The Languages of the Philippines: an overview

(draft 4/10/26 – comments welcome!)

Daniel Kaufman
City University of New York & Endangered Language Alliance
dkaufman@qc.cuny.edu

1. Introduction

There are roughly 177 indigenous languages of the Philippines (Eberhard et al. 2025), all of which belong to the Austronesian family and all but two of which are still spoken today. Given its location, the northern Philippines must have been the first stopping point of the Proto-Malayo-Polynesian (PMP) speaking seafarers who sailed southwards from Taiwan approximately 4,000 years ago. This was the initial step of one of the greatest migrations in human history, eventually leading to the present day distribution of Austronesian languages from Madagascar in the west to Easter Island in the east.

While the political boundaries of the modern Philippines might contain an arbitrary set of subgroups from a phylogenetic perspective, the languages contained within this region in fact share certain typological features which have given rise to the useful designation "Philippine-type language". This term is understood to refer to the Austronesian languages of the Philippines and those of a similar typological profile in neighboring areas of Taiwan, northern Borneo and northern Sulawesi. All the features of "Philippine-type languages", including predicate-initial word order, a four-way voice distinction, the case marking of arguments with a genitive-ergative syncretism, the existence of second position pronominal and adverbial clitics, and the equational nature of predication (Himmelmann 2005) appear to be conservative traits inherited from PMP if not from an earlier proto-language. Various morphological and syntactic innovations take place just south of this zone among the Austronesian languages of Malaysia, Indonesia, East Timor and the Pacific.[1]

The highly conservative nature of Philippine morphosyntax, from a historical perspective, gives enormous importance to the languages of the Philippines for reconstructing the grammar of PMP and, consequently, understanding the grammatical innovations which define Malayo-Polynesian subgroups outside of the Philippines.

In this chapter, I give a broad typological overview of a subset of Blust's (2019) Philippine subgroup. For the most part, the languages of this putative subgroup fall within the modern political boundaries of the Philippines, although the Sangiric, Minahasan and Gorontalic subgroups are found on northern Borneo and Sulawesi.[2] The focus here is on those languages within the Philippines proper.

[1] The Gorontalic and Bolaang-Mongondow languages of northern Sulawesi, which clearly form a phylogenetic unit with a number of Philippine subgroups to their north (Blust 1991), show this most clearly, as they have lost many of the aforementioned features in more recent history.

[2] I have reglossed the functional morphology in many of the examples here so that the terminology employed is as uniform as possible throughout. I do not mean to impose a particular analysis on the data by the use of "nominative" and "genitive" case, nor do I mean to imply that all forms glossed as "actor voice" are syntactically identical across languages. I transcribe examples of nasal substitution (triggered by the PMP prefixes *paŋ-/maŋ-) with deleted consonants in square brackets, e.g. *maŋ-[k]uːha*. I have also aimed to represent all the data presented here in a broad IPA transcription to avoid confusion across orthographies, although I maintain the symbol <y> for the palatal glide, as opposed to IPA [j]. Finally, any numbered examples whose language is not specified in the first line are Tagalog.

2. History of linguistic studies

There may have been a precolonial tradition of metalinguistic analysis, considering that coastal cities like Manila and Butuan hosted large multilingual trading communities and that Buddhism, Hinduism and Islam, all of which have their own traditions of linguistic analysis, took root on Philippine soil before the arrival of the Spanish. Unfortunately, if such traditions existed, no traces have been documented, as so much of pre-colonial culture of the lowland regions had been erased by the Spanish conquest.

A number of Philippine languages throughout the region were written with a native script, most likely developed from Indic models and related to similar writing systems found in Indonesia. Some of these scripts, such as those of the Buhid and Hanunoo in Mindoro, and the Tagbanwa in Palawan, have been passed down from generation to generation up until the present. Others, such as those of Tagalog, Ilokano, Kapampangan and Bisayan, had fallen into disuse after the introduction of the Latin alphabet, but are now being revived through various grassroots efforts.[3] All of these writing systems are closely related to each other while showing distinct variations by language group.

Ancient writing appears to have been employed primarily for the purposes of sending brief messages and composing poetry. The primary writing material was bamboo, which decays rapidly in tropical climates without special care. It is thus not surprising that we lack evidence for the use of writing in long term record keeping. No authentic traces exist of these alphabets written on stone, metal or pottery, although the Laguna copperplate inscription, dating from 900 CE, shows that Old Malay written in the Kawi (Javanese) script may have been used in more official functions in certain parts of the Philippines (Postma 1992).

The Spanish colonial period begins with the arrival of Magellan in 1521 and proceeds until the end of the 19th century, when the United States took over control of the Philippines from Spain in 1898 under the Treaty of Paris and held power over it for 48 years until Philippine independence in 1946. The earliest preserved descriptions of Philippine languages date from the 16th century, during which Spanish missionaries learned Philippine languages in order to convert and subjugate the native population. A large number of catechisms, dictionaries and grammars were produced in languages of the lowlands such as Tagalog, Ilokano, Cebuano and Hiligaynon. As Manila became the center of the Spanish colony, much effort was expended in producing grammars and dictionaries of Tagalog. To a large extent, Spanish missionary grammars of Philippine languages conformed rigidly to Latin models, following Nebrija's (1492) highly influential *Grammatica de la lengua castellana*, one of the first grammars of any modern European language. But like Nebrija's grammar, which shoehorns Spanish into a Latin model, the Spanish grammarians also applied poorly fitting categories and concepts from Latin to Philippine languages. Reminiscent of Alice in Wonderland's application of Latin to conjugate 'mouse' ('a mouse, of a mouse, to a mouse, a mouse, O mouse!'), the Spanish grammarians ignored the underlying organization of the case system in favor of seeking equivalences for the nominative, accusative, dative, ablative, genitive and vocative cases of Latin. Despite a professed appreciation on the part of many missionary grammarians for the unique characteristics of languages like Tagalog and Cebuano, the templatic nature of early European vernacular grammars proved to be an insurmountable obstacle to treating Philippine languages in their own

[3] Postma's (1972) work on the Mangyan writing of southern Mindoro was the first book length work to examine the traditional but contemporary use of a native Philippine script and Pangilinan (2012) discusses the revival of the Kapampangan *kulitan* system.

terms. Nonetheless, early Spanish grammarians succeeded in capturing key elements of grammar, most importantly, the four-way voice system that typifies the vast majority of Philippine languages. For instance, Ridruejo (2011:25-26) shows that in one of the earliest colonial grammars, the *Arte de la lengua Sambala y Española* from 1601, the author already recognized the importance of uniqueness and referentiality in voice selection and attempted to attribute a meaning to each of the "passive" markers. Also noted in some of the colonial period grammars is the flexibility with which all roots can be turned into verbs, by dint of the voice morphology that is generally required of event-denoting predicates (cf. Zwartjes 2011:73 on Oyanguren's 1742 Tagalog grammar). Sebastian de Totanes's (1740) *Arte y reglas de la lengua tagala* is often cited as the apogee of this descriptive tradition and it had a considerable influence on 19th century European scholars who did not have the opportunity to conduct research in the Philippines.

The 19th century heralded a shift towards comparative and historical linguistics for its own sake and Austronesian languages attracted the attention of some of the most prominent language scholars of the time, in particular, Wilhelm von Humboldt and Franz Bopp. Humboldt (1838) had a remarkably keen intuition that Philippine languages preserved a rich morphology that had been largely lost in languages such as Malay and Javanese. He considered Tagalog verbal morphology, in particular, as "the highest possible level that the grammatical organism can achieve in this language family" (cited in Werlen 2011:127) and thereby brought new prominence to the study of Philippine languages. Previously, Old Kawi (Javanese), which had a very well preserved tradition of writing, had been seen as a more natural candidate for a conservative window on to the past, a fallacy which misled Bopp into drawing an erroneous connection between Austronesian and Sanskrit. Humboldt was also responsible for innovative remarks on the nature of lexical categories in Tagalog, proposing two possible theoretical approaches which are still very much debated today (cf. Starosta et al. 1982; Ross 2002; Foley 2017, 2023; Himmelmann 1991, 2008; Kaufman 2009b,c): either Tagalog roots are acategorial and are derived into nouns and verbs by morphological processes, or they are essentially nominal and all verbs are denominal. Humboldt did a good deal to free Philippine grammatical studies from the procrustean bed of the Spanish colonial grammarians, an impressive task given that these grammars provided his only source of information on the languages.

The 19th century saw the beginnings of a native tradition of linguistic analysis embodied in the work of the first Filipino lexicographer, Pedro Serrano-Laktaw, and to a lesser extent the linguistic sketchings and orthographic reforms of national hero Jose Rizal. This tradition was continued into the 20th century by native linguists and literary figures such as Lope K. Santos, Cecilio Lopez and later figures such as Videa de Guzman, Resty Ceña, Nelly and Ernesto Cubar, Ernesto Constantino, Teresita Ramos, Fe Otanes, Consuelo Paz, Ma. Lourdes Bautista, Andrew Gonzalez, Paz Buenaventura Naylor, Ricardo Nolasco, among others.[4] The 20th century also brought the advent of the Summer Institute of Linguistics (SIL), which entered the Philippines in 1953 and has produced numerous publications documenting, describing and classifying the lesser known languages of the Philippines.[5] For a more complete linguistic historiography of the

[4] For a representative sample of this work, see Santos (1940), Constantino (1977), de Guzman (1978), Ceña (1977,2012), Ceña & Nolasco (2011), Cubar & Cubar (1994), Ramos (1973), Fe Otanes (1966), Schachter & Otanes (1972), Paz (1981), Bautista (1975), Gonzalez (1981), Naylor (1975, 1986), Nolasco (2003).

[5] According to an early mission statement (Wolfenden 1961), "The Summer Institute of Linguistics is a non-profit corporation of over 1100 members engaged in language studies on five continents. Its concern is to analyze written languages, to provide literacy materials for them, and to climax this effort with translations of the Christian Scriptures."

Philippines, see Salazar (2012), who offers a detailed summary of European works on Philippine languages from the beginning of the colonial period until the mid 20th century, and Garcia (1992), who analyzes late 19th and 20th century grammars of Tagalog. Johnson (1996) provides a comprehensive bibliography of extant descriptive materials on Philippine languages until the mid 1990s.

3. Subgrouping

Linguists have attempted to understand the relations between Philippine and other Austronesian languages for over two hundred years but it was Otto Dempwolff's (1934-1938) Vergleichende Lautlehre des Austronesischen Wortschatzes that ushered in a new standard of rigor with over 2,200 novel reconstructions based on a variety of new data. Dempwolff's objective, however, was reconstruction rather than subgrouping and he thus chose languages that he felt were sufficiently disparate without a clear hypothesis about the Austronesian family tree. The introduction of glottochronology in 1950, just prior to the SIL's entry into the Philippines, led to a proliferation of lexicostatistical studies that informed the subgrouping of Philippine languages for many years. Use of the comparative method (still often in conjunction with lexicostatistics) and more through dialect geography in the years that followed led to more reliable results as well as new reconstructions (e.g., Tharp 1974, Paz 1981, Charles 1974; Reid 1974, 1989, 2006, 2017; Himes 1998, 2012; McFarland 1977, 1981; Lobel 2013; Robinson and Lobel 2013; Zorc 1974, 1977, 1986, 2019; Blust 1991, 1992, 2005, 2019, 2022; Gallman 1983; Burton 1996; Elkins 1986; Savage 1986; Fleischman 1981; Thiessen 1981, to name some of the more prominent works).[6]

Zorc et al. (2024) present a general consensus subgrouping of Philippine languages that includes the thirteen following branches from north to south, with the last two groups, Sangiric and Minahasan, being almost entirely located in North Sulawesi on the Indonesian side of the border (with the exception of the Sangiric languages of Sarangani Island).

1. Batanic (3)
2. Northern Luzon/Cordilleran (51)
3. Central Luzon (10)
4. Umiray Dumaget (1)
5. Manide-Alabat (2)
6. North Mangyan (3)
7. Greater Central Philippines (and possibly Palawanic) (94)
8. Kalamianic (2)
9. Inati (1)
10. Southwestern Mindanao (formerly 'Bilic')
11. Southeastern Mindanao (formerly 'Bilic')
12. Sangiric (5)
13. Minahasan (5)

[6] The literature on subgrouping is summarized by Zorc et al. (2024) and Blust (2019) and so will not be rehearsed here in detail.

As can be seen above, some of these subgroups contain a very small number of languages, for example, Kalamianic, Manide-Alabat and Inati, the latter of which contains a single language. On the other hand, two larger subgroups contain the bulk of Philippine languages: Northern Luzon (Tharp 1974, Reid 2006) and the Greater Central Philippines (GCP) subgroups (Blust 1991). Northern Luzon contains the following subgroups:

1. Arta (1)
2. Cagayan Valley (12)
3. Iloko (1)
4. Meso-Cordilleran (32)
    a. South-Central Cordilleran (30)
        i. Central Cordilleran (23)
        ii. Southern Cordilleran (7)
5. Southern Alta (1)
6. Northeastern Luzon (4)

The GCP subgroup contains the following subgroups, with some of their better known members in parentheses.[7]

1. Central Philippine (Bisayan, Bikol, Tagalog, Mansakan…)
2. South Mangyan (Hanunoo, Buhid, Eastern Tawbuwid, Western Tawbuwid, Bangon…)
3. Palawanic (Batak, Aborlan Tagbanwa, Molbog, Brooke's Point Palawan…)
4. Subanen (Northern S., Eastern S., Central S., Western Subanon, Kolibugan)
5. Danao (Maranao, Maguindnao, Iranon)
6. Manobo (Sarangani Manobo, Talaandig, Higaonon, Kinamiging, Kagayanen…)
7. Gorontalo–Mongondow (Mongondow, Gorontalo, Buol, Ponosakan…)

A recent proposal by Blust (2019) building on ideas put forth by Zorc (1986), Charles (1974), and his own earlier work (Blust 2005), argues that all the languages of the Philippines, together with certain adjacent MP subgroups in northern Sulawesi and Sabah but excluding the Sama-Bajaw languages, descend from a uniquely shared ancestor, termed Proto-Philippines. This proposal has met with varied reactions. While some have endorsed it on the basis of the large number of proposed lexical innovations (Zorc 2020), others have voiced doubts based on the lack of phonological and morphological innovations (Liao 2020, Ross 2020, Smith 2017, Reid 2020, Kaufman 2026, Chen et al. 2026). It is well known that lexemes are the most easily borrowed element of language (in comparison to sound change and morphological innovations) and it is thus hard to know what cognates may be lurking outside the Philippine group given that the lexicons of many key border languages have not been documented, and those that have been documented may have not been thoroughly examined. Problematically, the Sama-Bajaw languages, which all agree must be external to any potential Philippine subgroup, contain many of the lexemes that Blust (2019) claims to be Proto-Philippine innovations, seemingly more than some of the member languages of northern Sulawesi, demonstrating how easily words spread. In sum, the hypothesis that higher level subgroups of the Philippines are primary branches of

[7] GCP is held together mostly by a small number of lexical innovations in addition to a single sound change, PMP *R>g, which occurs frequently outside of the group as well, and may face the same controversies that have bedevilled the Proto-Philippine hypothesis itself, discussed below.

Malayo-Polynesian still stands as a viable contender to the Proto-Philippine hypothesis. On the former approach, major Philippine subgroups are no more closely related to each other than they are related to Malayo-Polynesian subgroups of the Indo-Malaysian region. Whatever unique lexical items they share may either be due to network effects over millennia, or due to certain lexemes being commonly lost outside the Philippine zone (as most probably occurred with certain grammatical and prosodic features). With thousands of years of trading, migration and feuding within a complex maritime network, it is no simple matter to tease apart the dual roles of contact and inheritance in the formation of subgroups.

Relatively few works have focused on contact effects between indigenous languages of the Philippines. Pallesen (1985) remains one of the most detailed examinations of such phenomena in the Philippine context, focusing on bidirectional convergence between Tausug, a Central Philippine language, and the Sama languages of the Sulu zone. Other works on contact, such as Wolff (1967) and Lobel (2026), focus on linguistic traces of earlier Bisayan languages that have been largely submerged by the expansion of Cebuano. Gallman (1983) and Burton (1966) focus on borrowing and subgrouping in Mindanao, where several subgroups of Greater Central Philippine languages have long been in contact. Finally, Gallego (2022) examines contact between Batanic languages, Ilokano and Tagalog in the northern tip of the country. Far more work focuses on contact effects between Philippine languages, especially Tagalog, and the colonial languages, Spanish and English, which is touched upon in the following section.

4. Language vitality and the sociolinguistic situation

Until the late 18th century, the Spanish friars who ruled the Philippines actively discouraged Spanish language literacy in the fear that direct communication between their subjects and the Spanish crown would jeopardize their grip on power. As a result, only a relatively small percentage of Filipinos were ever fluent in Spanish and so, unlike in Latin America, Spanish never threatened the indigenous languages of the country. It did however serve as the everyday language of the mestizo elite and as the official language of government and education and thus prevented the advancement of indigenous languages as official languages in their own right.

