that the Central Bikol language is also spoken in various closely related dialects along the northern coast of most of the Bikol Region, from Daet in the west to Prieto Diaz in the east, and in specific certain portions of the southern coast that are mostly contiguous with these areas. This Central Bikol language, in the dialect of Naga City and surrounding areas, is the Bikol that is used in MTB-MLE instruction in the Bikol Region of the Philippines, and which will therefore be the primary concern in this section of the report. However, for the sake of clarity, the name "Central Bikol" is used in this section instead of the ambiguous term "Bikol."[12]

The Central Bikol language has the same phoneme system as Tagalog (**Table 2**), except for the Southern Catanduanes dialect (**Table 5**), which has an additional consonant: an interdental approximant /ð̩/. This additional consonant, which is pronounced like an "L" while placing the tip of one's tongue between the teeth, originated as an allophone of earlier /l/ and /r/ in specific environments, but it has become a distinct phoneme because of the subsequent borrowing of words with /l/ and /r/ in those same environments. Although this sound was not traditionally distinguished in writing, recent materials being developed for MTB-MLE instruction have used an "L" with a diacritic above it, <Ĺ>, to write this interdental approximant.

Table 5. The Phoneme System of the Southern Catanduanes Dialect of Central Bikol

	Consonants					Vowels			
	Bilabial	**Inter-dental**	**Dental, Palatal**	**Velar**	**Glottal**		**Front**	**Central**	**Back**
Stop	p		t	k	ʔ	High	i		u
	b		d	g		Mid			
Nasal	m		n	ŋ		Low		a	
Lateral			l						
Flap			r						
Approximant		ð̩							
Fricative			s		h				
Glide	w		y						

[11] Note that although some writers use the name "Bikolano" for the language, this term actually only correctly refers to the people of the Bikol Region and not to the languages, which are only properly referred to as "Bikol".
[12] Note that the official government spelling of the geographical region is "Bicol", although many Bikolano authors use the spelling "Bikol", as is also used in this section.

Other than the additional consonant in the Southern Catanduanes dialect of Central Bikol, the only other phonology-related peculiarity of Central Bikol vis-à-vis the four previously discussed MTB-MLE languages is in the behavior of the glottal stop in consonant clusters. Although the glottal stop can only occur post-consonantally and never pre-consonantally in the four previously discussed languages outside of reduplicated monosyllables (e.g., Ilonggo *ba-bâ* /baʔbaʔ/ 'mouth'), the opposite is true of the Naga and Legaspi dialects of Central Bikol. In the Naga and Legaspi dialects, the glottal stop can only occur preconsonantally, never post-consonantally (except in reduplicated monosyllables): *ba-go* /baʔgu/ 'new,' *mata-bang* /mataʔbaŋ/ 'lacking salt, lacking flavor,' and *pu-son* /puʔsun/ 'lower abdomen.' However, in the Central Bikol dialects spoken in the Camarines Norte Province, the Partido District of Camarines Sur Province, southern Catanduanes, and the northern coast of Sorsogon Province (plus Magallanes town), the opposite is true (e.g., *bag-o* /bagʔu/ 'new,' *matab-ang* /matabʔaŋ/ 'lacking salt, lacking flavor,' and *pus-on* /pusʔun/ 'lower abdomen'). In the interest of making MTB-MLE materials dialect neutral, this issue appears to have been handled in these materials by not indicating the glottal stop clusters at all, thus <*bago*>, <*matabang*>, <*puson*>, and others. However, this convention then creates the additional problem of the lack of any orthographical distinction between single word-medial consonants, on the one hand, and clusters of a glottal stop plus another consonant, on the other.

As with the other MTB-MLE languages previously discussed, the additional letters and digraphs <c, f, j, ñ, q, v, x, z, ll, gu, qu, e, o> frequently occur in Central Bikol writing, in spite of not being necessary in writing native vocabulary, because of the abundance of proper names of foreign origin and their long history of use in spelling names of native origin. Some examples of these proper names include Camarines Norte, Camarines Sur, Catanduanes, Daet, Garchitorena, Legazpi (frequently also "Legaspi"), Magallanes, Masbate, Ocampo, Prieto Diaz, San Fernando, Sorsogon, and Villafuerte. Likewise, Central Bikol has a three-vowel system (/a, i, u/) in spite of using all five Spanish/English vowel graphemes (<a, e, i, o, u>) in spelling, a point of considerable inconsistency in the various orthographies that have been used in writing the language. Finally, similar to Tagalog and the other previously discussed languages, the Central Bikol orthography similarly lacks any marking of word stress and word-final glottal stop, both of which are phonemic.

Morphologically, Central Bikol is of similar complexity to Waray-Waray and is significantly more complex than Ilonggo and Tagalog. However, most of the morphology is highly regular, with the only significant morphophonemic alternations involving the previously mentioned prefixes *maN-*, *naN-*, *paN-*, and *pinaN-*. Similar to Ilonggo, the Actor Focus distinction between <um> and *mag-* paradigms has been reduced, with only the *mag-* paradigm remaining, although remnants of the <*um*> paradigm still appear as alternate forms. Otherwise, the Central Bikol verb system retains the full systems of focus and mode, and, similar to Waray-Waray, plurality is marked on Central Bikol verbs far more frequently than in Cebuano, Ilonggo, or Tagalog. The tense/aspect system of Central Bikol is similar in complexity to Tagalog and Waray-Waray, and a single imperative form remains in modern Central Bikol, although no subjunctive form exists in the language.

Similar to the other languages traditionally considered to be major languages, Central Bikol has multiple Bible translations, the Catholic Mass, and numerous (mostly short) prayer books. No commercially produced (non-religious) books or magazines are currently available in the Central Bikol language, and although a limited number of newspapers were available in the cities of Naga and Legaspi as recent as 20 years ago, none appear to still be in publication. Only a single dictionary of professional quality (Mintz & Britanico, 1986) has been published in the Philippines for Central Bikol, but it has long since been out of print; however, a later revision (Mintz, 2004) has been published. Regarding television, Central Bikol is generally limited to a small amount of news broadcasts, whereas usage on the radio is found primarily on AM stations. No modern Central Bikol music is known to currently be being produced, with the most recent being that of the Virac-based band Isla, which as

recently as 2000, was recording and releasing music in both Bikol Naga and Southern Catanduanes Bikol. Orthography guidebooks have been developed over the past decade by the provincial government of Albay Province and by Ateneo de Naga University in Naga City.