Comparing the historical records from the earliest days of European contact (e.g. the *Doctrina Christiana* of 1593) with the current linguistic landscape of the Philippines, the stability and continuation of indigenous languages in the face of multiple heavy-handed colonial regimes is striking. It is even more surprising given the forced migration of so much of the indigenous population into encomiendas controlled by the Spanish authorities. The displacement and intermixing of communities on such a scale would, at the very least, be predicted to give rise to widespread dialect levelling. Yet when we look at the modern dialect geography of places such as the Bikol region (McFarland 1974, Lobel et al 2000, Anderson 1974), we find an intricate network of dialects involving features that make nearly each town linguistically unique, most of which appear to have been inherited from the precolonial era. Compared to the large-scale changes that affected English over the last 500 years, the contrast is indeed stark. With some exceptions, the changes we do find over the last 500 years have remained relatively understudied, as it is only in the last two decades that many of the key colonial grammars have been made more widely available through various digitization initiatives and republication efforts.

Another *prima facie* surprising fact concerning linguistic stability in the Philippines is the relatively small number of languages that have gone silent over the last several centuries. Those

languages that have been most affected by language shift are those of the aboriginal pre-Austronesian populations (most often referred to as 'Negrito') who already occupied the archipelago for tens of thousands of years when speakers of Austronesian languages first arrived. This aboriginal population exists in separate groups throughout the Philippines. Traditionally, they are nomadic hunter-gatherers rather than rice farmers and are often referred to with an appellation derived from PMP *qaRta(q), which Blust (1972) reconstructs with the meaning 'outsiders, alien people', but which replaces the generic words inherited from PMP *tau 'person' in many languages.

There is little trace of a pre-Austronesian signal in the languages of aboriginal populations in the Philippines, although Reid (1994b) argues that a small handful of pre-Austronesian lexemes can be identified across several geographically disparate languages. In a series of publications, Reid (1987, 1989, 1994a,b, 2013) argues more generally that the aboriginal languages preserve early Austronesian morphology more faithfully than the languages spoken by farming populations (whose descent comes in larger part from the original Austronesian travellers themselves), attesting to a period of early contact. Most strikingly, several aboriginal languages preserve a disyllabic form of the verbal prefixes containing the infix *<in> while most other Philippine languages have converged in reducing these disyllables to monosyllables in various ways. A clear example of this is seen in Reid's (1987) comparison of Casiguran Dumagat, a language spoken by aboriginal hunter-gatherers, with Paranan, a language spoken by neighboring rice agriculturalists in Table 1.

Table 1. The development of verbal prefixes (Reid 1987:47)

PMP	Casiguran Dumagat	Paranan
*m<in>aR-	minag-	nag-
*m<in>aN-	minaN-	naN-
*m<in>a-	mina-	na-

Additional evidence for early contact comes from the fact that three of the thirteen subgroups presented by Zorc et al. (2024) are spoken by small aboriginal populations that do not share in the linguistic innovations that define the surrounding languages spoken by agriculturalists. The PMP *Z/*d merger, which is the single phonological innovation seemingly shared by all the subgroups descended from the putative proto-Philippines (Blust 2019), does not, in fact, appear to have taken place in Inati (Pennoyer 1986-1987), a language spoken by an aboriginal people of Panay island in the western Visayas.[8]

It is also the languages of aboriginal Filipinos that are the most endangered due to their long history of marginalization and exploitation (Headland 2003, 2010, Lobel and Surbano 2019). Whereas in previous centuries this marginalization may have led to the preservation of these languages, the encroaching world of 'mainstream' Filipino culture and the rapid destruction of the natural environment that sustained hunter-gatherers over millennia has pushed aboriginal languages towards a tipping point.

Language endangerment in the Philippines has a very different appearance when compared to Latin America or other parts of the formerly colonized world in which a European colonial language displaces an indigenous language through language policy. In the Philippines,

[8] Nonetheless, it must be emphasized that the languages of aboriginal Filipinos have often been unjustifiably exoticized in the literature, considering the fact that they typically subgroup closely with neighboring languages spoken by non-aboriginal populations. This tendency is investigated critically by Zamar (2023).

it is not Spanish or English, or even Filipino, the Tagalog-based national language, which are the primary targets of language shift. Rather, speakers of endangered Philippine languages are shifting to the more prestigious languages of the surrounding community, which are typically the regional lingua francas such as Cebuano, Hiligaynon and Ilokano, in addition to Tagalog. But as Headland (2003) points out, language shift need not represent a wholesale abandonment of the mother tongue. It can take the form of massive borrowing to the point where only a few functional items and conservative lexical items betray the original affiliation of the local language. Wolff (1967) and Lobel (2026) examine two such cases in the Visayas, where speakers of Porohanon and Bantayanon have been shifting to Cebuano for at least several decades. Only through careful analysis of the morphology, syntax and conservative lexicon can we trace the ancestry of the language, which could otherwise easily be mistaken for a dialect of Cebuano, given large-scale borrowing. Two other examples are Kinamiging and Kagayanen, Manobo languages that have been under similarly heavy contact with Bisayan languages. From what little is known of Kinamiging, spoken on the island of Camiguin facing the city of Cagayan de Oro in northern Mindanao, it is rapidly converging with Cebuano as the population has assimilated culturally to the surrounding Bisayan culture. Kagayanen is geographically isolated in the Sulu Sea between the much larger islands of Palawan and Negros, but had been under the jurisdiction of Bisayan speaking Antique Province for roughly 300 years. The language is relatively stable on the island of Cagayancillo according to Pebley and Payne (2024), yet the language itself is rife with Bisayan loan words and looks substantially different from its Manobo relatives in Mindanao (leading King et al [2024] to incorrectly subgroup it with the Bisayan languages based on a Bayesian analysis of the lexicon).

The most extreme product of language contact is found in the Spanish lexified creole languages, Zamboangueño Chavacano, spoken in Zamboanga City, Mindanao, and Caviteño and Ternateño Chavacano, spoken in the province of Cavite, Luzon (Forman 1972, Pérez 2015, Steinkrüger 2008 *inter alia*). Each of these areas contains a distinct dialect but the overall structure of the language is very similar. The creole shows the word order patterns of the Philippine substrate but has been relexified by Spanish.

English was introduced into the Philippines in 1898 at the outset of the American colonial period and remains an official language in all capacities until today. The national language, now called Filipino as per the 1987 Philippine Constitution, is impossible to differentiate from Tagalog, despite having been planned as an amalgam of indigenous languages. In practice, Philippine multilingualism typically involves constant code-switching, mixing and spontaneous borrowing between indigenous languages as well as English and other non-indigenous languages. In a typical rural town, one can expect mixing between the local language, the regional lingua franca (e.g. Cebuano, Ilokano, Ilonggo), Tagalog/Filipino and English, depending on context, addressee, and purpose.

Sociolinguistic studies in the Philippines have largely been restricted to questions of language choice, lexicon and language contact (Bautista 1975, 1996, 1998; Pascasio 1977, 1984; see Gonzalez 1985 for an overview until the mid-20th century). There are few if any variationist studies of Philippine languages in the Labovian style and this is an area of study which could be very fruitful in the Philippine context, where social class is typically reflected across an array of sociolinguistic variables.

5. Phonology

5.1 Segment inventories

Philippine vowel inventories are typically simple. The Central Philippine languages typically either preserve the Proto-Austronesian four vowel system (*i, *u, *a, *ə) or conflate it to a three-vowel system by merging *ə with one or more of the other vowels. In the languages of Mindanao, *ə is often preserved as a high central vowel (ɨ), and this was clearly the case in the not so distant past for many of the Central Philippine subgroups (Wolff 1967, 1968). In rare cases, the inherited vowel inventory has been expanded in complex ways, for instance, Tboli and Blaan, with their seven-vowel systems (Porter 1977; Forsberg 1992; Abrams 1963), and Casiguran Dumagat, with its nine-vowel system (Headland and Wolfenden 1967).[9]

Several languages have developed an allophonic relationship between the high vowels and their mid counterparts. In many languages, including Tagalog, a generally word-final process of vowel lowering turns *i* and *u* into *e* and *o*, respectively (although this is more common with the back vowels). Kapampangan of the Central Luzon group has innovated a new set of mid vowels not from lowering of high vowels but rather through monophthongization of **aj* > *e* and **aw* > *o* (cf. Ibanag *don* 'leaf' from PMP *dahun after *h*-loss). Another source of mid-vowels includes optional harmonic processes triggered by final lowering, as in Ilokano *potoŋ* 'bamboo used as a storage container' and *poŋo* 'bunch, cluster' (from PMP *putúŋ and *puŋu, respectively (ACD)).

The falling diphthongs /aj/, /uj/, /aw/, /iw/ are very common across Philippine languages and are almost always treated as closed syllables, disallowing a coda consonant. Exceptions to this include Tausug (e.g., *lawm* 'inside'), where the historic loss of intervocalic /l/ led to innovative phonotactics. Many languages also exclude the homorganic diphthongs /uw/ and /ij/, as well as diphthongs containing the high-central vowel, /ɨw/ and /ɨj/.

Consonant inventories are also relatively simple. We typically find voiceless, voiced and nasal stops in three points of articulation, labial, alveolar and velar in addition to a glottal stop. Most languages also show two fricatives, /s/ and /h/, although the loss of /h/ takes place independently in several languages. In Batad Ifugao (Newell and Poligon 2005), among several other languages, *h>∅ is followed by *s>h, filling the former gap. The voiceless glottal fricative /h/ is typically restricted to onset positions except for Itbayat and a small handful of other languages (Blust 2018), where it can appear in coda positions, as well.

Other common sonorants include an alveolar lateral and bilabial and palatal glides. Many languages contain a tap or flap but this is most commonly an allophone of one of the previously mentioned alveolar phonemes (either /d/ or /l/).

We also find palatal obstruents at various stages of phonemicization, typically resulting from the combination of alveolars preceding /ij/, e.g. Tagalog 3SG.NOM /sija/ → [ʃ(j)a], 'there' /dijan/ → [dʒ(j)an], 'stomach' /tijan/ → [tʃ(j)an]. In a rarer development, Boholano has developed a voiced alveopalatal affricate from a historical palatal glide (i.e. PMP *y > dʒ).

Among the more unusual sounds described for Philippine languages we find the interdental approximant [ð̞] in a number of geographically far flung languages including Kalinga, Kagayanen and Kalagan, as described in detail by Olson et al (2010). In all these languages, the sound arose as an allophone of historical *l and in some languages it actively alternates with it.

The fortis/heavy stops of Maranao, described by Lobel and Riwarung (2009), and the aspirated stops of Subanen, described by Lobel and Hall (2010) form part of a larger set of

[9] See Reid (1973) for a detailed exposition of Philippine vowel systems and their development.

unusual reflexes of consonant clusters in the languages of Mindanao and northern Borneo, a fact that Lobel and Hall (2010:336-337) tentatively attribute to language contact. An unrelated phonological rarity in the same region that has not yet been noted is pharyngealization of nasals in Northern Subanen. The backing of the voiceless velar stop in at least certain varieties of Northern Subanen, a process common to many Philippine languages, has extended to voiced stops and nasals. Most remarkably, historical *ŋ takes on a pharyngeal articulation which could be represented as either [nˤ] or [ɴ̃ˤ], a sound which does not appear to have been recorded in any other language of the world.

5.2 Phonotactics

The canonical lexical root in Philippine languages is a disyllable with the following template: CV(C).CV(C), although many trisyllabic roots are also attested, in addition to a few quadrisyllabic roots. On one common analysis, there are no true vowel-initial syllables in lexical roots (Zorc 1977:52). Roots that appear to be vowel-initial (and are treated as vowel-initial orthographically) begin with a glottal stop.[10] Many Philippine languages lack vowel hiatus, as vowel hiatus relies on the possibility of onsetless syllables.[11] Root-initial glottal stops, whether they are underlying or epenthetic, surface predictably with prefixation in most languages, as in Tagalog /mag-(ʔ)abut/ (AV-reach) → [magʔabot], rather than *[magabot]. On the most transparent analysis, all syllables in lexical roots begin with a consonant while codas are optional. In languages that allow monosyllabic roots, many of these roots are transparently derived from disyllables either synchronically or historically. In Western Albay Bikol, this comes about by the loss of the glottal stop and fricative, PMP *dahun 'leaf' > *don*. In Tagalog and Ivatan, this comes about by the reduction of vowel-glide sequences *uw* > *w* and *ij* > *i* so that words like /buwa/ 'areca nut' and /bijaj/ 'life' are now [bʷa] and [bʲaj] in Ivatan (Maree 2007:24). Similarly, words like /buwan/ 'moon' and /bijak/ 'split' become [bwan] and [bjak] in Tagalog in the synchronic phonology. In several languages of the Sulu archipelago, the deletion of intervocalic /l/ has created monosyllables with long vowels, e.g. PMP *beli 'buy' Tausug /bi:/. Finally, Tboli shows another pattern of historical truncation via vowel deletion, e.g. PMP *epat > *fat* 'four', PAN *kaen > *ken* 'eat'.

Affixes do not have the same constraints as lexical roots; they are often monosyllabic and need not contain onsets. Onsetless affixes are typically provided with an onset either through epenthesis or infixation, the latter which only applies at the left edge of the base. When onsetless suffixes attach to stems that end in a vowel, either deletion or epenthesis avoids vowel hiatus. This latter process can be seen in Tagalog and Tagakaulo in (1a) and (b), respectively. The fricative /h/ is often used in this epenthetic capacity as it is not phonemic in root final position but glides also fulfill this role as in Tagakaulo.

		Tagalog		Tagakaulo (Burton 2018)
(1)	a.	*bagu-<u>h</u>in*	b.	*bagu-<u>w</u>un*
		new-PV		new-PV

[10] Central Tagbanwa is apparently one of the very few Philippine languages described as contrasting vowel initial syllables with glottal initial syllables (Scebold 2003:30). See Zorc (1982, 1996) for the historical challenges posed by the glottal stop in Philippine languages.

[11] Words that are written with two vowels orthographically in languages like Tagalog, e.g. *bait* 'goodness', contain an intervening glottal stop, e.g. [baʔit]. Zorc (1977:54) mentions Cuyunon and certain dialects of Tausug as exceptional in allowing vowel hiatus. Wolff (2005:70) further cites Itbayatan, Ivatan, Dumagat and Pangasinan, as allowing hiatus.

In Ilokano and other North Luzon languages, root final high vowels often become glides under similar conditions, as seen in (2).

		Ilokano (Rubino 1997:28)
(2)		/laki-ɯn/ → [lakjɯn]
		man-PV

Infixes typically are of a VC shape but obtain an onset from the stem, as shown again for Tagalog and Tagkaulo in (3).

		Tagalog		Tagakaulo (Burton 2018)
(3)	a.	*s<um>agot*	b.	*t<um>ubag*
		<AV>answer		<AV>answer

Gemination is found in Bagobo, Mansakan, Kagayanen, the Sama-Bajaw languages and at least one dialect of Bikol (Blust 2013:229), but is more common among the languages of the north. In some cases, gemination is triggered by a preceding prefix, as in Tuwali Ifugao /maka-higid/ (proclivity-sweep) [makahhigid] 'loves to sweep' (Hohulin and Burquest 2014), or Botolan Sambal where gemination is contrastive both morpheme internally and across morphemes, e.g. /ma-galaw/ (ADJ-play) 'playful' vs. /mag-galaw/ (AV-play) 'to play' (Antworth 1979:3). In other languages, gemination comes about due to a preceding schwa, which, as a weak/short vowel may trigger compensatory lengthening.

Phonemic glottal stop often arises from the historical change PMP *q > ʔ which took place widely throughout the Philippines although its phonotactic distribution varies by language and region.[12] For instance, PMP *baqeRu 'new', reduced historically to a disyllable, yields (Naga) Bikol *baʔgo*, Cebuano *bagʔo* (with metathesis), and Tagalog *baːgo* (with deletion and compensatory lengthening). These changes follow a general pattern as Cebuano does not allow ʔC clusters and Standard Tagalog does not allow either Cʔ or ʔC clusters.[13] Similar cases of metathesis are triggered by syncope when the resulting cluster is excluded by the general phonotactics of a language. Some of these clusters are universally absent in certain subgroups. For instance, Zorc (1977) cites **nm* as an unattested cluster in Bisayan roots and one that is actively avoided in forms that undergo syncope, as in /inum-an/ drink-LV which yields [imnan] with metathesis of the nasal consonants after deletion of medial /u/ (cf. Blust 1979, who makes a broader claim regarding the markedness of coronal-noncoronal clusters and their avoidance).

In many languages of the northern Philippines, root final glottal stop has been lost completely. In Tagalog, word final glottal stop is less phonetically salient than in the south and it is often lost in phrase medial position with compensatory lengthening. Even within small subgroups, we find variation in the distribution of glottal stop. In the three members of the Danao languages, Maranao allows stem/word final glottal stop but Iranun and Maguindanao have both eliminated it in this position.

The Bilic languages of Southern Mindanao are exceptional with regard to the typically simple syllable margins of Philippine languages. Tboli allows for a large number of typologically

[12] Zorc (1996) has argued for the need to reconstruct glottal stop separately from reflexes of PMP *q.
[13] Very few Philippine languages allow both Cʔ and ʔC clusters. Manide (Lobel 2010) is one such exceptional language.

rare onset clusters that violate the principle of sonority sequencing with regard to manner and voicing, e.g. /btaŋ/ ‘fall’, /tboli/ ‘Tboli’.[14] Similarly, a neighboring languages of Southwest Mindanao, Bagobo-Klata, shows unusual onset clusters, as seen in /gkaːwaŋ/ ‘hips’ and /gtoɔjɔ/ ‘kill-IRR.PV’ (Estrera 2004:61).