2.6 Ilokano

Besides being the native language of approximately six and a half million inhabitants of northern Luzon, Ilokano is also a lingua franca spoken and understood to varying degrees by dozens of other ethnolinguistic groups in surrounding areas. A primary dialect split in Ilokano can be identified between the more historically conservative southern dialects that retain the vowel /ə/ (the “schwa”) and the northern, more urban and more prestigious dialects, in which the schwa is absent, having merged with /e/.

The phoneme inventory of the northern dialect of Ilokano (**Table 6**) is generally similar to Tagalog, except for the addition of a trilled /rr/ in Spanish borrowings, at least for some speakers, and the vowel /e/ (as a reflex of Proto-Philippines *ə) as a fully phonemic vowel in native vocabulary. In the southern dialect (**Table 7**), the schwa /ə/ is also a fully phonemic vowel. The vowels /e/ and /o/ are also generally considered phonemic in Spanish borrowings (and likely in recent English borrowings), although it is unclear whether this is the case for all speakers or is primarily a characteristic of highly literate speakers only. Written Ilokano follows the northern dialect, in which the earlier /ə/ has completely merged with /e/, meaning that two distinct vowels in the southern dialect, /e/ and /ə/, are both written as “e” and pronounced as /e/ in the northern dialect.

Table 6. The Phoneme System of the Northern Dialect of Ilokano

	Consonants				Vowels			
	Bilabial	**Dental, Palatal**	**Velar**	**Glottal**		**Front**	**Central**	**Back**
Stop	p	t	k	ʔ	High	i		u
	b	d	g		Mid	e		o
Nasal	m	n	ŋ		Low		a	
Lateral		l						
Flap		r						
Trill		(rr)						
Fricative		s		h				
Glide	w	y						

Table 7. The Phoneme System of the Southern Dialect of Ilokano

	Consonants				Vowels			
	Bilabial	**Dental, Palatal**	**Velar**	**Glottal**		**Front**	**Central**	**Back**
Stop	p	t	k	ʔ	High	i	ə	u
	b	d	g		Mid	e		o
Nasal	m	n	ŋ		Low		a	
Lateral		l						
Flap		r						
Trill		(rr)						
Fricative		s		h				
Glide	w	y						

Beyond the phoneme inventory itself, however, a major difference between Ilokano and most of the other MTB-MLE languages is that Ilokano allows geminates (sequences of two identical consonants) such as *kalluto* /kal.lu.tu/ ‘just cooked,’ *lengnga* /leŋ.ŋa/ ~ /ləŋ.ŋa/ ‘sesame,’ and *ubbing* /ʔub.biŋ/ ‘children’ (Rubino, 2000). All consonants may occur as geminates in Ilokano, except for the glottal stop. Likewise, unlike in the other MTB-MLE languages, the glottal stop does not occur word-finally. Although the glottal stop also does not occur preconsonantally in Ilokano, it does occur post-consonantally, in which it is written with a hyphen, as in other major Philippine languages (e.g., *salun-at* /sa.lun.ʔat/ ‘health’; Rubino, 2000). Similar to the other MTB-MLE languages previously discussed, there are a few restrictions regarding word-medial consonant clusters, of which more than 100 pairings are found in native Ilokano words.

Orthographically, Ilokano differs from the previously discussed languages in writing short personal and demonstrative pronouns as part of the preceding word, by adding an extra syllable or two at the end of the written word (contributing to longer average word length). However, it should be noted that this is only a difference orthographically and that short pronouns do actually also behave this way grammatically (i.e., as enclitics) in other Philippine and Philippine-type languages. Overall, however, there are no symbols used in writing that are unfamiliar to speakers and readers of Tagalog, although non-native speakers of Ilokano are usually unable to properly pronounce the geminate consonant clusters.

As with the other MTB-MLE languages, a number of foreign letters are commonly used in writing Ilokano because of the presence of proper names of foreign origin and the history of writing even local place names in the Spanish orthography. Some examples of the names include Bangui, Cabarroguis, Cauayan, Muñoz, Quirino, Roxas, San Fernando, Tuguegarao, and Vigan. As such, even though these additional letters are not strictly necessary in spelling native vocabulary, there are no literate speakers of Ilokano who are unfamiliar with these letters, their meaning, and how to read and write them.

Although word-final glottal stops are not allowed in Ilokano, the lack of an orthographic indication of word stress is still a problem in written Ilokano as in the other languages discussed in this report, meaning that some written words taken out of context could be pronounced in two different ways.

The verbal morphology of Ilokano is of similar complexity to Tagalog, with a fully productive system of focus and mode, but a highly regular affix system nevertheless. No morphologically distinct imperative or subjunctive forms appear in the language, but Ilokano does have a five-way contrast in its tense/aspect system.

Similar to the other major languages spoken by Christian populations, Ilokano has more than one Bible translation, numerous prayer books, and Catholic Mass services. Also, similar to Cebuano, Ilonggo, and Tagalog, Ilokano has a weekly magazine, *Bannawag*, and a handful of small local newspapers. There is also the Ilokano Writers Association of the Philippines (*Gunglo dagiti Mannurat nga Ilokano iti Filipinas* [GUMIL]). Ilokano is also used on the radio and on television news broadcasts. Finally, although a handful of dictionaries have been published in recent decades, those of higher quality (e.g., Rubino, 2000) are not generally available in the Philippines.

2.7 Kapampangan

The Kapampangan language is spoken by approximately two million native speakers primarily in Pampanga Province, in the southern portion of neighboring Tarlac Province and in some adjacent towns in neighboring provinces. Little is known about the dialectal variation within the language itself, but its area appears to be receding as younger generations in many areas shift to Tagalog.[13]

As shown in **Table 8**, the phoneme inventory of Kapampangan is similar to that of Tagalog, except that /e/ and /o/ are fully phonemic in the native stratum, as monophthongizations of earlier diphthongs /ay/ and /aw/.

Table 8. The Phoneme System of Kapampangan

	Consonants				Vowels			
	Bilabial	**Dental, Palatal**	**Velar**	**Glottal**		**Front**	**Central**	**Back**
Stop	p	t	k	ʔ	High	i		u
	b	d	g		Mid	e		o
Nasal	m	n	ŋ		Low		a	
Lateral		l						
Flap		r						
Fricative		s						
Glide	w	y						

Compared with other MTB-MLE languages, there is a considerably larger amount of phonotactic alternation in Kapampangan, primarily of two types. In the first type, sequences of /a/ + /u/ become /o/, and /a/ + /i/ become /e/; in the second type, in the reverse direction, /e/ reverts to its underlying source /ay/, and /o/ reverts to its source /aw/ when suffixed with *-an*.

Orthographically, as with the other languages previously discussed, a number of letters not necessary in spelling native vocabulary are nevertheless commonly used in Kapampangan writing because of their presence in proper names of foreign origin and in local place names that were traditionally spelled according to Spanish orthographical rules. Some examples of these names include Concepcion, Dela Paz, Floridablanca, Mexico, Quebiauan, San Fernando, San Jose, and Tarlac. As such, no literate native speaker of Kapampangan would be unfamiliar with these letters. Similarly, place names such as “Anao” and “Magao” retain their orthographical “ao” diphthongs, even though the earlier /aw/ diphthong that they represented has long since been monophthongized to /o/ (thus, Anao /ano/ and Magao /mago/).

[13] A DepEd source claims that there are anywhere from two to five dialects of Kapampangan, as well as a Pampanga dialect of Tagalog/Filipino.

As in the orthographies of Tagalog and other MTB-MLE languages, the Kapampangan orthography indicates neither word stress, which is phonemic, nor word-final glottal stops.[14] Morphologically, one thing that distinguishes Kapampangan from most other Philippine languages is the high degree of morphosyntactic complexity in verb forms conjugated with reflexes of earlier infixes <*um*> and <*in*>. The result is that the conjugations of a large percentage of the most common verbs are irregular in their past, present, and future conjugations. Furthermore, because of historical processes, the suffixes marking both the Object Focus and the Location Focus have merged as *-an* in the infinitive form, although the non-infinitive conjugations are still differentiated.[15] Otherwise, the degree of complexity in the Kapampangan verb system is similar to that of Tagalog, with fully productive systems of focus and mode, but with no morphologically distinct imperative or subjunctive forms.

Similar to the other major Philippine languages spoken by Christian populations, there are multiple Kapampangan Bible translations, numerous prayer books, and the Catholic Mass in Kapampangan. Dictionaries of the language also exist, both locally produced and international. To date there are a meager availability of non-religious books, magazines, or newspapers, nor are there any modern forms of music widely available in the language. Elsewhere in the media, Kapampangan usage on television is limited to the usual television news broadcasts and radio programming primarily on AM and FM radio. Finally, it is noteworthy that there is a growing number of learning institutions and associations committed to the development, promotion, and preservation of the Kapampangan language and culture. These include the Juan D. Nepomuceno Center for Kapampangan Studies at Holy Angel University, the Institute of Kapampangan Studies at Angeles University Foundation, Akademyang Kapampangan, and Katipunan ding Talasalisksik at Talaturung Kapampangan among others.

2.8 Maranao (Also Known as Meranaw and Mëranaw)

With just under a million native speakers, the Maranao language (“Meranaw” /məranaw/ in the language itself, or <Mëranaw> in the orthography currently used in Department of Education MTB-MLE teaching materials), spoken by a Muslim population in one of the few remaining culturally homogenous areas of the large southern Philippine island of Mindanao. Maranao is highly unique in a number of ways, such as its one-of-a-kind phonology and its history of local publishing by native scholars.

Phonologically, Maranao is unique among Philippine languages because of its series of “heavy” consonants (/p’, t’, k’, s’/), which are generally spelled as <ph, th, kh, z>, respectively. However, in Maranao, <sh> is often used in place of <z>, which corresponds to the clusters /bp, dt, gk, ds/ in Maguindanaon and Iranun and manifest themselves most saliently in the tensing and raising of the following vowel (Lobel & Riwarung, 2009 and 2011). The result is a register-like system in which one series of consonants (the heavy consonants plus the fricative /h/, which only occurs in Malay loan words and voiced stops, albeit only allophonically) is followed by one set of vowel allophones, whereas the other consonants are followed by a different set of vowel allophones. **Table 9** lists the phoneme inventory of Maranao, and **Table 10** lists the two vocal registers in the Maranao language.

[14] A DepEd source suggests that diacritics can be used to disambiguate meanings; for example, the word 'masakit' has ambiguous meanings, but with diacritical marks, meanings are well defined: *'masakit'* (difficult), *masákit* (sick), and *másakit* ‘hard or difficult’. However, this is not part of the official orthography of MTB-MLE materials.

[15] As with other Philippine and Philippine-type languages (Blust, 2013), the suffix is dropped in one or more of the non-infinitive Object Focus conjugations.