Complex tautosyllabic clusters have also entered Philippine languages through Spanish and English borrowings, e.g. Tagalog *plato* ‘plate’, *preno* ‘brake’. An illustrative example is seen in the Spanish loan *sombrero*, which enters Tagalog at a very early stage as *sambalilo*, fully adapted to native Tagalog phonotactics, and again at a later stage as *sombrero*, with the non-native *br* cluster and free distribution of mid-vowels, which were originally word final allophones of high vowels.

All Philippine languages allow heterosyllabic clusters although each language exhibits its own constraints and tendencies. Interestingly, such clusters may be innovative and do not generally reconstruct to PMP (Blust 2013:62). The only clusters found at the PMP level as reconstructed in Blust et al. (2023) are either nasal+stop sequences, e.g. **simbuR* ‘to sprinkle’, or the result of reduplicated monosyllables, e.g. **taktak* ‘to fall, of many things at once’. However, there are many apparently reconstructable lexemes in Philippine languages which contain clusters that do not fit into either of the above patterns. It should be noted that gradient phonotactic patterns have not been examined systematically in Philippine languages and present a rich area for further study.[15]

5.3 Phonological processes

The phonology of many Philippine languages is relatively transparent in that surface forms do not differ substantially from what would be posited as underlying forms. Attested processes include palatalization, lenition, fortition, metathesis, and compensatory lengthening, discussed below.

5.3.1 Lenition

Tapping, a type of lenition, takes place in Tagalog morpheme internally, between a prefix-stem boundary as well as between a word-enclitic boundary. Tapping does not occur in Tagalog between proclitics and their following hosts, as seen in (4).

Tagalog

(4) a. /daː~datiŋ/ → [daːɾatiŋ] IPFV~arrive

b. /aku=din/ → [ako ɾin] 1SG.NOM=also

c. /maŋa=dagaʔ/ → [maŋa=dagaʔ], *[maŋa=ɾagaʔ] PL=rat

The Central Cordilleran languages and some of their neighbors, like Ibaloy, show a set of alternations that are not always easy to characterize as lenition or fortition. In Ibaloy (Ruffolo 2004:17), a simple phonemic inventory like the one sketched out above leads to a far greater

[14] While these can be broken up with a schwa, according to Awed et al. (2004), schwa insertion is optional. Whether this schwa should be analyzed as underlying or epenthetic has not been addressed in the literature.

[15] Benton (1971) examines the gradient phonotactics of Pangasinan while Zorc (1977:53) notes the existence of phonotactic constraints in heterosyllabic clusters but laments the lack of data to address its nature.

number of surface allophones due to this set of processes. To take one example, /d/ becomes [tʃ] at the beginning of a word or stressed syllable but is lenited to [r] intervocalically, and remains [d] word finally and before a consonant. A gemination rule further targets the first allophone when preceded by a schwa. While onsets are often subject to fortition at the beginning of words and stressed syllables, this environment triggers affrication and spirantization of stops in Ibaloy and the Central Cordilleran languages. This frication appears to be the result of a fortition process that devoiced and aspirated voiced stops. However, the aspiration in turn led to spirantization giving the ultimate impression of lenition as opposed to fortition. This type of process has led to an odd distribution of voiced stops in some languages. For instance, Reid (1973) notes that in Kalinga, the voiced bilabial stop only occurs as a syllable coda. In onset position, it surfaces as a voiceless affricate or stop, depending on the quality of the following vowel.

5.3.2 Palatalization

Allophonic processes of palatalization are found in a range of Philippine languages in which the palatal glide triggers assimilation of a preceding alveolar obstruent. This can be seen with Tagalog /sija/ 3sg.NOM → [ʃja] ~ [ʃa], /tijan/ ‘stomach’ → [tʃjan] ~ [tʃan]. /dijan/ ‘there’ → [dʒjan] ~ [dʒan].

Fukuda (1997:15) notes a palatalization rule before low vowels in the Kadaclan and Barlig dialects of Eastern Bontoc such that /baboj/ becomes [fʲafoj] and /gasot/ becomes [kʲasot] (together with separate processes of initial devoicing and lenition). As Liao and Reid (2024) note, this process may be connected to the typologically odd process of low vowel raising after voiced stops in the Northeastern Luzon languages, as described by Robinson and Lobel (2013).

The Bashiic languages in the north and certain Central Philippine languages of the south have acquired palatal consonants from different allophonic processes. Boholano, as mentioned earlier, has developed a /dʒ/ phoneme from the palatal glide and Itbayat, a Bashiic language, has developed palatals from velar stops (voiceless, voiced and nasal) adjacent to a high front vowel (Yamada 2002:8). Kapampangan, alone among Philippine languages, preserves PMP *ñ as a palatal nasal while all other Philippine languages have merged PMP *ñ with PMP *n.

5.3.3 Syncope and metathesis

The canonical Austronesian root is disyllabic and trisyllabic stems are reduced to disyllables through an active rule of syncope in many languages, exemplified by Agutaynen in (5).

(5) Agutaynen (Quakenbush et al. 2010:41)

a.	/balet-en/	→	[balten]	b.	/b<in>etaŋ/	→ [bintaŋ]
	respond-PV				<PFV>put	

When syncope creates a cluster that is otherwise unattested, a phonological process typically repairs the output. In Agutaynen, a debuccalization process C → ʔ repairs certain clusters, as shown in (6), while in other cases, metathesis is employed, as in (7).

(6) Agutaynen (Quakenbush et al. 2010:42)

/te~teled/ → tetled → [teʔled]

PROG~enter

Agutaynen (Quakenbush et al. 2010:41)

(7) a. /pa-belag/ → pablag → [palbag]

CAUS-separate
b. /pa-belet/ → pablet → [palbet]
CAUS-borrow

Syncope can also feed place assimilation as in (8a), where the penultimate vowel has been deleted leading to assimilation of the nasal to the following bilabial stop. In (8b), the geminate in the infix blocks syncope, as the resulting form would yield an illicit cluster.

Bugkalot
(8) a. /l<in>ɨpɨd/ → [limpɨd]
<PRF>cut
b. /l<imm>ɨpɨd/ → [limmɨpɨd]
<PRF.AV>cut

In contrast, Ilokano syncope occurs with degemination to accommodate regular phonotactic constraints, as shown in (9).

Ilokano (Rubino 1997:28)
(9) a. /na-punno/ → [napno]
STA-full
b. /na-ləppas/ → [nalpas]
STA-finish

Neither syncope nor metathesis are productive in Tagalog but both processes are richly attested in allomorphy, as seen in (10) (see Blust 1971 for the complex interaction of metathesis and assimilation in this pattern).

Tagalog (Bloomfield 1917:391)
(10) a. /atip-an/ → atpan → [aptan]
roof-LV
b. /silid-an/ → sildan → [sidlan]
room-LV

A more limited type of metathesis is seen with bilabials in Western Subanon. Seemingly as a strategy to avoid a sequence of syllables beginning with bilabial stops, Western Subanon shows metathesis of the adjectival prefix *mo-* when it attaches to stems that begin with a bilabial stop (Bulalang 2025:40). Thus, /mo-bogan/ 'ADJ-lightness' becomes [ombogan]. Pangasinan shows a similar metathesis with a cognate prefix, at least historically. This appears most commonly with bilabial initial roots (e.g. [amputi] 'white', [ampetaŋ] 'warm', [ambetel] 'dry', [ambasa] 'wet', [ambelat] 'heavy') but this is also attested with roots that start with non-bilabials and so the trigger is not as clear as in Western Subanon.

Syncope can be seen as a strategy of creating favored disyllabic stems from disfavored trisyllabic ones. Another strategy for achieving the same ends is syllable deletion. While rare, Estrera (2024) describes such a process in Bagobo-Klata, where it is fed by the addition of a select group of suffixes (irrealis locative voice/nominalizer *-a* and irrealis patient voice *-ɔ*), as exemplified in (11).

Bagobo-Klata (Estrera 2004:66)

(11)a. /ka:pɛʔ-a/ → [pɛʔa]
hold-IRR.LV

b. /ha:kɔj-a/ → [kɔ:ja]
ride-IRR.LV

5.3.4 Vowel reduction and harmony

Vowel reduction is found throughout the Subanen languages (Arms 1996:5, Bulalang 2025), as well as Bornean languages south of the Philippines. In Western Subanon (Bulalang 2025:28), /a/ and /o/ merge in penultimate position. Western Subanon /o/ is the regular reflex of PMP *ə and its status as a weak vowel remains after the historical change in articulation.

A similar process turns /a/ into [ə] in Arta although it can take place in the penultimate syllable, as well, so long as the following syllable contains a high vowel or schwa (Kimoto 2017:47).

Lobel and Riwarung (2009, 2011) describe vowel harmony in Maranao where two complementary sets of vowels have developed, a "lax" set, [ɪ, ə, o, a], and a corresponding "tense" set, [i, ɨ, u, ʉ]. They show that the set of consonants they term "heavy", represented as /p', t', k', s', h/, obligatorily trigger the tense allophones of the following vowels. The voiced stops /b, d, g/ optionally trigger the tensing of the following vowel, and all other consonants condition the lax set. Because the heavy/light distinction on consonants plays an important role in the morphology, there are minimal pairs for every verb, as exemplified in (12). The "future" is signaled by the change of a light stem initial consonant to its heavy counterpart, and the consequent vowel harmony.

(12) Maranao (Lobel and Riwarung 2011:40)

a.	[t̪a.ʔa.man]	b.	[t̪'ʉ.ʔʉ.man]
	/taʔam-an/		/t'aʔam-an/
	taste-LV		FUT/taste-LV

5.4 Morphophonology

5.4.1 Infixation

Two productive infixes inherited from PAn, *<*um*> ACTOR VOICE and *<*in*> PERFECTIVE, continue to play an important role in Philippine languages (Reid 1992). They are positioned after the first consonant of the stem, as shown in (13) for Tagalog. Historically, both of these infixes could co-occur although unreduced combinations, as still seen in Naga Bikol (e.g. *k*<*um*><*in*>*ua*), are relatively rare among living languages (Lobel 2004).

Tagalog

(13)a. *k*<*in*>*u:ha*-Ø
<BEG>take-PV
'taken'

b. *k*<*um*>*u:ha*
<AV>take
'take'

Infixation is often externalized altogether in a process which turns *<*um*> into *mu-* and *<*in*> into *ni-*, as found in Cebuano. Reflexes of *<*in*> have also been reduced to a single segment in Danao languages (e.g. Maranao *t*<*i*>*abas* <PFV>cut), Tboli, Mansakan, and elsewhere in Mindanao. Many languages have innovated new infixes, nearly all of them of the shape CV or C

and entering the stem from the left edge. Bugkalot (known in the literature by the exonym Ilongot) shows what appears to be the only case of a right aligned infix among the languages of the Philippines. As seen in (14), the Bugkalot intensive infix is of the shape /Vl/ and is inserted before the final vowel of the stem while copying all the features of that vowel.

Bugkalot

(14)a.	*bës<il>ik*	'run intensively'	f.	*awd uk*	'hunt intensively'
b.	*maŋ<al>an*	'eat intensively'	g.	*pas<il>i*	'kill intensively'
c.	*tag<ɨl>ɨm*	'dance intensively'	h.	*kɨw<al>a*	'walk intensively'
d.	*pokn<al>a*	'hit intensively'	i.	*bëz o*	'mash with pestle intensively'
e.	*aɣ<al>a*	'take intensively'	j.	*klaŋ<il>it*	'shout intensively'

5.4.2 Reduplication

Philippine languages make heavy use of various types of reduplication for a vast number of purposes. Tagalog has two types of CV reduplication, one with and one without vowel length, as well as foot reduplication. CV reduplication without vowel length is found in agentive nominalization, shown in (15a), intensive formation, and elsewhere. CV reduplication with vowel length, shown in (15b), is used chiefly for imperfective/progressive aspect.

Tagalog

(15)a.	*mag-na~na:kaw*	b.	*mag-na:~na:kaw*
	AV-NMLZ-steal		AV-IPFV-steal
	'thief'		'will steal'

Foot reduplication in many cases is indistinguishable from full reduplication of the root, as shown in (16a), as most roots are disyllabic. However, larger stems, as in (16b), demonstrate that no process of reduplication in Tagalog copies more than a foot.

Tagalog

(16)a.	*ma-ganda~ganda=sila*	b.	*bali:~bali:taʔ*
	ADJ-MODER~beauty=3PL.NOM		MODER~news
	'They are moderately beautiful.'		'gossip'

Other languages, such as Central Tagbanwa, possess full word reduplication without such a maximality constraint, as seen in (17).

Central Tagbanwa (Scebold 2003:42)

(17)a.	*naka-tohod*	b.	*naka-tohod~naka-tohod*
	LOC-forest		LOC-forest~LOC-forest
	'in the forest'		'deep in the forest'

Multiple processes of reduplication can take place in the same word, as shown in Tagalog (18a), where (aspectual) CV reduplication applies to a stem that has already undergone (iterative) foot reduplication and in (18b), where (imperfective) CV: reduplication has applied to a stem that has undergone (intensive) CV reduplication.

Tagalog

(18)a. *mag-haː~hanap~hanap*
AV-IPFV~ITER~search
'will keep searching'

b. *p<in>ag-saː~sa~sabi*
<BEG>TR-IPFV~INTNS~search
'what is being said (intensively)'

Whereas Tagalog reduplication simply truncates a base that has more than two syllables, Cebuano and Bikol employ reduplication with fixed segmentalism for the same aim. Thus, for a trisyllabic Cebuano stem like *padala* 'send' we find *p<ulu>~padala*, where the first consonant of the stem has been copied and the following *ulu* is infixed, instead of **padala~padala* or **pada~padala* (see Mattes 2014:76 for additional complexities).

Word-based reduplication should be differentiated from a robustly syntactic process of reduplication which employs the linker or genitive case marking. These types of reduplication, shown for Central Tagbanwa in (19) and Tagalog in (20) (cf. Schachter and Otanes 1972:398), usually indicate repetitive action and are never affected by maximality constraints. Such constructions typically allow pronominal and other clitics to intervene between the base and the reduplicant, as in (20).

Central Tagbanwa (Scebold 2003:57)

(19) *t<um>umpok a t<um>umpok*
<AV>pile LNK <AV>pile
'kept piling up'

(20) *k<um>aːʔin ako naŋ k<um>aːʔin*
<AV.BEG>eat 1SG.NOM GEN <AV.BEG>eat
'I kept eating and eating.'

Most North Luzon languages show a rich inventory of reduplicant morphemes, some of which are highly uncommon outside this subgroup. Liao and Reid (2024) offer a thorough overview of the attested types and their varied functions across the languages of the northern Philippines, including habitual aspect, recent perfective aspect (with verbs), plurality (with kinship terms and human nouns), intensification (with property words), distributive and restrictive interpretations (with numerals), among other uses.

5.4.3 Nasal substitution

Philippine languages, like many other Malayo-Polynesian languages, display a morphophonological process termed "nasal substitution" with cognates of the sister prefixes PMP **paŋ-* DISTRIBUTIVE and **maŋ-* ACTOR VOICE + DISTRIBUTIVE. Nasal substitution refers to assimilation of the final nasal of these prefixes to the place of articulation of the stem-initial consonant accompanied by deletion of the latter, as in Tagalog (21).[16]

Tagalog

(21) /maŋ-baril/ → [mamaril]
AV.DIST-gun
'shoot'

[16] The nasal coda of the prefixes that trigger nasal substitution are often represented by N, a placeless nasal with special morphophonological properties. Blust (2004) reviews nasal substitution patterns across Malayo-Polynesian languages.

The deletion of the stem onset after nasal assimilation is not entirely predictable in Tagalog and other Central Philippine languages. Zuraw (2000) proposes a multifactorial analysis of this deletion for Tagalog, which must take into account the features of the first segment of the stem, as well as the stem's semantics and frequency. In other languages, nasal substitution patterns are completely predictable on the basis of phonology alone, typically with stem-initial voiceless segments undergoing deletion and voiced segments being maintained (Blust 2004). There may be unexpected phonological outcomes with the combination of **maŋ-* and certain stem-initial consonants. For instance, in Kapampangan, /s/ and /d/ trigger a palatal nasal instead of maintaining their alveolar articulation: /maŋ-dakáp/ [maɲakáp], /maŋ-salítaʔ/ [maɲalítaʔ] (Forman 1971:22).

5.5 Stress and prosody

The vast majority of Philippine languages have a phonemic stress/prominence distinction on roots which has long posed a challenge for reconstruction. As discussed in Kaufman & Himmelmann (2024), the basic feature that underlies the Philippine penultimate vs. final (aka paroxytone vs. oxytone) stress distinction is often best analyzed as a vowel length contrast in the penultimate syllable, as per Schachter and Otanes's (1972) analysis of Tagalog, Wolff's (1972) analysis of Cebuano, among others. This explains the otherwise puzzling complementary distribution of apparent penultimate 'stress' with penultimate syllable coda consonants in many Philippine languages. Nonetheless, some languages, like Kapampangan and Ivatan, do allow for vowel length distinctions in closed syllables, albeit only in derived contexts, e.g. Ivatan /igpit/ (hold) [ig.pit] vs. /i-igpit/ (CV-hold) [i:g.pit] (Maree 2007:29). Given their restricted distribution, such long vowels in closed syllables might be best treated as the result of coalescence between two vowels heading distinct syllables, at least on a more abstract level.

Central Philippine languages differ in whether closed penultimate syllables attract stress or, more neutrally, pitch prominence. In Tagalog, penultimate closed syllables do not attract pitch prominence nor can they co-occur with a long vowel and are thus predictably unaccented. In the Bisayan languages, on the other hand, closed penultimate syllables cannot bear a long vowel but do attract pitch prominence and are thus predictably accented. Thus, a root like /dakdak/, in isolation, would surface as [dakˈdak] in Tagalog but [ˈdakdak] in Cebuano and other Bisayan languages.