Table 9. The Phoneme System of Maranao

	Consonants				Vowels			
	Bilabial	**Dental, Palatal**	**Velar**	**Glottal**		**Front**	**Central**	**Back**
Stop	p	t	k	ʔ	High	i	ə	u
Heavy	p’	t’	k’		Mid			
Voiced	b	d	g		Low		a	
Nasal	m	n	ŋ					
Lateral		l						
Flap		r						
Fricative		s		h				
Glide	w	y						

Table 10. Two Series of Vowels by the Preceding Consonant

	Vowel Phoneme			
	/a/	/ə/	/i/	/u/
1st register—Non-raising (/k, p, t, s, m, n, ŋ, r, w, y/)	[a]	[ə][16]	[ɪ]	[o]
1st register—Non-raising, but transparent (/l, ʔ/)	[a]	[ə]	[ɪ]	[o]
2nd register—Obligatory raising (/k’, p’, t’, s’, h/)	[ɤ]	[ɨ]	[i]	[u]
2nd register—Allophonic raising (/b, d, g/)	[ɤ]	[ɨ]	[i]	[u]

This phoneme system has complicated both the analysis of the language (particularly by non-Maranaos) and outside attempts to standardize its orthography for pedagogical purposes. For the sake of clarity, **Table 11** lists all of the consonant phonemes of Maranao and the letters traditionally used by native Maranao authors over the past 60 years to represent them. **Table 12** lists the only four vowels that are phonemic in the Maranao language and the graphemes used by native Maranao authors to represent them.

Table 11. Graphemes Representing the Consonant Phonemes of Maranao

Consonant phoneme	b	d	g	h	k	k’	l	m	n	ŋ	p	p’	r	s	s’	t	t’	w	y	ʔ
Grapheme	b	d	g	h	k	kh	l	m	n	ng	p	ph	r	s	Z (or sh)	t	th	w	y	—

Table 12. Graphemes Representing the Vowel Phonemes of Maranao

Vowel phoneme	a	ə	i	u
Grapheme	a	e (or u)	i	o

[16] In **Table 10**, the symbol “ə” marks a true schwa, which contrasts allophonically with the high, unrounded tense central vowel /ɨ/, as opposed to elsewhere in this report, in which the symbol “ə” is used as a more easily readable alternative representation of the high, unrounded tense central vowel [ɨ].

As with other Philippine languages, no Maranao orthography represents the word-final glottal stop, in spite of its phonemic status, and many orthographies that have been used for Maranao over-specify the vowel system by marking a false distinction between “o” and “u” (although in some native orthographies, the letter “u” is used instead of “e” to represent the schwa, which is written as <ë> in current Department of Education pedagogical materials).

Beyond the phoneme inventory of the language, a number of historical processes have reduced the number of permissible consonant clusters in the Maranao language from the 100-plus in most other major Philippine languages, to only seven, all of which consist of a stop or /s/ preceded by a homorganic nasal: /mp, mb, nt, nd, ns, ŋk, ŋg/. Note that this amount is fewer than even the 11 permissible clusters in the closely related languages of Maguindanaon and Iranun, and in general, the reduction of consonant clusters in Maranao and Maguindanaon should hypothetically reduce the difficulty for students learning to read these languages. The same historical processes that reduced the number of permissible consonant clusters and that created the “heavy” consonants of Maranao from the earlier clusters /bp, dt, gk, ds/ also operate at the boundaries of prefixes and root words. This makes a proper understanding of Maranao phonology an integral part of any analysis of the verbal morphology of the language, and has thwarted all previous attempts to analyze the verb system and orthography of the Maranao language by non-Maranao researchers. Morphophonemic complexity is therefore highest in Maranao among Philippine languages as a whole, and only a small number of possible root word forms are not intimately intertwined with the complex morphophonemics of the Maranao language. Beyond the morphophonemic issues, Maranao has a fully productive system of focus and mode, as well as numerous imperative and subjunctive forms that further complicate the verbal conjugations.

As previously mentioned, the Maranaos are the only ethnolinguistic group in the Philippines known to produce and publish a significant amount of native-language literature. The best known example is the Maranao version of the Islamic Qur’an translated[17] by Shiek Abdul Aziz Guroalim Saromantang in the 1990s. This version was first made available on the local level in paperback installments, but it was later printed as a single, hardbound volume that is still widely available. However, a large amount of other lengthy Maranao books has been published and made commercially available in the Islamic City of Marawi over the past few decades. Some of these books are religious in nature, others are simply textbooks for studying Arabic and/or English, and others were studies of various aspects of Maranao history and culture. A substantial amount of traditional and modern Maranao music is also constantly being produced, as are DVDs of various cultural events. In contrast, no newspapers or magazines are published in the Maranao language.

2.9 Maguindanaon (Also Known as Magindanawn)

Maguindanaon is the language of another Muslim population closely related to Maranao. Maguindanaon (or Magindanawn in Department of Education MTB-MLE teaching materials) is spoken in various dialects by just over one million native speakers in several portions of the western half of Mindanao. Maguindanaon is spoken primarily in the areas almost immediately south of the Maranao-speaking region (albeit with Iranun located in the middle), but also in specific portions of the Zamboanga Peninsula.

The phoneme inventory of Maguindanaon (**Table 13**) is similar to that of Tagalog, save for the absence of /h/ and the addition of the vowel /ə/.

[17] It is a widespread belief in Islam that the only true Qur’an is one written in original Arabic and that anything else is not a “translation” but simply an “interpretation” to help non-Arabic-speaking Muslims understand the original Arabic. However, because both Maranao and Saudi Arabian materials refer to the Maranao Qur’an with words such as “translation” and “translated,” those terms are also used in this report for the sake of simplicity.

Table 13. The Phoneme System of Maguindanaon

	Consonants				Vowels			
	Bilabial	**Dental, Palatal**	**Velar**	**Glottal**		**Front**	**Central**	**Back**
Stop	p	t	k	ʔ	High	i	ə	u
	b	d	g		Mid			
Nasal	m	n	ŋ		Low		a	
Lateral		l						
Flap		r						
Fricative		s						
Glide	w	y						

Similar to many of the languages of western Mindanao and adjacent portions of northern Sulawesi in Indonesia, and northern Borneo in Indonesia, Maguindanaon has a drastically reduced number of permissible consonant clusters, with only 11 such clusters permitted. The 11 clusters are as follows: /bp/, /dt/, /ds/, /gk/, /mb/, /mp/, /nd/, /nt/, /ns/, /ŋk/, and /ŋg/. The four voiced-voiceless pairs, /bp, dt, ds, gk/, correspond to the “heavy” consonants of Maranao, but are much more transparent in pronunciation because they actually are clusters of /b/, /d/, or /g/ followed by /p/, /t/, /s/, or /k/. Similar to Maranao, however, the impermissibility of other consonant clusters means that the reflexes of widespread prefixes such as *mag-, nag-, pag-, maN-, naN-,* and *paN-* are often barely recognizable, and dependent on the form of the root word to which they are attached.