As noted by Blust (2013:251) and Kaufman & Himmelmann (2024), prosody is not phonemic in several languages of the southern Philippines. Revel-Macdonald (1979:63) describes a general absence of phonemic accentual distinctions in Palawan but the presence of final syllable lengthening, which gives the impression of final stress. The lack of contrastive prosody (penultimate long vowels) appears to be a contact feature in this area. Pallesen (1985) observes that the Tausug of Sulu lacks the prosodic distinctions found in Central Philippine languages but that the Tausug of Palawan, which originated in 19th century Sulu, maintains the distinctions found in other Central Philippine languages, concluding that the loss of this distinction in Sulu is a relatively recent phenomenon that came about through contact with Sama languages, which show predictable penultimate word stress.

Other languages without contrastive accent include Central Tagbanwa, which shows variable stress (Scebold 2003:27), Agutaynen, with penultimate phrase-based stress (Quakenbush et al. 2010:40), Matigsalug Manobo (Wang et al. 2006:3), and the Danao languages (Lobel and Riwarung 2011, Kaufman 2025), which have been described as having penultimate word-based stress, and Tboli, with regular word final stress (Forsberg 1992). Among the North

Luzon languages, loss of contrastive stress is one of the innovations that define the Southern Cordilleran subgroup (Himes 1998).

Kaufman (2025) argues that length is the phonemic element in what is described as contrastive stress in Philippine languages and that pitch accents serve merely to enhance the vowel length contrast in languages that maintain it.[17] Contrast enhancement explains why 'default stress' appears on the final syllable when the penultimate syllable lacks a long vowel but on the penultimate syllable when it contains such a vowel. Default stress only shifts to the penultimate syllable in languages that have lost contrastive length on the penultimate syllable, when it can no longer interfere with the perception of vowel length.

Many Philippine languages employ vowel length as a prosodic morpheme. Zorc (1977:64-67) discusses three types of morphological accent in the Bisayan languages which he takes to be part of the exponence of certain affixes. He notes, for instance, that in the Warayan subgroup of Bisayan, a prefix *ha-*, which derives adjectives indicating dimension and distance, co-occurs with penultimate stress. Thus, a root like *ra'yuʔ* 'distance' which shows final stress in isolation surfaces with penultimate stress with this prefix: *ha-'rayuʔ* 'far'. This apparent accent shift is likely due to the addition of vowel length to the penultimate syllable of the prefixed form (e.g. /ha-ra:yuʔ/). Other Bisayan affixes co-occur with final stress and Zorc terms these "ultima-accent affixes", for instance, the prefix *manog-* 'on the verge of'. When attaching to a stem with penultimate stress like *'tapus* 'finish', the derived form *ma,nog-ta'pus* has final stress. Finally, Zorc discusses affixes that appear to flip the stress of the stem with final stress stems taking penultimate stress and vice versa.

Nearly all Philippine languages that preserve contrastive vowel length show a process of rightwards length-shift with suffixation, as in (22a) and (22b), giving the impression of a metrical stress system. Arta provides an interesting exception to this pattern, as seen in (22c), where length remains on the root penult despite being suffixed.

	Tagalog		Ilokano		Arta
(22)a.	[basa:hin]	b.	[basa:en]	c.	[pe:nasan]
	/ba:sa-hin/		/ba:sa-en/		/pe:nas-an/
	read-PV		read-PV		wipe-LV

In Arta, certain roots that have no long vowels in their bare form take on a long vowel in their final syllable when suffixed. Kimoto (2017) analyzes this as in (23), where an underlyingly long vowel surfaces as short if it would create a superheavy syllable. Thus, a root like /adu:p/ can only express the vowel length when the final consonant is syllabified as the onset of a following consonant.

Arta (Kimoto 2017:114)

(23)a.	[adup]	b.	[adu:pan]
	/adu:p/		/adu:p-an/
	'help'		'help (LV)'

[17] An exception to this may be the Batanic languages which have been described as having inherited the vowel length contrast in the penultimate syllable but placing the accent on the final syllable regardless of whether the penult is long or short. Maree (2007:20) offers several minimal pairs of the following type, /gɨgɨt/ [gɨ'gɨt] 'house lizard' vs. /gɨ:gɨt/ [gɨ:'gɨt] 'to cut by sawing', in which only one member shows penultimate vowel length but both take final stress.

Similar to Tagalog and other Central Philippine languages, Newell and Poligon (1993) notes that Batad Ifugao has the potential for underlyingly long vowels restricted to open penultimate root syllables and that this length shifts to the right with suffixation (but not enclisis). Overall, this agreement between geographically disparate languages which have not been in contact historically strongly suggests the reconstruction of a similar pattern to the common ancestor of these Philippine subgroups, whether that is the putative Proto-Philippines or PMP.

As Zorc (1993) points out, length can be lost historically but recouped via various phonological processes. In Arta, which had lost the historical vowel length contrast (Kimoto 2017:70), vowel length was reintroduced via compensatory lengthening and coalescence of similar vowels, which gave rise to novel environments for long vowels, as seen in (24). In (24a), we find contrastive vowel length in word-final position due to historical coda loss and in (24b) we find vowel length derived from reduplication of V-initial roots and subsequent coalescence. Similarly, Estrera (2024) shows how Bagobo-Klata has maintained an old vowel length distinction while gaining a new length distinction from compensatory lengthening after word-final coda loss, as seen in (25).

Arta (Kimoto 2017:51,74)

(24)a. *anak > ana: ‘child’

b. *a~anak > a:na: ‘children’

Bagobo-Klata (Estrera 2024:63)

(25) *bu:lan > bu:la:

The prosody, stress and intonation of the North Luzon languages is highly understudied and the complex facts described in the literature beg for further analysis. Even Ilokano, the best studied of these languages, has a number of poorly understood conditions on length-shift (Rubino 1997:21-23; see also Ruffolo 2004:63-68 for Ibaloy). Furthermore, the morphological use of vowel length alterations in Philippine languages is still largely uncharted territory. Relatively little progress has been made in this area since Zorc (1977).

6 Morphology

6.1 Word classes and the lexicon

Overall, the morphology of most Philippine languages is highly complex along several dimensions. First, a large proportion of morphemes are multifunctional and take on distinct meanings in different morphological contexts. Second, much of the morphology is portmanteau, yielding a prototypical “fusional” language in Sapir’s (1921) typology. Third, the exponence of a morpheme, i.e., how a set of features are expressed on the surface, is often dependent on what other morphemes are present in the word. On a more surface level, Philippine languages make rich use of infixes, reduplication, circumfixes and prosodic morphology (especially vowel length alternations), which are less well attested outside this group.

It is crucial to understand the disconnect between root class and syntactic class before we proceed to a more detailed look at the morphology of event-denoting words. All previous works on Philippine languages have recognized a basic distinction between lexical and functional words (cf. Bloomfield’s 1917 ‘full word’ vs. ‘particle’ for Tagalog), the former providing the ‘real-world’ content of an expression and the latter acting to some extent as grammatical glue. However, the precise nature of lexical categories in Philippine languages has generated much

debate. What is clear is that the categorization of words in Philippine language differs significantly from familiar European languages. Predicates do not favor verbs but rather treat property, entity and event denoting words equally. Any of them may fulfill the predicate function when taking the generally clause-initial position of the predicate, as can be gleaned from the Tagalog examples in (26).

(26)	**PREDICATE**	**PIVOT**	
a.	*baba:ʔe*	*si liga:ya*	‘Ligaya is a woman.’
	woman	P.NOM Ligaya	
b.	*ma-taŋkad*	*si liga:ya*	‘Ligaya is tall.’
	ADJ-tall	P.NOM Ligaya	
c.	*na:sa labas*	*si liga:ya*	‘Ligaya is outside.’
	LOC.PRED outside	P.NOM Ligaya	
d.	*gutom*	*si liga:ya*	‘Ligaya is hungry.’
	hungry	P.NOM Ligaya	
e.	*nag-su:~sulat*	*si liga:ya*	‘Ligaya is writing.’
	AV.BEG-IMPRF~write	P.NOM Ligaya	

Similarly, arguments need not be headed by nouns or any other particular lexical category. Arguments are determined by case markers, which take any lexical category as their complement, as shown in (27).[18]

	Tagalog		
(27)a.	*aŋ baba:ʔe*	b. *aŋ ma-taŋkad*	c. *aŋ na:sa labas*
	NOM woman	NOM ADJ-tall	NOM LOC.PRED outside
	‘the woman’	‘the tall one’	‘the one outside’
d.	*aŋ gutom*	e. *aŋ nag-su:~sulat*	
	NOM hungry	NOM AV.BEG-IMPRF~write	
	‘the hungry one’	‘the one writing’	

Finally, there is no unmarked modifier, as seen in (28). All lexical categories may serve as modifiers through the mediation of the nasal linker (which in Tagalog, has two allomorphs: *=ŋ* after vowels, *n* and *ʔ* and *na* after all other consonants).

	Tagalog		
(28)a.	*ba:ta=ŋ baba:ʔe*	b. *ba:ta=ŋ ma-taŋkad*	c. *ba:ta=ŋ na:sa labas*
	child=LNK woman	child=LNK ADJ-tall	child=LNK LOC.PRED outside
	‘girl’	‘tall child’	‘child outside’
d.	*ba:ta=ŋ gutom*	e. *ba:ta=ŋ nag-su:~sulat*	
	child=LNK hungry	child=LNK AV.BEG-IMPRF~write	
	‘hungry child’	‘writing child’	

[18] It is very likely that the obligatory case marking determiners are responsible for the flexibility found in (27). Those languages south of the Philippine zone which have lost such determiners appear to have simultaneously developed dedicated relativizers employed obligatorily for turning non-nominals into arguments (Kaufman 2018a).

Himmelmann (1991, 2008), Foley (2008, 2023) and Kaufman (2009b, 2017) treat word-level categorization (or lack thereof) as a fundamental characteristic of Philippine morphosyntax that is tied to the aforementioned flexibility as well as to the nature of Philippine argument structure and the voice system. From a different perspective, the flexible distributions above could be due to the presence of null relativizers and copulas (Richards 2009a,b; Hsieh 2019). Neither approach implies that there is no categorization at all on the root level. Roots are most certainly categorized, but often in unusual ways. For example, one of the few rigid root-based categories in Tagalog divides property denoting roots into static and dynamic classes, a split which cannot be easily predicted on the basis of their meaning (Himmelmann 2008; Kaufman 2011a, 2012; Kaufman et al. 2026, Sabbagh 2005, 2011). An example of this can be seen with the two Tagalog roots /tuwaʔ/ and /saja/, which can both be translated as 'happy'[19] but whose morphological potentials are in near total complementary distribution with each other, as seen by the acceptability of the various formations in (29ii-v). The dynamic root gives rise to a predicate that is more verb-like, in requiring aspect inflection in a typical finite main clause. The stative root, on the other hand, while taking the same *ma-* prefix, does not allow it to be inflected for aspect.

	Tagalog	
(29)a. i.	/tuwaʔ/ 'happy' (DYN)	b. i. /saja/ 'happy' (STAT)
ii.	**ma-tuwa=ŋ baːtaʔ* STAT-happy=LNK child	ii. *ma-saja=ŋ baːtaʔ* STAT-happy=LNK child 'happy child'
iii.	*na-tuwaʔ* STAT.PRF-happy 'was happy'	iii. **na-saja* STAT.PRF-happy
iv.	*tuwa=ŋ~tuwaʔ* happy=LNK~INTNS 'very happy'	iv. **saja=ŋ~saja* happy=LNK~INTNS
v.	**naːpaka-tuwaʔ* INTNS-happy	v. *naːpaka-sajaʔ* INTNS-happy 'very happy'

Other languages have a special category for dimension adjectives, which take a descendant of the PMP *ha- prefix. This can be seen in Ilokano *a-baːbaw* 'shallow', *a-daːlem* 'deep' (Rubino 1997:126) as well as Naga Bikol *ha-rayoʔ* 'far'. Many languages have a class of bare property-denoting words that exist alongside adjectives prefixed with *ma-*, for instance, Ilokano *dakkəl* 'big' and Tagalog *pandak* 'short'. Cebuano has shed most instances of adjectival *ma-*, a pattern that is taken to its extreme in some languages south of the Philippine zone, such as Malay. Despite the diverse morphology of property-denoting words in Philippine languages, many languages show evidence for a distinct adjectival class that has a unique plural marking strategy, like CV reduplication as in Ilokano *na-la~laiŋ* (ADJ-PL~good), *da~dakkəl* (PL~big) (Rubino 1997:140) and Tagalog *ma-ga~ganda* (ADJ-PL~beauty), *pa~paːŋit* (PL~ugly).

The most commonly used morphemes found with event-denoting words (e.g. voice and aspect), property-denoting words (e.g. adjectival *ma-*) and entity-denoting words (e.g. number marking) do not make good diagnostics for root categories in Philippine languages, as they

[19] Alternatively, the dynamic root *tuwaʔ* may be best thought of as 'to become happy' alongside the stative root *saya* 'to be happy'.

typically take roots of any category. Rather, it is the peripheral morphology that is more selective (Kaufman et al. 2026). For instance, Tagalog *ma-* appears on property-denoting roots to form words such as (30a), but *ma-* can attach to any entity-denoting root to form a property-denoting word, as seen in (31a). However, it is only the first root, which patterns with other property-denoting roots, in allowing the intensive prefix, as seen in (30b) versus (31b).

	Tagalog		
(30)a.	ma-ya:man	b.	na:paka-ya:man
	ADJ-wealth		INTNS-wealth
	‘wealthy’		‘very wealthy’

(31)a.	ma-pe:ra	b.	*na:paka-pe:ra
	ADJ-money		INTNS-money
	‘moneyed/rich’		

Space precludes a fuller account of lexical categories in Philippine languages but it is worth taking a brief look at the morphological potential of numerals, which is unusually rich. In addition to being able to form verbs with the help of voice morphology, numerals are associated with a unique set of morphological processes, only some of which are shown in (32).[20]

(32)		Ilokano	Tagalog	gloss
a.	root	*dua* ‘two’	*dalawa*	‘two’
b.	distributive	*sag-du~dua*	*tag-dalawa*	‘two each’
c.	ordinal	*maika-dua*	*ikalawa/pangalawa*	‘second’
d.	X-times	*mamindua*	*makalawa*	‘twice’
e.	restrictive	*du~dua*	*da~dalawa*	‘only two’
f.	approximate	*s<um>ag-du~dua*	*maŋa dalawa*	‘around two’

On the level of the root, numerals clearly form a class, as they license a number of morphological processes that are not found with other root types. But just as with other root categories, numerals do not give rise to a unique syntax. They form unmarked predicates and can function as modifiers or arguments in the same as the lexical words seen above. With this much background, we now turn to functions commonly expressed as bound morphology on event-denoting predicates in Philippine languages.

6.1 Aspect

Although often described in terms of tense in the literature (e.g. McKaughan 1958, Wolff 1973, Zorc 1977 *inter alia*) the temporal inflections of Philippine languages uniformly indicate aspect rather than tense, with the possible exception of Iraya (Reid 2017). Voice and aspect are grammatically prominent and paradigmatically interconnected in most Philippine languages (cf. Reid 1992; Ross 2002; Himmelmann 2005). This can be seen in the Tagalog voice/aspect paradigm shown in Table 2, where the voice marker disappears unexpectedly in the prospective aspect of the actor voice paradigm and in the perfective of the patient voice paradigm.

[20] Note that almost all of the morphs in (31) (e.g. *tag-, i-, ka-, <um>, ma-*) are familiar from other areas of the grammar. This overlap is too frequent to be a coincidence but at the same time it is difficult if not impossible to unite the numeral functions of these morphs with their non-numeral functions.

Table 2. Fragment of the Tagalog voice aspect paradigm

baːsag 'break'	**Actor** *<um>*	**Patient** *-in*	**Locative** *-an*	**Conveyance** *i-*
Neutral	*b**<um>**aːsag*	*basaːg**-in***	*basaːg**-an***	***i-**baːsag*
Perfective	*b**<um>**aːsag*	*b<in>aːsag*	*b<in>aːsag**-an***	***i-**b<in>aːsag*
Progressive	*b**<um>**aːˀbaːsag*	*b<in>aːˀbaːsag*	*b<in>aːˀbaːsag**-an***	***i-**b<in>aːˀbaːsag*
Prospective	*baːˀbaːsag*	*baːˀbasaːg**-in***	*baːˀbasaːg**-an***	***i-**baːˀbaːsag*

A subset of Central Philippine languages display three primary aspects which can be termed perfective, progressive and prospective.[21] The three way distinction may arise from two atomic features corresponding to reflexes of *<in> and *CV reduplication, as in (33).

(33)

atomic features		compositional meanings	
<in>	BEGUN	<in>	**perfective**
CV~	IMPERFECTIVE	<in>CV~	**progressive**
		CV~	**prospective**

The feature combination [+begun, –imperfective] is interpreted as perfective, [+begun, +imperfective] as progressive, and [–begun, +imperfective] as prospective. Thus, while none of the surface aspects are indicated uniquely by a morpheme, they are derived in a compositional manner (see Otanes 1966, De Guzman 1978 and Reid 1992 for different feature based approaches to this paradigm). Aspectual CV-reduplication may have originally marked the imperfective or durative in PMP while *<in> appears to have marked the perfective (Wolff 1973, Zorc 1977, Reid 1992, Ross 1995, 2002). Reid (1992) argues that *<in> innovatively spreads into the progressive in Central Philippine languages, where it comes to signal [+begun]. In languages outside of this subgroup, the use of *<in> reflexes are rarely seen in the progressive.

Aspect marking is most often obligatorily on finite verbs although in some languages, such as Cebuano, Agutaynen and Ilokano, a single form will be used for the imperfective/prospective and the infinitive thus yielding a two-way distinction. Such languages can be said to conflate the historical unmarked and prospective aspects into a general 'unrealized' inflection (Reid 1992:74).

In addition to the major aspects shown in the above tables, most languages also possess minor aspects like Tagalog's recent perfective and immediate prospective, shown in (34).

Tagalog

(34)a. *ka-raːˀ~ratiŋ ko laŋ*
RCT_1-RCT_2~arrive 1SG.GEN only

b. *pa-ratiŋ na ako*
IMMD-arrive already 1SG.NOM

[21] The prospective, which is used for unbegun action, is also referred to as "contemplated", "future" and "irrealis", all of which are, strictly speaking, inappropriate labels. "Contemplated" suggests cognition on the part of an agent; "future" designates a tense rather than an aspect; "irrealis" suggests that the form would be obligatory in negated and counterfactual contexts, although this is not the case.

‘I just arrived.’ ‘I’m about to arrive.’

The recent perfective and immediate prospective are minor both in their frequency and in their emphatic interpretation, in contrast to the basic aspect categories. The syntax of the recent perfective is also distinct from the major aspects. In Tagalog and most other Central Philippine languages, voice is neutralized and genitive case is assigned to what would normally be the nominative case marked argument. The recent perfective cannot be negated and may show additional syntactic restrictions, as well.

In many languages throughout the Philippines, the progressive or imperfective is signaled by a reflex of PMP **paR*- (cf. Zorc 1977 for Aklanon and Kerr 1988:8 for Cotabato Manobo). This can be seen in the Arta actor voice and patient voice verbs in (35) and (36), respectively, where a reflex of *<in> indicates the perfective in (35b) and (36b) and the prefixation of paC- indicates the imperfective in (35c) and (36c).

Arta (Kimoto 2017:328)

(35)a.	*man-di:muy*	b.	*nan-di:muy*	c.	*pad-di:muy*
	AV-bathe		PERF.AV-bathe		PROG-bathe
	‘take a bath’		‘took a bath’		‘be taking a bath’
(36)a.	*bisag-en*	b.	*b<in>isag*	c.	*pab-bisag-en*
	break-PV		<PERF>break		PROG-break-PV
	‘break’		‘broke’		‘be breaking’

The combination of **<in>* with the composite actor voice markers beginning with *m*- (i.e. PMP **maŋ*- AV.DIST, **maki*- AV.SOC, **maR*- AV.MID, **maka*- AV.POT) typically yields *n*- initial forms without infixation (e.g. *naŋ*-, *naki*-, *nag*-, *naka*-). This “externalization” of *<in> postdates PMP, as we also find languages throughout the Philippines that reflect disyllabic forms for **m<in>aR*- as well as languages that display reductions reflecting **miR*- rather than **naR*-, showing that the full historical form was reduced in diverse ways after the break-up of the major Philippine subgroups.

In negated clauses, aspect is often indicated by the choice of negator and the verb is left unmarked or marked with an aspect neutral inflection. An example of this is seen in Sarangani Manobo, where aspect is marked on the verb in (37a-b) but through negation in (37c-d). Similar examples could be produced for most Bisayan languages, as well.

(37) Sarangani Manobo (Dubois 1976:20)

a.	*t<om>edogi se bayi*	b.	*t<im>edogi se bayi*
	<AV>sleep NOM woman		<AV.PFV>sleep NOM woman
	‘The woman will sleep.’		‘The woman slept.’
c.	*edek tedogi se bayi*	d.	*weda? tedogi se bayi*
	NEG sleep NOM woman		NEG.EXT sleep NOM woman
	‘The woman will not go to sleep.’		‘The woman didn’t sleep.’

A more holistic understanding of aspect in Philippine-type languages must take into account both “inner aspect”, i.e. perfective, progressive, prospective, as marked with bound morphology, together with “outer aspect”, the relation of event time to expected reference time,

as marked by enclitics, typically descendants of PMP *=*dena* ‘already’ (almost always reduced to a monosyllable) and *=*pa* ‘still’ (Kaufman 2011b).[22] The expression of inner aspect and outer aspect are completely independent from each other and can occur in any combination. However, as far as has been documented, most Philippine languages appear to treat reflexes of PMP *=*dena* and *=*pa* ‘still’ as a pair in opposition and thus incapable of co-occurring on the same predicate. The introduction of an additional functor may allow for such cooccurrence by separating the clitics, as in (38).

Tagalog
(38) *hindi?=**na**=ako* *mag-ta:~tanoŋ=**pa***
NEG=ALREADY=1S.NOM AV-IPFV~ask=STILL
‘I won’t ask any more.’

Only in rare cases do these clitics cooccur on the same predicate, as exemplified in (39), where Ilokano =*en* reflects PMP *=dena and =*pay* descends from PMP *=pa.

Ilokano (Rubino 1997:226)
(39) *ag-tak~takaw=ka**=pay=en*** *tapno adda i-pa-lamut=mo kenkuana*
AV-PROG~steal=2S.NOM=still=now so EXT CV-CAU-food=2S.GEN 3S.OBL
‘You are stealing so you can feed her?’

6.2 Voice

A conservatively rich voice system is a defining feature of the morphosyntax of all Philippine-type languages.[23] This system is generally understood to select a particular clausal participant as the nominative argument (or absolutive, depending on the analyst) using one of several voice morphemes. This argument is typically interpreted definitely and can stand alone without an associated predicate. It is in some sense a privileged argument but its cross-linguistic status vis a vis subject and topic remains debated after several decades of investigation (Schachter 1976, Kroeger 1993, Richards 2000, Rackowski 2002, Schachter 1996, Aldridge 2004, Chen 2017, Erlewine et al. 2017, *inter alia*).

Agents of non-actor voice verbs are uniformly expressed in the genitive case in Philippine languages. Undergoer participants, when not selected by the voice morphology to become nominative arguments, are generally expressed either as genitives (as in Tagalog) or as obliques (as in Cebuano). Some Batanic languages appear to possess a dedicated object/accusative marking although this is very rare. Other languages may introduce indefinite actor voice objects with the linker that typically signals modification, as in the Naga Bikol and Kapampangan examples in (40a) and (41a).

[22] There also exist Northern Luzon languages which employ clitics for tense marking, e.g. Ilokano =nto FUTURE.
[23] On the Bornean side, Lobel (2013:150) locates the southern border of the full voice system in the area of “Brunei Dusun, Kolod, Tingalan, Abai Sembuak/Tubu, Bulusu, and Tidung languages, although a handful of non-Philippine-type languages exist north of this hypothetical line.” In Sulawesi, the full voice system seems to be continued only in the Mongondow-Gorontalo (or “Gorontalic”) languages, as well as the Minahasan and Sangiric subgroups.

The four primary voices are the actor voice, patient voice, locative voice and conveyance voice, shown in Table 3.[24]

Table 3. Common reflexes of PMP voice markers

Voice	**PMP reconstruction**	**Common reflexes**
actor voice	*<um>	<um>, m-, mu-
patient voice	*-en	-in, -un, -ɨn, -en, -n
locative voice	*-an	-an, -n
conveyance voice	*(h)i-[25]	ʔi-, hi-, ∅

The basic use of the voice markers is very consistent with few exceptions. The system can be illustrated with the Naga Bikol examples in (25).

(40) Naga Bikol (McFarland 1974:104-105)

a. *nag-bakal aku=ŋ bagas*
AV.BEG-buy 1SG.NOM=LNK rice
'I bought rice.'

b. *b<in>akal-∅ ko an bagas*
<BEG>buy-PV 1SG.GEN NOM rice
'I bought the rice.'

c. *b<in>akal-an ko si hwan ki bagas*
<BEG>buy-LV 1SG.GEN P.NOM Juan OBL rice
'I bought some rice from Juan.'

d. *i-b<in>akal ko si hwan ki bagas*
CV-<BEG>buy 1SG.GEN P.NOM Juan OBL rice
'I bought some rice for Juan.'

Kapampangan (Roswell 1983:45)

(41)a. *Mān-agal=ya-ng pusa ing=asu*
AV.IPFV-chase=3s.NOM-LNK cat NOM=dog
'The dog chased a cat.'

b. *Ta~tagal-an=ne ning=asu ing=pusa.*
IPFV-chase-LV=3s.GEN>3s.NOM GEN=dog NOM=cat
'The dog chases the cat.'

As can be seen, one participant is selected by the predicate to be the nominative argument while other participants are expressed in non-nominative cases. The actor voice selects a proto-agent as

[24] What is termed here conveyance voice, following Wolff (1973), goes by several other names as well: circumstantial, instrumental, benefactive, secondary object, and theme voice, among others. See Blust (2002) and Ross (2002) for a review of the terminology and its history.

[25] The PMP cognate of the PAn conveyance voice marker **Si-* is predicted to be **hi-*, but this form only surfaces as such in Tausug and Samareño. Everywhere else, the initial *h* seems to have been eliminated in favor of a (possibly epenthetic) glottal stop. Nonetheless, because *h* is expected and these two languages were not in close contact with each other, the more common form *ʔi-* is thought to have come about through parallel innovation.

the nominative argument; the patient voice selects a proto-patient; the locative voice selects a locative, directional or other type of oblique argument as well as an unaffected object; the conveyance voice selects a theme moving away from the agent as well as an instrumental or benefactee as the nominative argument.

The proper treatment of these voice markers remains an area of copious theorization and major debate in Austronesian linguistics. The earliest published analyses carried out by Spanish linguists and inherited by Bloomfield (1917) treated the patient, locative and conveyance voices as types of passive (e.g. direct passive, locative passive, etc.). It was recognized from the earliest point, however, that the putative "passives" of Philippine languages, which are fully transitive, were not equivalent to the Indo-European passive, a marked detransitive construction used primarily to background the agent. In the symmetrical analysis of Philippine-type voice (Foley 2008; Himmelmann 2005; Riesberg 2014; Chen 2017), the system represents a unique type of alignment where all voices are equally marked, standing in natural opposition to accusative and ergative languages which typically display unmarked transitive and intransitive clauses. For the vast majority of Philippine languages, it also holds true that there is no morphologically unmarked voice, just as in the Tagalog paradigm seen earlier. Proponents of an ergative analysis of the Philippine voice system (Starosta et al. 1982, De Guzman 1988, Gerdts 1988, Brainard 1997, Aldridge 2004, Liao 2004) argue that the actor voice is less transitive than its non-actor voice counterparts. Although the arguments for one analysis over another cannot be reviewed here (but see Kaufman 2017), the principles of voice selection require basic explication.

There is widespread agreement that some type of referentiality largely determines voice selection in all Philippine languages (see Wolfenden 1961, Wolff 1966, Schachter 1976, McFarland 1978, for early modern treatments). Table 4, based on Tagalog but applicable more widely, abstracts away from many complications, additional factors, and cross-linguistic variation (cf. Schachter 1976, Naylor 1986, Adams and Manaster-Ramer 1988, McFarland 1978, Latrouite 2011, Nolasco 2003) but captures what I believe to be the core basis for the alternation. When the agent is definite and the theme/patient is indefinite or absent, the predication will be expressed in the actor voice. When the theme/patient is definite, there is a strong tendency to employ the patient voice, regardless of the definiteness of the agent. With a verb of transfer and similar predicates, when the theme is indefinite but the recipient is definite, the locative voice will be selected. When a conveyed theme is definite, the conveyance voice will be selected, regardless of the definiteness of the agent and recipient.

Table 4. Voice selection in a typical Philippine language

Agent	**Theme/Patient**	**Locative**	**Preferred Voice**
def	(indef)	–	actor voice
def/indef	def	–	patient voice
def/indef	(indef)	def	locative voice
def/indef	def	def	conveyance voice

Genitive agents and nominative arguments may always receive a definite interpretation, while directional arguments are felicitously expressed in the oblique case regardless of their definiteness. What the pattern in Table 4 conspires to avoid is the expression of a definite

undergoer as a non-nominative object.[26] If a previously introduced or otherwise familiar argument does surface as a non-nominative object, it typically receives a partitive interpretation or is understood to be less affected by the action (Nolasco 2003).

In an intransitive predication with an indefinite subject, the subject is typically introduced with the use of an existential, as shown in (42a) (Schachter and Otanes 1972:279, but see Adams and Manaster-Ramer 1988, Bell 1978 and Kaufman 2026a for additional wrinkles). The same holds for a bivalent predication in which neither argument has been previously introduced, as seen in (42b). This strategy is necessary to avoid the ordinarily definite interpretation of the nominative phrase.

Tagalog

(42)a. *may d<um>atiŋ*
EXT <AV.BEG>arrive
'Someone arrived.'

b. *may k<um>a:ʔin naŋ sa:giŋ*
EXT <AV.BEG>eat GEN banana
'Someone ate a banana.'

The patient voice is not restricted to semantically bivalent predicates. Examples of the type in Tagalog (43) show that patient voice also selects affected subjects of monadic and even entity denoting predicates (Ceña 1977).

Tagalog

(43)a. *la:~laŋgam-in aŋ asu:kal*
IPFV~ant-PV NOM sugar
'The sugar will be "anted".'

b. *s<in>i:~sipon-Ø ako*
<BEG>IPFV~flu-PV 1SG.NOM
'I have the flu.' ('I'm being "flued".')

Similarly, the locative voice can select a recipient or location that we would consider part of the lexical semantics of the verb, as in (44), but it can just as easily "promote" an adjunct to become the nominative argument, as in (45).

Tagalog

(44) *b<in>igy-an ni=rori naŋ=pe:ra si=pe:peŋ*
<BEG>give-LV GEN=Rory GEN=money NOM=Pepeng
'Rory gave Pepeng money.'

(45) *in-iyak-an ni=rori si=pe:peŋ*
BEG-cry-LV GEN=Rory NOM=Pepeng
'Rory cried to Pepeng.'

The locative voice can also alternate with the patient voice to indicate that the nominative argument is less affected by the action than would normally be assumed, as seen in the minimal pair in (46).

[26] This pattern holds throughout most Philippine languages but some languages, such as Cebuano (Bell 1978), are argued to have a more flexible correspondence between the syntactic status of an argument and its definiteness. Reid (2020) and Kaufman (2026a) discuss ways in which Philippine languages have accommodated a more Indo-European style of definiteness marking under Spanish and English influence.

Tagalog

(46) a. *k<in>a:ʔin-Ø ni=maria aŋ=isdaʔ*
<BEG>eat-PV GEN=Maria NOM=fish
'Maria ate the fish.'

b. *k<in>aʔi:n-an ni=maria aŋ=isdaʔ*
<BEG>eat-LV GEN=Maria NOM=fish
'Maria ate from/at the fish.'

On the other hand, with inherently oblique/directional arguments, the difference between actor voice and locative voice can be more subtle and does not relate directly to the definiteness of the participants in many languages. We can compare the two Tagalog examples in (47), which correspond to the same English translation, but which differ in the affectedness of the goal, Helen, and the accomplishment of the action. With the actor voice in (47a), the action of returning may have been initiated but not yet completed whereas the preferred interpretation of (47b) is that the return is complete. In (47a), we are only told that the agent made a return to Helen but in (47b) there is an implication that they got back together, perhaps as a couple.[27]

Tagalog

(47)a. *b<um>alik=siya kay=helen*
<AV.BEG>return=3S.NOM P.OBL=Helen
'He returned to Helen.'

b. *b<in>alik-an=niya si=helen*
<BEG>return-LV=3S.GEN P.NOM=Helen
'He returned to Helen.'

The conveyance voice (PAn *Si-) selects benefactees, instrumentals and objects conveyed away from the agent as the nominative argument, and is thus difficult to characterize in a semantically unified manner. These seemingly disparate functions can be disambiguated in languages that combine a reflex of *Si- with other affixes. For instance, alongside the ambiguous *i-* voice marker, Tagalog also has *i-pag-* BENEFACTIVE and *i-paŋ-* INSTRUMENTAL. In Northern Luzon languages, the benefactive function is disambiguated with a combination of the historical conveyance voice and locative voice (*-an), as in Ilokano (48), where it is analyzed as a benefactive voice circumfix (Reid & Liao 2002:460).

Ilokano (Rubino 1997:240)

(48) *i<in>-gatang-an=na=kayo iti=kalding*
BV_1-PRF-buy-BV_2=3S.GEN=2P.NOM OBL=kalding
'She bought the goat for you.'

In addition to the indicative/independent voice forms, there also exists a non-indicative/dependent paradigm. Wolff (1973:88) reconstructs this paradigm for the imperative and after certain "preverbs" while later work by Ross (2002) reconstructs it with a slightly wider range of functions. In most MP languages outside the Philippines, the distinction between the indicative and non-indicative forms have merged. In the northern and central Philippines, the paradigms are generally merged in favor of the indicative paradigm and are

[27] As Nolasco (2003) points out, these differences align well with Hopper and Thompson's (1980) transitivity parameters, indicating that the non-actor voices are more transitive than the actor voice.

reduced in various ways south of the Philippines.[28] Wherever the non-indicative paradigm is preserved, it is at least used in the imperative. This is seen in (49)-(52). The dependent paradigm imperatives are distinguished from independent paradigm imperatives in most languages by the obligatory omission of a second person singular addressee pronoun, as in Batangas Tagalog, although there are exceptions to this, as in Maranao, Keley-i and Ibaloy.