Orthographically, Maguindanaon is much more transparent than its close relative Maranao. Similar to any other Philippine language, the Maguindanaon orthography lacks marking of word-final glottal stops, and a number of letters of foreign origin appear in writing, especially in place names and personal names of Arabic/Islamic origin. Some examples of these names include Cotabato, Datu Abdullah Sangki, General Santos, Maguindanaon, Rajah Buayan, Shariff Aguak, and Zamboanga.

The verb system of Maguindanaon is, in many ways, of a similar complexity to that of Ilonggo, Tagalog, or Waray-Waray, but the previously mentioned historical processes cause a considerable amount of morphophonemic alternations in the reflexes of the extremely common prefixes *mag-, nag-, pag-,* and *pinag-*. Still, compared with Maranao, these alternations in Maguindanaon are much more transparent. Similar to the other MTB-MLE languages, Maguindanaon retains a full system of focus and mode, and also has imperative and subjunctive forms.

A Maguindanaon translation of the Qur’an was completed a decade ago, but it is not as widely available as the Maranao equivalent, and before its completion, no Maguindanaon written materials were available. There are also no commercially available books, magazines, or newspapers in the Maguindanaon language, although a dictionary of the language was produced by Cotabato’s Notre Dame University (Sullivan, 1986). Some locally produced Maguindanaon music is sold in areas with considerable Maguindanaon populations, and additional material is available on the Internet.

2.10 Tausug (Also Known as Bahasa Sūg)

Spoken by nearly one million Tausug in the southwestern Philippines and northern Borneo, the Tausug language is also the lingua franca for most of the Sulu Archipelago and some coastal areas in northern and eastern Sabah in Malaysia. It is important to note in the

Tausug language itself, *Bahasa Sūg* refers to the language and *Tau Sūg* refers to the people of the Sulu Archipelago.

In terms of phoneme inventory (**Table 14**), Tausug is similar to Tagalog and to the Bisayan languages (e.g., Cebuano, Ilonggo, and Waray-Waray) to which it is closely related. The only difference in phoneme inventory is the addition of the phoneme /ʤ/, which is spelled <j> as is common in English and is not simply an allophone of /dy/ sequences as in many other Philippine languages such as Cebuano and Tagalog. The historical loss of intervocalic /l/ has resulted in long vowels that are independent of word stress, similar to those found in many of the L-dropping dialects of Cebuano. Beyond the phoneme inventory, it is noteworthy that all consonants except /ʔ/ and /h/ may occur as geminates.

Table 14. The Phoneme System of Tausug

	Consonants				Vowels			
	Bilabial	**Dental, Palatal**	**Velar**	**Glottal**		**Front**	**Central**	**Back**
Stop	p	t	k	ʔ	High	i		u
	b	d	g		Mid			
Nasal	m	n	ŋ		Low		a	
Lateral		l						
Flap		r						
Fricative		s		h				
Affricate		ʤ						
Glide	w	y						

The orthography of Tausug is similar to those of other Philippine languages in lacking representation of the word-final glottal stop. Long vowels are written with a macron above the vowel (e.g., <ā, ī, ū>), which distinguish them from vowel-glottal-vowel clusters which, as in other Philippine languages, are written as sequences of two consecutive vowels (e.g., <uu> representing /uʔu/, <aa> representing /aʔa/, and <ii> representing /iʔi/). Geminate consonants are written as sequences of two consecutive consonants such as <dd>, <ll>, <mm>, and others. As with other Philippine languages, several foreign letters and spelling conventions are found in written Tausug because of the historical usage of Spanish spelling conventions in spelling place names (e.g., Caluang, Jolo, Old Panimao, Talipao) and personal names of Arabic/Islamic origin.

Tausug, similar to the other MTB-MLE languages previously discussed, has a fully developed system of verbal focus and mode and a degree of complexity similar to languages such as Ilonggo, Tagalog, and Waray-Waray. In addition to the usual morphophonemic alternations in the prefixes *maN-, naN-,* and *paN-,* alternations also occur due to the dropping of the nasal consonants in the common Philippine verbal infixes <*um*> and <*in*> when affixed to root words of specific phonological characteristics in Tausug.

There are no known commercially available Tausug language books, newspapers, or magazines, although the use of the Tausug language is widespread on the Internet. There also does not appear to be any widely distributed religious materials in the Tausug language, although a Qur'an translation may be in progress, or may even have been completed in recent years. There is no information available about local usage of the Tausug language on television or the radio in the Sulu Archipelago.

2.11 Chavacano

The only non-Austronesian language native to the Philippines, and the only designated MTB-MLE language not genetically part of the Philippine subfamily, is Zamboangueño Chavacano (henceforth referred to as Chavacano[18]), with more than 300,000 native speakers. Chavacano is a Spanish-based creole in which the vast majority of vocabulary is Spanish in origin, yet absent are such Spanish grammatical features as the verb conjugations and the morphological agreement in the marking of gender and plurality. However, although its sentence structure is more Philippine-like than Spanish-like, the Philippine-type systems of focus, mode, and case are nowhere to be found. In summary, with its three- if not four-century time depth in the Philippines, Chavacano is like no other language in the Philippines, whether orthographically, morphologically, or grammatically.

Phonologically, for most speakers, Chavacano has a phoneme inventory similar to genetically Philippine languages such as Cebuano and Tagalog (**Table 15**; Steinkrüger, 2013). However, the palatal nasal /ɲ/ and the palatal lateral /ʎ/ appear to be phonemic in Chavacano, although no evidence has been presented that they actually contrast with the clusters /ny/ and /ly/, respectively. For at least some speakers, the vowels /e/ and /o/ contrast with /i/ and /u/, respectively, and several other consonants (/ʧ, ʤ, rr, x, f/) may similarly be phonemic for some portions of the Chavacano population, although it is unclear whether this is because of renewed (if not sustained) exposure to Spanish (and, more recently, English).