Batangas Tagalog

(49) *buks-i (*mo) aŋ pintuʔan*
open-LV.DEP 2SG.GEN. NOM door
'Open the door!'

Maranao (McKaughan 1958:25)

(50) *tabas-a ŋka so dinis*
cut-PV.DEP 2SG.GEN NOM cloth
'Cut the cloth!'

Keley-i (Hohulin et al. 2018)

(51) *ekal-i=m tep me-belʔat*
remove-LV.DEP=2s.GEN because ADJ-heavy
'Remove some because it is heavy!'

Ibaloy (Ruffolo 2004:426)

(52)a. *karamak kalti*
kada=m=ak kalat-**i**
PRHB=2s.GEN=1s.NOM bite-LV.DEP
'Do not bite me.'

b. *karajo kena*
kada=jo kan-**a**
PRHB=2p.GEN eat-PV.DEP
'Do not eat (it).'

In many Central Philippine languages, the dependent paradigm is also used in the negated perfective, as shown by Wolff (1973) for Samareño (53). This paradigm does not co-occur with imperfective reduplication or the perfective/begun *<in> infix (although they can occur in the recent perfective, see below).

Samareño (Waray) (Wolff 1973)

(53)a. *waraʔ lakaw-Ø a baːtaʔ*
NEG.EXT go.away-AV.DEP NOM child
'The child did not go away.'

b. *waraʔ ku balik-a a sibiːsa*
NEG.EXT 1SG.GEN return-PV.DEP NOM beer
'I did not go back after the beer.'

c. *waraʔ ku hiŋalimt-i a isturya*
NEG.EXT 1SG.GEN forget-LV.DEP NOM story
'I did not forget the story.'

[28] In the majority of Austronesian languages, the independent locative voice *-an* survives with a nominalizer function and some remnant of *<um>* (typically melded with one of the mode prefixes as *m-*) survives in the actor voice. On the other hand, patient voice **-en* and conveyance voice **Si-* are widely lost as productive voice markers south of the Philippine languages, although the distinction may be carried out through different morphological means.

d. *waraʔ niya pilak-an an basuːra*
NEG.EXT 3SG.GEN throw.away-CV.DEP NOM garbage
'He did not throw the garbage away.'

The dependent forms are also employed in temporal adjuncts, as seen in (54) and (55) (Stevens 1969, Zorc 1977) and the recent perfective (not shown here). These contexts are particularly interesting as the voice morphology selects a particular argument to promote, *kanya suwildu* in (54) and *baŋku* in (55), although no argument surfaces with nominative case. Rather, the argument selected by the voice morphology surfaces with oblique or genitive case, as it would in the actor voice (Kaufman 2026c).

Samareño (Waray) (Zorc 1977:139)

(54) *pag-ta-tág-an=niya [sa kanya suwildu] [kanya nanay]...*
SBJV-ASP-give-CV.DEP=3SG.GEN OBL 3SG.GEN earning 3SG.GEN mother
'When he gives all of his earnings to his mother…'

(55) *pag-liŋkur-i=niya han baŋku, na-rubaʔ*
SBJV-sit-LV.DEP=3SG.GEN GEN.DEF bench STA.PFV-break
'When he sat on the bench, it broke.'

The Tboli voice system has been reshaped by the general loss of suffixes and case marking on full noun phrases. Here, there exists a general actor voice marked by *me-/* and a general undergoer voice marked by *ne-/<en>*, while the conveyance voice is left unmarked morphologically but still considered distinct. Tboli agent voice, undergoer voice and instrumental voice clauses are exemplified in (56), respectively.

Tboli (Awed et al. 2004:79, 25)

(56)a. *s<m>akay=le owoŋ yo ken ŋaʔ*
<AV>ride=3PL airplane that PL child
'The children rode in the airplane.'

b. *gel n-boʔ maʔ ɔu*
always UV-carry_on_back Father me
'Father always carried me on his back.'

c. *Ø-ɔfɔk Walan du asay*
CV-chop_down Walan it axe
'Walan chopped it down with an axe.'

Composite minor voices, which appear to have been innovated more recently, are formed from combinations of inherited morphemes, and select adjuncts such as purposive clauses as the nominative argument. The Tagalog prefix *ika-* (< PMP **(h)i-* CONVEYANCE VOICE + **ka-* STATIVE) and its cognate Sarangani Manobo exemplify this in (57) and (58) (Kaufman 2024c).

Tagalog

(57) *ano aŋ ik<in>a-puː~punta niya duʔun?*
what NOM <BEG>REAS-IPFV~go 3SG.GEN there
'What's his reason for going there?'

Sarangani Manobo (Dubois 1976:67)

(58) *yan se iŋke-opal ko*
that NOM REAS-anger 1SG.GEN
'That's why I became angry.'

Philippine languages typically allow only one voice marker per word, but we have already seen how Northern Luzon languages employ a combination of conveyance and locative voice to form the benefactive voice. Note however that this morphological combination does not promote multiple arguments nor does it promote a single argument twice (for instance, from oblique to object and from object to nominative argument). The combination functions for all intents and purposes as a single morpheme.

Such cases are markedly different from combinations of actor voice and conveyance voice in Northern Luzon languages, e.g. *maŋ-i-* (AV-CV-), in which both the actor voice and the conveyance voice do play an independent voice related role, the first determining the voice of the entire predicate and the second functioning as a true object applicative with the same semantic range as the conveyance voice (instrumental, conveyed object, benefactive). In (59a), a plain actor voice verb in Ibaloy selects a genitive theme. In (59b), with the addition of the conveyance voice/applicative, an instrument has now been promoted to object.

Ibaloy

(59)a. *memonas ni damisaan*
məŋ-ponas ni lamisaʔan
AV-CV-wipe GEN table
'she will wipe a table'

b. *mengiponas ni dopot*
məŋ-**i**-ponas ni lopot
AV-CV-wipe GEN cloth
'she will use a piece of cloth to wipe'

A similar pattern is found in most other Northern Luzon languages. It is not always clear in these languages that an applicative has 'promoted' an oblique to object position. In many cases, certain predicates select their canonical complements with the conveyance voice, especially when it is an unaffected object/theme. For instance, in Ilokano, as in many languages, the theme of a verb like 'write' is selected with the conveyance voice. Unlike languages outside this region, Northern Luzon languages may retain the conveyance voice to select a non-affected or centrifugal theme even in the actor voice, as in *ag-i-surat* (AV-CV-write), which still selects the agent as the nominative argument (Rubino 1997:197).

Another type of double voice marking is found among certain North Luzon languages in nominalizations like (60) and (61). Here, an actor voice infix co-occurs with the locative voice suffix to form an event nominalization, which appear to be required in 'when' questions, among other contexts.

Kankana'ey (Allen 2014:166)

(60) *Pigʔan di s**<om>**aa**-an**=da?*
when NOM <AV>do-**LV**=3p.GEN
'Their going home (is) when?'

(61) Kalinga (Gieser 1963:22)
*nan l**<um>**ayug**-an**=na*
NOM <AV>jump-LV=3s.GEN
'where it jumps down'

Overall, multiple voice marking in Philippine languages has not been studied from a comparative perspective so little more can be said here.

6.3 Mode

There are several common verbal morphemes in Philippine languages that are often treated under the vague header of "mode", a questionable practice I continue here. Under this header, I cover the potentive (§6.3.1), the distributive (§6.3.2), the sociative (§6.3.3), agent plurality (§6.3.4), the multifunctional *paR-/maR- prefixes (§6.3.5), reciprocal marking (§6.3.6) and the inchoative (§6.3.7).

6.3.1 Potentive

Nearly all Philippine languages have a potentive paradigm, which is used to indicate possible, unintentional action, as well as action that is out of the actor's control (Baradjí et al. 2022). This polysemy, which is remarkably stable across Austronesian languages, can be seen in the Tboli sentences in (62) and (63) with the *g(e)-* prefix, a reflex of PMP **ka-* (cf. Bennásar 1892:38-39 for the Tiruray cognate).

Tboli
(62) *nə g-tutuk kulu nib*
and POT-nail head Nib
'And Nib accidentally bumped his head.'

Tboli (Forsberg 1992:92)
(63) *g-uŋɔl-u udɛl sdoʔ fatu ləm law*
POT-hear-1SG.GEN voice pig across in cane
'I was able to hear the squeal of a pig in the cane across (the river).'

The potentive does not simply provide a way of emphasizing the accidental or unintentional nature of an action; it is obligatory in contexts where control is feasible or "at issue" (to use Baradjí et al.'s 2022 term) and as a corollary, the unmarked (non-potentive) form unambiguously denotes intentional action by an animate agent in these cases. This paradigm, which is contrasted with the unmarked "dynamic" voice paradigm, has a very distinct history involving the PAn prefix **ka-*, which appears to have been an existential marker (Kaufman 2011a, 2012). The four basic voices in Tagalog are shown in their unmarked dynamic form alongside their potentive counterparts in Table 5.

Table 5. The Tagalog potentive paradigm

	dynamic	**potentive**
ACTOR VOICE	*<um>*	*maka-*
PATIENT VOICE	*-in*	*ma-*
CONVEYANCE VOICE	*i-*	*ma-i-*
LOCATIVE VOICE	*-an*	*ma- -an*

The Tagalog potentive is transparently derived from the basic voice paradigm in the conveyance and locative voices with the addition of *ma-* but the actor and patient voices do not show clear correspondences. The potentive patient voice does not include a reflex of patient voice *-*en* and the potentive actor voice is not obviously related to other forms in the paradigm. This somewhat confusing picture, typical for Central Philippine languages, has a straightforward historical explanation. The **ma-* prefix was originally a reduction of stative **ka-* combined with actor voice **<um>*, as a general *non-actor voice* potentive (Ross 1995:741). Historically, there was an opposition between an active clause such as (64a) and a passive-like stative clause, as in (64b), where the logical object would be the nominative argument. The latter is derived with the stative prefix *ka-* combined with the actor voice *<um>* followed by apheresis of the first syllable.

		Tagalog		
(64)	a.	*ʔ<um>u:bos*	b.	*k<um>a-ʔu:bos* → *ma-ʔu:bos*
		<AV>finish		<AV>STA-finish
		'to finish'		'to get finished'

One type of formerly actor voice **ma-* forms were reinterpreted as transitive in many Philippine subgroups, including the Central Philippine languages. In these languages, any root that can take the patient voice can also take the reflex of **ma-* as its potentive counterpart. Just as the patient voice licenses genitive marked agents, so does potentive **ma-*, yielding an opposition like the one illustrated in (65).

	Tagalog				
(65)a.	*na-ʔu:bos*	*ni*	*bo:boy*	*aŋ*	*pagka:ʔin*
	STA.BEG-finish	GEN	Boboy	NOM	food
	'Boboy finished the food (accidentally)'				
b.	*<in>u:bos-Ø*	*ni*	*Bo:boy*	*aŋ*	*pag-ka:ʔin*
	<BEG>-finish-PV	GEN	Boboy	NOM	food
	'Boboy finished the food (purposefully)'				

The use of the genitive agent in (65a) must have been an innovation, although it remains unclear at what stage it took place, and if it spread horizontally across languages of different subgroups. In many Northern Luzon languages, the potentive paradigm with **ma-* functions more like a true passive than a transitive clause (cf. Reid & Liao 2004:464) and its use has a true potentive meaning even when forming verbs of perception. In (66a), with a genitive agent, the clause is interpreted as a happenstance occurrence. Without an agent, as in (66b), it is interpreted as a stative.

	Ibaloy (Ruffolo 2004:306)				
(66)a.	*naon'an*	*ni*	*dedaki*	*sota*	*bibiid* *Batan*
	na-ʔonəj-an	ni	CV-laki	sota	CV-biʔi=d batan
	POT-see-LV	GEN	PL-man	NOM	PL-woman=LOC Batan
	'the men happen to see the women of Batan'				
b.	*naon'an*	*i*	*bawaja*	*chi*	*Crocodile Farm*
	na-ʔonəj-an	ʔi	bawaja	di	crocodile farm
	POT-see-LV	NOM	crocodile	LOC	Crocodile Farm
	'the crocodiles were visible at the Crocodile Farm'				

In such languages, perception verbs typically take the patient voice in their normal use, as exemplified in (67).

Tuwali Ifugao (Hohulin and Burquest 2014)
(67) *Nganne tut-uwa-y in-angʔang-∅=mu?*
what truly-LNK PFV-see-PV=2S.GEN
'Now, what did you really see?'

In contrast, Central Philippine languages typically employ the potentive paradigm perception verbs. In some cases, the patient voice of a perception verb will require the potentive form, for example, Tagalog *ma-ki:taʔ* (POT-see) versus the ungrammatical **kita:ʔ-in* (see-PV), while in other cases, the difference signals involuntary perception in contrast to the action of trying to perceive, as with *ma-amuy* (POT-smell) 'to happen to smell' versus *amuy-in* (smell-PV) 'to sniff/smell something (voluntarily)'.

It is likely that prior to the reanalysis of **ma-* forms as transitive, actor arguments could only be expressed as obliques, as this strategy still exists as an option in a number of languages, like Ilokano (68).

Ilokano (Rubino 1997:34)
(68) *na-duaya=ka iti siŋgit dagiti ag-si~s<inn>uŋbat a kundidit*
PRF-lullaby=2S.NOM OBL shrill PL AV-PL~<RECP>answer LNK cicada
'You were lullabied by the shrills of the conversing cicada.'

The reanalysis of **ma-* from its original actor voice stative function to a potentive undergoer voice marker goes hand in hand with its appearance in other voices. The spread of **ma-* can be seen clearly in the comparison between Toratán (a Sangiric language of North Sulawesi, Himmelmann and Wolff 1999), Bikol Naga, and Tagalog (both Central Philippine), shown in Table 6. Note that *ma-* is labelled as a patient potentive in Table 6 due to sharing a case frame with the patient voice in many Philippine languages, as seen above in (48), but it is historically an actor voice form and is considered intransitive by Reid and Liao (2002:462).

Table 6. Potentive paradigms for three Philippine-type languages

	Toratán	**Bikol Naga**	**Tagalog**
Actor Voice	*maka-*	*maka-*	*maka-*
Patient Voice	***ma-***	***ma-***	***ma-***
Locative Voice	*ka- -an*	***ma-*** *-an*	***ma-*** *-an*
Conveyance Voice	*ka-*	*i-ka-*	***ma-****i-*

Toratán shows the most conservative paradigm, with *ka-* still used in both the conveyance and locative voices. It is innovative in having lost the *i-* in the potentive conveyance voice, but this is a recurring change seen to take place in Mindanao, as well. The *ma-* prefix has spread to the

locative in Bikol Naga and additionally to the conveyance voice in Tagalog and, seemingly, the majority of Philippine languages including those of the north.[29]

The other oddity of the potentive paradigm is the actor voice counterpart to *ma-*, namely, *maka-*, which is derived from the combination of PMP *<*um*> with the PMP causative **pa-* and the stative **ka-*. The original opposition between today's patient and actor voice potentive was thus not one of voice at all but rather one of causation.[30]

6.3.2 Distributive

Many Philippine languages express a distributive or pluractional meaning with a reflex of the PMP prefix **paŋ-* and its actor voice counterpart **maŋ-*. For certain predicates, this could be obligatory. For instance, the act of fishing, by its nature, involves repeated action and does not have a single fish as its target. The use of the pluractional has thus become obligatory in many languages for forming predicates such as 'to fish' and 'to harvest', which are naturally pluractional, as shown in (69).

Ibaloy (Ruffolo 2004:238)

(69) *jet eŋabot ni damot*
jət ʔəN-kabot ni lamot
then AV.PFV-harvest GEN root
'then he harvested some roots'

For other predicates, such as Tagalog *kuha* 'take', shown in (70), it is optional and adds a meaning ranging from repeated action, action on plural generic objects and unwanted persistence (De Guzman 1978).

Tagalog

(70) a. *k<um>uːha*
<AV>take
'to take'

b. *maŋ-[k]uːha*
AV.DIST-take
'to take (many)'

Although the distributive most often occurs in the actor voice form with a cognate of **maŋ-*, it is not restricted to the actor voice. As exemplified by Tagalog (71) and Sarangani Manobo (72), the distributive can co-occur with any voice in many languages.

---

[29] The replacement of *ka-* with *ma-* in the locative and conveyance voices appears to have been a gradual and messy process in the Central Philippine languages. In many languages, including Tagalog, the conservative *ka- -an* and *i-ka-* coexist alongside the innovative *ma- -an* and *ma-i-* but are used with innovative meanings or with a limited set of roots.

[30] Liao (2011b) argues that *maka- only bears a surface similarity to *paka-, but the evidence in favor of an underlying connection is in fact quite strong. Firstly, a number of phylogenetically disparate languages show an alternation between *maka- and *paka- in exactly those contexts where *<um> regularly disappears (including certain types of negation, imperatives and nominalizations), as seen for example in Casiguran Dumagat (Headland and Healey 1974), the Danao languages (McKaughan 1958) and Ilokano, among others. Secondly, the derivation of the actor voice potentive from the combination of the causative and the stative is entirely compositional with regard to its semantics (Kaufman 2009a, 2011, 2024c). The potentive poses many interesting puzzles of both a diachronic and synchronic nature which have only been touched upon lightly here. For more thorough discussion (from differing perspectives), see Dell (1983), Himmelmann (2006), Kaufman (2011b, 2012), Hauk (2019), Kroeger (1990), Baradjí et al. (2022).

Tagalog

(71) *i:log na la:bis na p<in>aŋ-isdaʔ-an*
river LNK overly LNK <BEG>DIST-fish-LV
'an over-fished river'

Sarangani Manobo (Dubois 1976:76)

(72) *i-m-pem-[b]egay dan se libro*
CV-PFV-DIST-give 3PL.GEN NOM book
'They gave out books.'

In Western Subanon, and most likely other languages as well, the distributive has been reanalyzed as a plain plural marker in reference to the nominative argument (Bulalang 2025:275). South of the Philippines, the distributive takes on new functions, such as that of a dedicated antipassive in certain South Sulawesi languages (Kaufman 2017), as well as the default marker of actor voice, as in Malayic languages.[31]

The actor voice distributive **maŋ-* also commonly plays a role in the formation of professional/habitual agent nominalizations, oftentimes with imperfective reduplication, as in (73).

(73) Ilokano (Rubino 1997:92)

a.	b.	c.
mannaray	*maŋŋalap*	*mannursuro*
maŋ-taray	maŋ-kalap	maŋ-su~suro
AV.DIST-run	AV.DIST-catch_fish	AV.DIST-IMPRF~teach
'racer'	'fisherman'	'teacher'

6.3.3 Sociative

A morphological category found most commonly among Philippine languages is the so-called 'sociative', expressed with a reflex of PMP *paki- or its actor voice counterpart, *maki-, discussed by Liao (2011a), who terms them 'comitative verbs'. In most cases, this morpheme can be translated into English as 'with others', as in Tagalog (74), although this often does not capture the relation between the agent and the others.

Tagalog

(74)a.	b.	c.
maki-hiŋiʔ	maki-ta:wa	maki-upoʔ
AV.SOC-request	AV.SOC-laugh	AV.SOC-sit
'to request'	'to laugh with others'	'to sit with others'

The sociative often connotes copying the action of others for social purposes, a meaning which is more salient for some predicates, such as (74b), than for others. The predicate *maki-ta:wa* is typically interpreted as laughing because other people are laughing whereas the predicate *maki-upoʔ* is simply to sit among others. The sociative need not denote a social activity in a positive sense. For instance, 'to fight' is often expressed with the sociative, e.g. Tagalog *maki-pag-a:way* (AV.SOC-TR-fight), Cebuano *makig-a:way* (AV.SOC-fight).

[31] Rubino (1997:205) describes Ilokano *maŋ-* as a detransitivizing prefix but it is not clear that *maŋ-* affixed verbs are really any less transitive than other actor voice verbs in this language.

The sociative also functions as a 'petitive' (Blust 2009:364), signalling a request on the part of the agent, as in (75).

(75) Ibanag (Dita 2011:200)
Maki-pisag na' nge?
AV.SOC-urinate 1s.NOM please
'May I use the bathroom?'

The imperative of the sociative, *paki-, has developed in another direction, now signaling a polite request in a number of Philippine languages, which Liao (2011a) suggests spread through Tagalog.[32]

6.3.4 Plural agent marking

It appears possible to reconstruct a PMP marker **si-* which necessitated a plural subject (reconstructed by Kitada 2019 as a sociative and by Liao 2011c as simultaneous aspect). In Central Philippine languages, we find a reflex in such forms as Tagalog *mag-si-takbo* (AV-PL-run), where it serves to mark plurality. In the Bisayan languages, a reflex of this prefix indicates individuated action over a group, translated with 'each (subject)' (Zorc 1977:143).

Although it is seemingly unattested for Philippine languages to show obligatory number agreement with any argument, plural marking can often be indicated simultaneously by several morphemes for emphasis. In Agutaynen (76), we find that the distributive **maŋ-* prefix has been reinterpreted as a plural agent prefix, which can co-occur with another plural marker *<Vr>*, commonly found in nearby Central Philippine languages, and the locative voice *-an* suffix used in its reciprocal function.

Agutaynen (Quakenbush et al. 2010:43)
(76) *mam-[p]ag-s<or>oay-an*
AV.PL-TR-<PL>fight-LV
'They will fight each other.'

Plural marking is often not uniform across word classes. In Tagalog, Agutaynen and elsewhere, adjectives with the uninflectable *ma-* prefix indicate plurality via CV-reduplication (without vowel length), e.g. Tagalog *ma-taba?* (ADJ-fat), *ma-ta~taba?* (ADJ-PL-fat). In Maranao, plurality on adjectives is marked with the *<aŋ>* infix, and in Cebuano, the *<g>* infix carries out the same function on dimension adjectives, e.g. *mu<g>bo?* (<PL>short), *da<g>ko?* (<PL>large). Ilokano also has a plural <g> formant, although this only emerges in demonstratives, shown in (77).

[32] As mentioned in fn.30, Liao (2011b:220) argues that there need not be a derivational relationship between all apparent *p-/*m- prefix pairs: "It is very likely that reflexes of *maka-/*paka- and *maki-/*paki- have become variants of one another due to the striking similarity in their forms and functions." She observes that only **paki-* forms are used in polite requests and only **maki-* forms are used in "permissive" social constructions. This, however, can be derived from the fact, also noted by Liao, that *<um> was not present in imperatives in the PMP voice paradigm, leaving underlying *p- initial prefixes to surface unobstructed. The fact that polite requests are strictly *imperatives* predicts that they may conservatively lack *m-*, even in languages like Tagalog, which have lost *p-* initial imperatives in other paradigms.

Ilokano (Rubino 1997:41)

(77)

	SING	PL
PROX	*daytoy*	*da**g**itoy*
MED	*dayta*	*da**g**ita*
DIST	*daydiay*	*da**g**idiay*

6.3.5 Multifunctional *paR-/maR-

Reflexes of **paR*- (**maR*-, in the actor voice) can be found in almost all Philippine languages although the range of functions associated with these morphemes differ from language to language. As Pittman (1966) first noted, Tagalog *mag*- has apparently contradictory functions, in some cases increasing valency, e.g. <*um*>*akyat* 'to ascend' vs. *mag-akyat* 'to bring something up', and in other cases, e.g. <*um*>*ahit* 'to shave others' vs. *mag-ahit* 'to shave one's self', decreasing valency. Kaufman (2009a, 2018b) derives the apparently contradictory functions of this affix by viewing it as a historically complex combination of two components: the well attested causative prefix **pa*- and a middle voice prefix **R*-, which fused with the former. With some roots and paradigms, it is the causative *pa*- function which is meaningful while in other cases it is the middle voice whose interpretation prevails. The middle function of **R*- is also implicated in the durative, reciprocal and reflexive functions found with the ***paR-/maR***- prefix. This can be seen in the contrast between (78a) and (b), where *nag*- (the perfective form of *mag*-) carries a reciprocal meaning.

Ilokano (Rubino 1997:204)

(78)a. *s<imm>ina=da iti guŋlo*
<AV.PRF>separate=3p.NOM OBL club
'They separated from the organization.'

b. *nag-sina=da*
nag-sina=da
'They separated from each other.'

In many languages of the southern Philippines, the "plain" actor voice *<*um*> paradigm increasingly gives way to a **maR*- paradigm (Liao 2004:106; Lobel 2004, 2013:46-47). This prefix also appears to have been borrowed in several areas in the Philippines as the reflex of **R* often does not match regular sound correspondences (Liao 2004:107-12, Reid and Liao 2004:457). Specifically, subgroups which show *R>y (e.g. Central Luzon and Bashiic), *R>l (e.g. Southern Cordillera) and *R>r (e.g. Ilokano) also show sporadic instances of a *g reflex of *R in PMP *maR-. Gallego (2022) investigates this in detail with respect to the Batanic languages, which generally show *R>y but which also tend to employ *mag*-, as opposed to the native form of the prefix, with stems borrowed from *R>g languages.

6.3.6 Reciprocals

There are several recurring strategies for forming reciprocals in Philippine languages. The first, shown in Tagalog (79), involves an apparent circumfix formally consisting of the actor voice prefix together with the locative nominalizer/voice suffix, i.e. **maR*-√-*an*, a formation which is also found in Malay (e.g. *bər-təŋkar-an* AV-fight-RECP). The second, exemplified by Samar-Leyte (80), involves the **maR*- prefix together with the **ka*- prefix, one of whose functions is similar to English *co*-, deriving a partner in sharing something denoted by the stem. This later formation may only happen to overlap semantically with the reciprocal proper in (79), as it more often refers specifically to two agents sharing in an activity.

Tagalog

(79) *nag-patay-an sila*
AV-kill-RECP 3PL.NOM
'They killed each other.'

Samareño (Waray)

(80) *nag-ka-du:rug hira*
AV-CO-sleep 3PL.NOM
'They slept together.' (Zorc 1977:144)

In some cases, the **maR-* prefix expresses a reciprocal on its own, as in Tagalog *mag-kita?* AV-see 'to meet' and Ibaloy (81), where *nan-* reflects the perfective form of PMP **maR-*.

Ibaloy (Ruffolo 2004:23)

(81) *jet nan'aseba ira*
jət nan-ʔasəwa ʔida
then AV.PFV-spouse 3p.NOM
'then, they got married (to each other)'

There are other reciprocal markers whose etymologies are not so clear. For instance, Tboli marks reciprocals with an *s-* prefix (likely derived from PMP **si-* discussed in §6.3.4), e.g. *tagak* 'to leave behind' *s-tagak* 'to leave each other'; *toboŋ* 'to help' *s-toboŋ* 'to help each other' (Forsberg 1992:91). In Binukid, as well as several Bisayan languages, the reciprocal is expressed with a circumfix whose first part is the <*in*> infix and the latter part is *-aʔ* or *-ay* (with *-ay* also appearing in the Bisayan languages), as seen in (82). Although both components of this circumfix occur in other derivations, they do not seem to be semantically related.

Binukid (Post and Gardner 1992:xxiv)

(82) *m<i>g-b<in>ulig-aʔ*
<PFV>AV.DUR-<RECP$_1$>help-RECP$_2$
'They helped each other.'

The presence of an <in> formant in verbal reciprocals across distantly related subgroups of the northern and southern Philippines suggests that it must be reconstructed in this function. If it is somehow related to the perfective PMP **<in>* infix, the semantic connection is far from obvious. Examples from the Northern Luzon languages can be seen in Ilokano forms such as *t<inn>aray* (<RECP>run) 'racing with one another' (Rubino 1997:86) as well as the Ibaloy pair in (83).

Ibaloy (Ruffolo 2004:231)

(83)a. *mantocho ira*
man-todo ʔida
AV.IPFV-teach 3p.NOM
'they will teach'

b. *mantinoro ira*
man-t<in>odo ʔida
AV.IPFV-<ITER>teach 3p.NOM
'they will teach each other'

Finally, we find what appear to be reciprocals derived with a *ki-* formant, as in Cebuano *maki-g-away* 'to fight' which can be analyzed as containing the actor voice sociative prefix *maki-* plus *g-*, which most probably derives from PMP middle voice marker **<R>*. A similar analysis may also be possible for the Arta *maŋiC-* prefix (Kimoto 2017:306), which suggests a composition of the actor voice distributive prefix *maŋ-* with the sociative and middle formatives, as in Cebuano. The reason reciprocals have such varied exponence in Philippine languages is most likely that the semantics of the reciprocal overlaps with several distinct functions that were

signalled independently in PMP (or later proto languages), namely, the middle voice *<R> (Kaufman 2009a), the sociative **paki-/maki-**, the counterpart prefix *ka-*, various plural marking strategies, including reduplication, and even the pluractional/distributive **paŋ-/maŋ-**, especially where it has been reanalyzed as a plain plural marker. Teasing apart which patterns are reconstructable to what proto-languages amongst these recurring developments is thus no easy matter.

6.3.7 Inchoative

The inchoative, termed by Zorc (1977:142) "essive", has barely been investigated from a comparative perspective. In many languages of the central and southern Philippines, the inchoative function with an entity denoting root is signaled by a unique prefix, as shown in (84) and (85).[33]

Aklanon (Zorc 1977:142)

(84) *nagiŋ-rayna si neli*
AV.PFV.INCH-queen PL.NOM Neli
'Nellie became a queen.'

Tiruray (Bennásar 1892:40)

(85) *mente-eteu*
AV.INCH-person
'to become a person'

In other languages, the inchoative is expressed either with one of the regular actor voice morphemes (usually a simple reflex of **maR-**) or periphrastically with a predicate meaning 'change/transform', as exemplified by the pairs in (86) and (87).

Ibaloy (Ruffolo 2004:235)

(86)a. *man-doktor sota anak=ko*
AV-doctor NOM child=1S.GEN
'My son will become a doctor.'

b. *nan-balin ni tilay*
AV.PRF-transform GEN lizard
'He transformed himself into a lizard.'

Maranao

(87)a. *m<iy>ag-ɤtor*
AV<PFV>-rock
'changed into a rock'

b. *mim-bɤloy a ator*
AV.DIST-change LNK stone
'became a rock'

In some cases, the inchoative form allows for non-actor voice derivations. These were still current in the Tagalog of the early 20th century, as seen in (88), but are now obsolete.

Tagalog (Lendoyro 1909:256)

(88) *aŋ maynila aŋ p<in>agin-pari:ʔ-an niya*
NOM Manila NOM <BEG>INCH-priest-LV 3SG.GEN
'It was in Manila where he was ordained a priest.'

With property denoting roots, there is far more cross-linguistic agreement in the use of PMP *<um> reflexes to obtain an inchoative meaning, as in Ilokano *t<um>ayag* and Tagalog *t<um>angkad* 'to become tall', Dupaningan Agta *d<um>akal* 'to become big' (Robinson

[33] Ilokano, Ibanag, and other North Luzon languages show an apparent cognate of the Tagalog inchoative *magiŋ* where it functions as a "pretentaive". Such verbs signal that the agent pretends to have the quality indicated by the stem (Rubino 1997:296) or carry out the action indicated by the stem (Dita 2010:202-205).

2008:154), Tuwali Ifugao *d<um>ukke* ‘to become long’ (Hohulin and Burquest 2014:85), *l<um>anggas* ‘to become thin’ (Bulalang 2025:109).

6.4 Causative

The PAn causative prefix ***pa-*** is perhaps the most stable affix in the entire PMP morphological inventory. The causative introduces a causer into the argument structure and can co-occur with any voice, mode and aspect. Abstracting away from various complications, Table 7 shows the canonical mapping of roles to arguments in a causative clause.

Table 7. Canonical role/case correspondences in the causative

	GEN	**NOM**	**OBL**
Actor voice	theme	causer	causee
Patient voice	causer	causee	theme
Conveyance voice	causer	theme	causee

In an actor voice causative clause, as in (89), the nominative argument is the causer while the theme is expressed just as an actor voice object would be expressed. The causee, on the other hand, is expressed as an oblique argument.

Tagalog

(89) *nag-pa-suːlat ako naŋ liːham sa estudyaːnte*
AV.BEG-CAUS-write 1SG.NOM GEN letter OBL student
‘I had a student write a letter.’

In a patient voice causative clause, as in (90), it is always the causee that is selected as the nominative argument rather than the theme. The agent is assigned genitive case, as expected, and the theme, if expressed, is assigned genitive or objective case.

Tagalog

(90) *p<in>a-suːlat-Ø ko naŋ liːham aŋ estudyante*
<BEG>CAUS-write-PV 1SG.GEN GEN letter NOM student
‘I had the student write a letter.’

The conveyance voice consistently selects causative themes as the nominative argument, regardless of what voice is used to “promote” the notional object to nominative in a non-causative clause. The example in (91) shows how the causer is expressed as a genitive agent, as in the other non-actor voices, the causee is expressed as an oblique, and the theme or “notional object” becomes the nominative argument.

Tagalog

(91) *i-p<in>a-suːlat ko sa estudyante aŋ liːham*
CV-<BEG>CAUS-write 1SG.GEN OBL student NOM letter
‘I had the student write a letter.’

The causation of states is often expressed differently than the causation of activities, often with a form derived from **maka-*, the actor voice of the combination **pa-ka-* CAUS-STAT-. In Tagalog, such forms are generally 'frozen' in the progressive aspect. For instance, Tagalog *na-ka~ka-baliw* (AV.BEG.CAUS-STAT~IMPRF-crazy) 'maddening', does not have a perfective form **naka-baliw* or a prospective aspect **maka~ka-baliw*. In Ilokano, functionally cognate constructions are formed similarly (with or without a reflex of *<in>), as in *maka-sa~sadut* 'causing laziness' (Rubino 1997:220) as well as with the addition of a more transparent *pa-* causative, as in *maka-pa-salunʔat* 'healthy' (Rubino 1997:150).

The causative PMP **pa-* prefix appears to be implicated in a very large number of morphological functions, several of which are not transparently causative, as in the *mapa-* prefix, which indicates involuntary movements in Ilokano, Tagalog, and other languages, as well as the verbalization of certain non-eventive roots, like Ilokano *p<um>a-laem* and its Tagalog cognate *p<um>a-lalim* 'to go into the deep'.

6.5 Negation

Philippine languages are relatively rich in negators, especially the Central Philippine languages, which often have distinct negators for perfective events, prospective events, prohibitives (imperatives), identification and existential predication. Few if any languages possess five distinct negators for each of these functions, but many languages show three way distinctions (Tagalog, N. Subanen, Maranao, Tboli, Ilokano) and four-way distinctions (Aklanon), while only one language in the same shown in Table 8, Dupaningan Agta, has a single negator that can serve all five purposes (although Robinson (2008:22) reports that *limus* was cited as the original negative existential but not encountered in actual speech).