Table 15. The Phoneme System of Chavacano

	Consonants					Vowels			
	Bilabial	**Dental**	**Palatal**	**Velar**	**Glottal**		**Front**	**Central**	**Back**
Stop	p	t		k	ʔ	High	i		u
	b	d		g		Mid	e		o
Nasal	m	n	ɲ	ŋ		Low		a	
Lateral		l	ʎ						
Flap		r							
Trill		(rr)							
Fricative		s			h				
Affricate			ʧ						
			(ʤ)						
Glide	w		y						

Orthographically, although variation exists as in other Philippine languages, the most common conventions appear to be to generally follow the original Spanish spelling rules for words of Spanish origin (even in cases where the Chavacano word does not exist in that form in standard Spanish) and to follow Tagalog-like spelling rules for words originating from Tagalog or Bisayan languages. This mixture of two spelling systems and the complexity inherent in the Spanish-type spelling rules make the orthography of Chavacano arguably the most complex in the Philippines. **Tables 16 and 17** illustrate the consonant and vowel graphemes of Chavacano, respectively, and the phonemes that they represent.

[18] Another variety of Chavacano survives in Cavite Province just south of the Philippine capital of Manila. Several other now-extinct varieties also once existed elsewhere in the Philippines. However, only the Chavacano of Zamboanga City is currently relevant to MTB-MLE instruction in the Philippines.

Table 16. Graphemes Representing the Consonant Phonemes of Chavacano

Grapheme	b	c	ch	d	f	g	h	j	k	l	ll	m	n	ñ	ng	p	q	r	rr	s	t	v	w	x	y	z
Consonant phoneme	b	k, s	ʧ	d	p, (f)	g, x	—	h, (x)	k	l	ʎ	m	n	ɲ	ŋ	p	k	r	rr, hr	s	t	b	w	s, x	y	s

Table 17. Graphemes Representing the Vowel Phonemes of Chavacano

Grapheme	a	e	i	o	u
Vowel phoneme	a	e	i	o	u

As with other languages in the Philippines, word stress is not marked, even though it is phonemic, and there is no written indication of word-final glottal stops, which are only found in words of genetic Philippine origin. One key difference from the orthographies of other languages in the Philippines is that because of its use of the Spanish spelling rules, sequences of two non-identical vowels do not indicate the presence of an intervening glottal stop, but instead, a true diphthong (e.g., *cuando* /kwando/ ‘when,’ *despues* /despwes/ ‘after,’ *nuevo* /nwebo/ ‘new,’ *quien* /kiyen/ ‘who’).

As previously stated, Chavacano does not have a Philippine-type verbal system, in spite of its Philippine-type sentence structure. There are no focus affixes and no mode affixes, and tense is marked by proclitic particles, not affixes as in genetically Philippine languages. There are also no Romance-style verb conjugations for person, number, and tense.

Little print material is commercially available in the Chavacano language. A Bible translation is available, but apparently no other books, magazines, or newspapers are available, although the Zamboanga City Government has been proactive in promoting Chavacano, and the City Charter and other written declarations are printed in Chavacano (William Hall, personal communication, April 21, 2020).

3 Methodology for Developing Language Complexity Score

Based on the issues identified by Biber et al. (2011), Bulté & Housen (2014), Dahl (2004), and Miestamo (2003), and others and adapted to the goals of the current study, the study authors identified various objective factors of complexity that can be categorized as phonological, morphological, orthographical, or sociolinguistic. These factors are summarized in **Tables 19** and **20**.

In determining the phonological complexity of each language, both absolute and relative factors were considered. The number of vowels and consonants were counted, as were the number of vowels and consonants in each language that are not found in Tagalog, the national language of the Philippines. The total number of phonemes (vowels and consonants) were also counted for each language, as well as the total number of phonemes found in each language that are not present in Tagalog. Beyond the phoneme inventory, we counted the number of word-internal consonant clusters in the core vocabulary of each language (defined as the native stratum plus any long-standing loan words, for which centuries-old Spanish and Malay loan words would count, but not English borrowings, which had only entered the languages within the past generation or two). For most languages, well over 100 such clusters are permitted, but each language was only scored up to 100. The presence of various other aspects of the phonology were also considered, including the following:

- Whether word stress is phonemic
- Whether phonemic long vowels occur in the language independent of word stress

- Whether consonant gemination is allowed
- The maximum number of phonemes per syllable in the native vocabulary
- Whether voice register exists in the language
- The degree to which phonotactic alternations are found in the language.

Some of the orthographical factors included were the number of graphemes used in spelling the native vocabulary of the language that are not found in Tagalog, the degree of irregularity in the orthography, the number of non-transparent graphemes, and the number of digraphs found in the orthography.

Some of the morphological factors included were whether each language had an imperative mood and/or a subjunctive mood, the total number of verb moods in the language, the degree of morphological irregularity in the language, the degree of morphophonemic alternations, and the total degree of affix complexity, as determined by the total number of verbal affixes (and, in Ilokano, pronouns and demonstratives).

Some of the sociolinguistic factors considered were the availability of newspapers and/or magazines in the language and the commercial availability of non-religious books. Another sociolinguistic factor considered was because religious texts are often the only place where Philippine languages can be found in print, whether a Bible was available in the language for Christian ethnolinguistic groups, or a Qur'an translation, for Islamic ethnolinguistic groups.

The previously mentioned factors were scored individually, after which the totality of these complexity factors was considered as a whole. Languages were then scored on a scale of "1" to "3", with "1" indicating a lower degree of complexity, "2" indicating medium complexity, and "3" indicating a higher degree of complexity. The result is the categorization presented in **Table 18**.

Table 18. Language Complexity Categories

Complexity Group	Languages in Each Group
1	Tagalog
2	Bikol, Bahasa Sūg, Ilonggo, Sinugbuanong Binisaya, and Waray-Waray
3	Ilokano, Kapampangan, and Maguindanaon
4	Maranao

Glossary of Terms

Agglutination—When affixes or other morphology attach to root words.

Allophone—One of two or more distinct variations of a phoneme that has no effect on the meaning of the word.

Dialect—When a language has well-established variations in the form of different words (not just variation in pronunciation of words), but the variations remain mutually intelligible.

Digraph—two letters that represent one sound, as 'ph' [f].

Glottal stop—A consonant-like sound that results from an obstruction of air in the vocal tract.

Geminate clusters—Identical speech sounds pronounced one after the other.

Grapheme—The written symbols used to write the language.

Diphthong—A vowel sound that combines two vowel sounds in the same syllable (despite, at times, being written with only one grapheme), as in the British English word "go" [gəʊ].

Morpheme—The smallest unit of meaning in a word.

Morphology—The structure of words and parts of words, such as stems, roots, prefixes, and suffixes.

Morphophonemic alternation—When the combination of two morphemes results in a change in the sound produced by the morpheme (as in some English words that change when the plural or past tense suffixes (morpheme /s/) are added, such as with "*cats"* /s/, "*dogs"* /z/, and "*horses"* /ɬz/).

Monophthong—A vowel sound that remains constant from beginning to end, with no glide as a result of a change in position of articulation.

Phoneme—A unit of sound in a language.

Phonology—The system of sounds that make up a language.

Phonotactics—The rules governing the arrangement of allowable speech sounds within a given language.

References

Abadzi, H. (2012). *Developing cross-language metrics for reading fluency measurement: some issues and options.* (Global Partnership for Education Working Paper No. 6). Washington, DC: World Bank.

Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly, 45*(1), 5–35.

Borleffs, E., Maassen, B., Lyytinen, H., & Zwarts, F. (2019). Cracking the code: The impact of orthographic transparency and morphological-syllabic complexity on reading and developmental dyslexia. *Frontiers in Psychology*, *9*, 2534. doi.org/10.3389/fpsyg.2018.02534

Blust, R. A. (2013). *The Austronesian languages.* Canberra: Pacific Linguistics.

Brunette, T., Piper, B., Jordan, R., King, S., & Nabacwa, R. (2019). The impact of mother tongue reading instruction in twelve Ugandan languages and the role of language complexity, socioeconomic factors, and program implementation. *Comparative Education Review, 63*(4), 591–612. doi:10.1086/705426

Bulté, B., & Housen, A. (2014). Conceptualizing and measuring short-term changes in L2 writing complexity. *Journal of Second Language Writing, 26*, 42–65. [Note: See pages 45 and 46 of this current report for the definition of complexity.]

Coady, J. A., & Aslin, R. N. (2004). Young children's sensitivity to probabilistic phonotactics in the developing lexicon. *Journal of Experimental Child Psychology, 89*(3), 183–213. doi:10.1016/j.jecp.2004.07.004

Dahl, Ö. (2004). Chapter 3: Complexity, order, and structure. In *The growth and maintenance of linguistic complexity*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Dubeck, M. M., & Gove, A. (2015). The Early Grade Reading Assessment (EGRA): Its theoretical foundation, purpose, and limitations. *International Journal of Educational Development, 40*, 315–322. doi:10.1016/j.ijedudev.2014.11.004

Eberhard, D. M., Simons, G. F., & Fennig, C. D. (Eds.). (2020). *Ethnologue: Languages of the world.* 23rd edition. Dallas: SIL International.

Elbro C., & Arnbak, E. (1996). The role of morpheme recognition and morphological awareness in dyslexia. *Annals of Dyslexia, 46*, 209–240. doi:10.1007/BF02648177

English, L. J. (1986). *Tagalog-English dictionary*. Mandaluyong City: Congregation of the Most Holy Redeemer and National Book Store.

English, L. J. (1977). *English-Tagalog dictionary*. Mandaluyong City: Congregation of the Most Holy Redeemer and National Book Store.

Frost, R. (1998). Toward a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin*, *123*, 71–99.

Gonzalez, A. (2007). The Philippines. In Simpson, A. (Ed.), *Language and National Identity in Asia*, 360–373. Oxford: Oxford University Press.

Gonzalez, A. (1998). The language planning situation in the Philippines. *Journal of Multilingual and Multicultural Development, 19*(5&6), 487–525.

Lewis, M. P., Simons, G. F., & Fennig, C. D. (Eds.). (2013). *Ethnologue: Languages of the world.* 17th edition. Dallas, US: SIL International.

Lobel, J. W., & Riwarung, L. H. S. (2009). Maranao revisited: An overlooked consonant cluster and its implications for lexicography and grammar. *Oceanic Linguistics, 48*(2), 403–438.

Lobel, J. W., & Riwarung, L. H. S. (2011). Maranao: A preliminary phonological sketch with supporting audio. *Language Documentation and Conservation, 5*, 31–59.

McEachern, F. (2013). *Local languages and literacy in the Philippines: Implications for early grade reading instruction and assessment*. Prepared for EdData II Technical and Managerial Assistance, Task Number 17 Contract Number: AID-492-M-12-00001. Research Triangle Park, NC: RTI. Retrieved from https://shared.rti.org/content/local-languages-and-literacy-philippines-implications-early-grade-reading-instruction-and#

McFarland, C.D. (1994). Defining a Filipino lexicon. In Sibayan, B. & Newell, L. E. (Eds.), *LSP Special Monograph Issue 35* [Papers from the First Asia International Lexicography Conference] (pp. 241–254). Manila, Philippines: Linguistic Society of the Philippines.

Miestamo, M. (Ed.). (2003). Language complexity: Typology, contact, change. Amsterdam/Philadelphia: John Benjamins Publishing.

Milledge, S. V., & Blythe, H. I. (2019). The changing role of phonology in reading development. *Vision, 3*(2), 23. doi:10.3390/vision3020023

Mintz, M. W. (2004). *Bikol dictionary: Diksionariong Bikol*. Indonesian/Malay Texts and Resources.

Mintz, M. W, & Del Rosario Britanico, J. (1986). *Bikol-English dictionary: Diksionariong Bikol-Ingles.* Quezon City: New Day Publishers.

Quezon, M. L. (1937, December 30). Speech titled *Talumpati ng Kagalang-galang Manuel L. Quezon, Pangulo ng Pilipinas, sa paglikha ng pambansang wika*.

RTI International. (2016). *Early Grade Reading Assessment (EGRA) Toolkit. 2nd Edition*. Washington, DC: United States Agency for International Development. Retrieved from http://pdf.usaid.gov/pdf_docs/pa00m4tn.pdf

Rubino, C. R. G. (2000). *Ilocano dictionary and grammar: Ilocano-English, English-Ilocano*. PALI Language Texts: Philippines. University of Hawai'i Press.

Steinkrüger, P. O. (2013). Zamboanga Chabacano. In Michaelis, S. M., Maurer, P., Haspelmath, M., & Huber, M. (Eds.). *The survey of Pidgin and Creole languages, Volume 2: Portuguese-based, Spanish-based, and French-based languages.* Oxford: Oxford University Press.

Sullivan, R. E. (1986). *A Maguindanaon dictionary*. Cotabato: Notre Dame University Institute of Cotabato Cultures.

Storkel, H. L. (2001). Learning new words. *Journal of Speech, Language, and Hearing Research, 44*(6), 1321–1337. doi:10.1044/1092-4388(2001/103).

Wolff, J. U. (1972). *A dictionary of Cebuano Visayan*. Ithaca, NY: Cornell University.

Table 19. Phonological Measures of Complexity

	Language											
Complexity Measure	**CBik**	**CEB**	**CHA**	**HIL**	**ILK**	**KAP**	**MAR**	**MGD**	**TAG**	**TSG**	**WAR**	**Notes**
Total vowels	3	3	3 / 5	3	4	5	4	4	3 / 5	3	3	3 / 5 means three native vowels but adoption of two additional in loan words
Vowels: Non-Tagalog	0	0	0	0	1	2	1	1	0	0	0	
Total consonants (including [ʤ ʧ ɲ ʃ])	20	20	21	20	20	20	24	20	20	20	20	
Native phonemic consonants	16	16	21	16	16	16	20	16	16	17	16	
Rare/non-Tagalog consonants	0	0	5	0	0	0	4	0	0	1	0	
Total phonemes (including [ʤ ʧ ɲ ʃ])	23	23	26	23	24	25	28	24	25	23	23	
Rare/non-Tagalog phonemes	0	0	5	0	0	0	5	1	0	0	0	
Total consonant clusters	100	100	—	100	100	100	7	11	100	100	100	[0–100]
Contrastive/phonemic stress	1	1	1	1	1	1	0	0	1	0	1	[0 = no; 1 = yes]
Phonemic vowel length	1	0 / 1	0	1	0	0	0	0	0 / 1	1	1	[0 = no; 1 = yes]
Consonant gemination allowed	0	0	0	0	1	0	0	0	0	1	0	[0 = no; 1 = yes]
Phonological irregularity	0	0	0	0	0	0	0	0	0	0	0	[0 = low; 1 = high]
Phonotactic alternations	0	0	0	0	0	0	1	1	0	0	0	[0 = no; 1 = yes]
Syllable structure (maximum overall)	4	4	5	4	4	4	4	4	4	4	4	
Syllable structure (maximum native strata)	3	3	5	3	3	3	3	3	3	3	3	
Voice register	0	0	0	0	0	0	1	0	0	0	0	[0 = no; 1 = yes]

Note: CBik = Central Bikol; CEB = Cebuano; CHA = Chavacano; HIL = Ilonggo/Hiligaynon; ILK = Ilokano; KAP = Kapampangan; MAR = Maranao; MGD = Maguindanaon; TAG = Tagalog; TSG = Tausug; WAR = Waray.

Table 20. Orthographical, Morphological, and Sociolinguistic Measures of Complexity

		Language											
Category	**Complexity Measure**	**CBik**	**CEB**	**CHA**	**HIL**	**ILK**	**KAP**	**MAR**	**MGD**	**TAG**	**TSG**	**WAR**	**Notes**
Orthographical	Orthographical irregularity	0	1	2	0	0	0	2	1	0	1	0	[0 = low, 1 = mid, 2 = high]
	Non-Tagalog graphemes	0	0	11	0	0	0	4	0	0	1	0	
	Non-transparent graphemes	3	4	3	3	1	2	4	1	3	1	3	
	Digraphs	5	5	4	5	5	5	9	5	5	16	5	
Morphological	Verbal moods	2	3	1	2	1	1	3	3	1	2	3	[total number]
	Imperative	1	1	0	1	0	0	1	1	0	1	1	[0 = no, 1 = yes]
	Subjunctive	0	1	0	1	0	0	1	1	0	0	1	[0 = no, 1 = yes]
	Irregularity	0	0	0	0	0	1	0	0	0	0	0	[0=low, 1=high]
	Morphophonemic alternations	0	0	0	0	0	1	1	1	0	0	0	[0=low, 1=high]
	Affix complexity	1	0	0	0	1	2	2	1	1	1	2	[0 = low, 1 = mid, 2 = high]
Sociolingistic	Newspapers and/or magazines	0	1	0	1	0	0	0	0	1	0	0	[0 = no, 1 = yes]
	Bible/Qu'ran	1	1	1	1	1	1	1	1	1	0	1	[0 = no, 1 = yes]
	Non-religious books	0	0	0	0	0	0	1	0	1	0	0	[1 = yes, 0 = no]

[a] The Tausug Bible translation was not counted because virtually no Tausugs are Christian; therefore, it would be highly unlikely for them to use it. Maranao and Maguindanaon numbers are for the availability of Islam's Qur'an, not Christianity's Bible.

Note: CBik = Central Bikol; CEB = Cebuano; CHA = Chavacano; HIL = Ilonggo/Hiligaynon; ILK = Ilokano; KAP = Kapampangan; MAR = Maranao; MGD = Maguindanaon; TAG = Tagalog; TSG = Tausug; WAR = Waray.