Table 8. Negators

	Tagalog	**Aklanon**	**Northern Subanen**	**Maranao**	**Tboli**	**Ilokano**	**Dupaningan Agta**
PERFECTIVE EVENT	*hindiʔ*	*ʔuwaʔ*	*ʔəndaʔ*	*diʔ*	*laʔ*	*di/saan*	*awan*
PROSPECTIVE EVENT	*hindiʔ*	*ʔindiʔ*	*ʔəndiʔ*	*diʔ*	*laʔ*	*di/saan*	*awan*
PROHIBITIVE	*huwag*	*ʔayaw*	*ʔəndiʔ*	*diʔ*	*béʔ*	*saan*	*awan*
IDENTIFICATION	*hindiʔ*	*bukon*	*gənnaʔ*	*kenaʔ*	*sundu*	*saan*	*awan*
EXISTENTIAL	*walaʔ*	*ʔuwaʔ*	*ʔəndaidun*	*daraʔ*	*(laʔ wən)*	*awan*	*awan* (*limus*)

What is termed here 'event' versus 'identification' negation is often framed in terms of lexical categories, e.g. verbal, nominal, and adjectival negation. Non-verbal negation can often be traced to a word meaning 'different'. For instance, Blust et al. (2023) reconstruct both PWMP **beken* 'negator of nominals, other, different' as well as PWMP **laqin* 'different', which also comes to function as a general negative marker in Sorsogon.

It is a common feature of Malayo-Polynesian languages outside the Philippines to combine the event negator with the existential to derive a negative existential (e.g. Malay *tidak*

ada NEG EXT and *ti-ada* NEG-EXT), but the vast majority of Philippine languages employ distinct unanalyzable roots for the existential and negative existential.[34]

In the northern languages, we find multiple declarative negators whose usage may overlap, for instance, in a plain declarative clause such as (92), the two Ilokano negators *saan* and *di* are judged to be semantically equivalent, although *saan* requires the linker.

Ilokano (Rubino 1997:55)

(92)a. *saan=da a naŋan*
NEG=3P LNK AV:eat
'They did not eat.'

b. *di=da naŋan*
NEG=3P AV:eat
'They did not eat.'

In a number of languages, certain negative contexts require the dependent verbal paradigm, as discussed by Wolff (1973) and Zorc (1977).

6.6 Gerunds

Most Philippine languages have a distinct morphological strategy for forming event nominalizations although a higher level reconstruction of this function has not yet been attempted. Gerunds are typically formed with the underlying mode prefixes that the event-denoting root would take in the actor voice, most often a reflex of *paR- or *paŋ-, as shown in (93). There is no dedicated gerund morpheme in these cases but rather the gerund resembles an intermediate form on its way to becoming an actor voice verb but without the addition of actor voice *m-*. This is the not the whole story, though, as Central Philippine languages often employ a reflex of *paR- for the gerund whose corresponding actor voice form takes <um> without any mode prefix. For instance, Tagalog *t<um>akbo* 'run' forms its gerund as in (93g), even though *mag-takbo* has an intensive or transitive meaning that is not present in (93g).[35] Other strategies exist, too. Notably, apparent reflexes of the CV prefix *(h)i- form gerunds with certain verbs in a number of North Luzon languages, as seen in (93f), and an old nominalizing prefix of the form *ta- appears to be implicated in Bagobo-Klata formations of the type seen in (93h).

(93)
a. *pag-doŋol* (Southern Alta, Abreu 2018:150)
GER-listen
b. *pən-pəltag* (Northern Alta, Garcia-Laguia 2018:54)
GER-fish
c. *pag-gimit* (Arta, Kimoto 2017:145)
GER-do
d. *paŋ-[k]ulug* (Tuwali Ifugao, Hohulin and Burquest 2014:52)
GER-believe
e. *p<an>ag-basa* (Ilokano, Rubino 1997:102)
<GER>GER-read

[34] Eastern Tawbuid provides a counterexample: the existential predicate *(k-)e*, is negated as any other verb or adjective (Fleming 2022:31), as in Malay.

[35] In Tagalog, the difference between actor voice predicates in <*um*> and *mag-* is expressed with reduplication in the corresponding gerund of the *mag-* form. Thus, in contrast to (93g), the gerund of *mag-takbo* would be *pag-ta~takbo* 'running (something)'.

f. *i-sa-saŋpet* (Ilokano, Rubino 1997:102)
CV-GER~arrive
g. *pag-takbo* (Tagalog)
GER-run
h. *tɔ~lalɛ* (Bagobo-Klata, Estrera 2024:109)
GER-run

Some languages have aspect inflected gerunds, which brings them even closer to the already noun-like voice marked predicates used to denote events. The Ibaloy gerund in (94), *indawmi*, and the Ibanag gerund in (95), *pinallubbe,* are true event nominals, but are marked for perfective aspect. Their status as gerunds is clear from their reference to the event as a whole and the lack of nominative case marking on a pivot argument.

Ibaloy (Ruffolo 2004:108)
(94) *nəm ʔəman-ʔodan nonta ʔin-law=mi*
but ACTV/PROG-rain when.PAST MNRGER/PRF-go=1pl.ex.GEN
'but it was raining during our departure'

Ibanag (Dita 2010:206)
(95) *in-istorya=ku i p<in>ag-lubbe=ku ta kagayan nira*
PRF-narrate=1s.GEN NOM <PRF>GER-go.home=1s.GEN OBL Cagayan 3p.OBL
'I narrated to them my going home to Cagayan.'

In the next section, we turn to the ordering relations between words and syntax proper.

7. Elements of syntax

In this section, I present the basic word order across various phrase types (§7.1), and then take a closer look at word order within the noun phrase (§7.2) and the clause (§7.3). Finally, I look at the syntax of referential expressions: pronouns, demonstratives, case markers and the positioning of pronominal clitics (§7.4).

7.1 Basic word order relations

All Philippine languages are robustly head-initial, as can be seen in the basic ordering relations exemplified by Tagalog in (96), where ≷ represents a robust tendency in linear order and > represents an even stronger generalization.

(96)a. **Pred > Subj**
matali:no si bo:boy
smart P.NOM Boboy
'Boboy is smart'

b. **Noun > Possessor**
aŋ na:nay ni keŋkoy
NOM mother P.GEN Kengkoy
'Kengkoy's mother'

c. **Adj ≷ Noun**
mataŋkad na baba:ʔe
tall LNK woman
'tall woman'

d. **Verb ≷ Adv**
t<um>akbo naŋ mabilis
<AV>run GEN fast
'to run fast'

e. **Adposition > Noun**
ga:liŋ sa gu:bat
from OBL jungle
'from the jungle'

f. **Title > Name**
ginoʔo=ŋ reyes
mister=LNK Reyes
'Mister Reyes'

g. **Complementizer > Clause**
aka:laʔ ni dodoŋ na matali:no siya
thought GEN Dodong COMP smart 3SG.NOM
'Dodong thinks he's smart.'

h. **Noun ≷ Relative Clause**
daga=ŋ p<in>atay-Ø ni=Kengkoy
rat=LNK <BEG>kill-PV GEN=kengkoy
'a rat killed by Kengkoy'

i. **Aux > Verb**
da:pat mag-madali=ka=na!
must AV.TR-hurry=2SG.NOM=already
'You should hurry up!'

j. **Comparative > Adjective > Standard**
lalo=ŋ mataŋkad sa kanya
more=LNK tall OBL 3SG.OBL
'taller than him/her'

k. **Negation > Verb**
hindiʔ s<um>ayaw
NEG <AV.BEG>dance
'didn't dance'

However, not all these relations are equal. Some, such as (e), (f), (g), (i), (j) and (k) are relatively strict or invariable. Others, such as (a) and (b), allow for alternative constructions with different semantic or pragmatic implications. A third category, which includes (c), (d) and (h), represent tendencies but co-exist with equally unmarked alternative orders. We examine these in the following subsections.

7.2 Word order within the noun phrase

In addition to a complex set of voice markers, nearly every language of the Philippines possesses a set of case marking determiners. Case markers always precede the noun phrase but many modifiers can be positioned flexibly, despite having an unmarked order. The canonical order of elements in a typical Central Philippine noun phrase is shown in (97), from Tagalog.

(97)

CASE	PRE-POSS	NUM	ADJ	ADJ	N
aŋ	*[kanya=ŋ]*	*maŋa*	*ma-ga~ganda=ŋ*	*pula=ŋ*	*bulaklak*
NOM	3SG.OBL=LNK	PL	ADJ-PL~beauty=LNK	red=LNK	flower

'his/her beautiful red flowers'

The case marker is in initial position, followed by the preposed possessor. Possessive pronouns can either be expressed as independent pronominals in the oblique case, typically preceding the possessum and connected via the linker (Hsieh 2023), or as second position clitic pronouns in the genitive case (not shown above). Only in very rare cases (e.g. Hanunoo, Epo 2014) has the preposed position become the norm.

Following this position we find a plural marker, which is ubiquitous in the languages of the central and southern Philippines but infrequent in the north, where a range of plural marking strategies are employed for arguments, including partial reduplication (e.g. Keley-i *baka* 'cow', a Spanish loan, becomes *babakka* in the plural, with CV-reduplication and gemination of the

second consonant).[36] Other languages mark number syncretically with the case marker, as in Arta *i* versus *tidi*, as seen in (98).

Arta (Kimoto 2017:190)

(98) Um-bèr {i/tidi} manuː=i ayti Dipintin=i
INTR-fly SG.NOM.DEF/PL.NOM.DEF bird=SPEC here Dipintin=SPEC
'The bird/birds will fly to Dipintin.'

In some languages, affixal plural morphology may be further accompanied by functional elements in the NP. This can be seen in Ibaloy (99), where the CV reduplication marks plurality on the noun itself while the clitic /ʔida/, which appears to have developed from the identical 3pl pronominal clitic and which can appear in several positions within the NP, expresses plurality again on the phrasal level.

Ibaloy (Ruffolo 2004:342)

(99) *inang ni paydeng para soni aanakto ira*
naN-ʔala ni pajləŋ pada so=ni CV-ʔanak=to ʔida
AV.PRF-get GEN fish for OBL=GEN PL-child=3.GEN PL
'he took some fish for his children'

Adjectival modifiers typically precede the noun, but flexibility is found in most languages that preserve the linker. The linker, which signals all types of modification, is common to the vast majority of Philippine languages but far rarer south of the Philippines, even among Philippine-type languages and those of Blust's (2019) Philippine subgroup. The presence of the linker correlates closely with freer word order within the noun phrase, including both the position of adjectival modifiers and relative clauses. The Bilic and Sama languages of southern Mindanao and Sulu pattern with their southern neighbors in lacking the linker and displaying a rigid word order within the NP.[37] In some North Luzon languages, the linker is homophonous with a genitive marker but in most Philippine languages the two functions are distinct.

Demonstratives are somewhat more difficult to generalize over. When case is marked syncretically on demonstratives (e.g. Tagalog *ito* 'this (neutral)', *nito* 'this (GEN)', *diːto* 'this/here (OBL)'; Ibaloy *ima* 'this (NOM)', *niya* 'that (GEN)'), they appear to always precede the head of the noun phrase. In their neutral forms, they may sandwich the noun phrase, as in Tagalog and

[36] Zorc (1977:103) claims that the plural marker (or "diversity marker") *maŋa* is found in all the Bisayan languages. Blust and Trussel (ongoing) reconstruct PMP **maŋa* as a prenominal plural marker. See Lynch et al. (2002: 90–91) for its history in Oceanic and Wu (2017) for a general look at plural markers in Austronesian, including the distribution of **maŋa*.

[37] As discussed by Donohue (2007:359-363), a rigid Noun-Adjective order emerges south of the Philippines and is common to languages of the Southeast Asian mainland. There is a marked difference between Central Philippine languages and those of the southern periphery in this regard, where the Bilic and Sama groups pattern similarly to languages of Indonesia. Even languages of northern Sulawesi belonging to Blust's Philippine subgroup appear to show Donohue's (2007) southern pattern, e.g. Buol *botu moitomo* stone black (Zobel 2005:633). On the Bornean side, Kroeger (2005:411) describes the Kimaragang order of elements within the NP as: Determiner (Number) N (Possessor) (Modifier). It is only the unmarked position of the modifier that has shifted to the right edge when compared with the Central Philippine languages.

other languages. Elsewhere, as in Subanen, the demonstrative must come on the right edge of the noun phrase (Daguman 2004:148, Bulalang 2025:284).

Despite referring here to the 'noun phrase', it is important to note that no language of the Philippines requires arguments to be headed by a lexical noun. Verbs and adjectives (that is, words inflected by voice, aspect or stative morphology) very often serve as the sole lexical head in the argument phrase, as exemplified by (100).

Batad Ifugao (Newell 2005:24)

(100) *nan in-anup-an nan linala:ʔi*
DEM PRF-hunt-LV DEM man.PL
'the thing which was hunted by the men'

Such structures have been interpreted as mere headless relatives by some and as evidence for a fundamental difference in the organization of lexical categories in Philippine-type languages (see Himmelmann 1991, Kaufman 2009b and responses in same volume, Kaufman 2017, Foley 2017, and references therein). Crucially, there is no evidence in any Philippine languages that structures like *nan inanupan* (100) are any more complex or marked than arguments with a nominal head such as *nan linala:ʔi*.

7.3 Word order within the clause

As with all conservative MP languages, Philippine languages are almost without exception predicate initial across lexical category and clause type. Beyond the simple predicate-initial generalization, the question of the basic order of phrases within the clause has never been answered definitively. Furthermore, as Himmelmann (2005:143) notes, there have been unwarranted claims of total freedom of phrasal order in the post-predicate domain. When considering clauses with full NP arguments, nearly all languages show the basic order shown in (101) for undergoer voice (i.e. non-actor voice) clauses and actor voice clauses.

(101) a. <u>Undergoer voices</u>
V A_{GEN} P_{NOM}

b. <u>Actor voice</u>
V ($P_{OBL/GEN}$) A_{NOM} ($P_{OBL/GEN}$)

In the undergoer voices, there is a very strong tendency for the genitive marked agent to immediately follow the predicate head. In languages with impoverished case marking, like Ilokano and Tboli, this tendency becomes a rule. In the actor voice, the ordering relations appear to be less fixed although if there is an unmarked order, it tends to be one in which the nominative argument follows the patient. Current generative literature often assumes [S[VO]] to be a universal base structure, deriving predicate initial order via verb movement, VP movement, subject lowering or some other syntactic process. None of these approaches, however, predict that the verb forms a constituent with the agent in the undergoer voices and with the undergoer in the actor voice, as evinced by coordination facts (Kroeger 1993, Kaufman 2009b) and even higher level prosody (Hsieh 2016).

The preverbal domain is typically reserved for pragmatically marked arguments and adjuncts (see Naylor 1975, Kroeger 1993, Kaufman 2005, Nagaya 2007 for Tagalog). All Philippine languages allow for topicalization of the nominative/absolutive argument to a preverbal position (Reid & Liao 2004:447). The fronted topic is often followed by a dedicated topic marker, but in languages like Cebuano, there is topic fronting without a topic marker. In

rare cases, an unmarked SVO word order emerges, presumably from the semantic bleaching of topicalization, as seen in Buhid of southern Mindoro (Zorc 1974; Lobel 2013:188-193).

While phrases taking nominative and oblique case can be topicalized in most languages, phrases taking the genitive case are typically banned from topicalization, as exemplified by Tagalog in (102).

Tagalog

(102)a. *k<um>a:ʔin naŋ maŋga aŋ ba:taʔ sa ku:sina*
<AV.BEG>eat GEN mango NOM child OBL kitchen
'The child ate the mango in the kitchen.'

b. *aŋ ba:taʔ ay k<um>a:ʔin naŋ maŋga sa ku:sina*
NOM child TOP <AV.BEG>eat GEN mango OBL kitchen
'The child, ate the mango in the kitchen.'

c. *sa ku:sina ay k<um>a:ʔin naŋ maŋga aŋ ba:taʔ*
OBL kitchen TOP <AV.BEG>eat GEN mango NOM child
'In the kitchen, the child ate the mango.'

d. **naŋ maŋga ay k<um>a:ʔin aŋ ba:taʔ sa ku:sina*
GEN mango TOP <AV.BEG>eat NOM child OBL kitchen
(For, 'A mango, the child ate in the kitchen.')

This restriction extends to relativization and cleft-like constructions in addition to topicalization. There is a sprawling theoretical literature on this pattern, which cannot be reviewed here but see Kaufman (2017) for an attempted summary.

The Central Philippine languages appear to have innovated a special focus position for fronted oblique arguments and adjuncts shown in (103).

Tagalog

(103) $_{\text{Foc}}$[*Doon sa bundok*]=*kayo maŋ-a:so!*
DIST OBL mountain=2p.NOM AV.DIST-dog
'It's *there in the mountain* where you hunt!'

The focus fronted oblique phrase attracts second-position clitics and receives a cleft-like "exhaustive list" interpretation (i.e. 'in the mountain and nowhere else') (Kaufman 2005). In many Bisayan languages, focus fronting of an oblique phrase in this manner requires using the dependent paradigm of the verb. This construction is generally uncommon, if attested at all, in the North Luzon languages. The northern analogue employs a fronted prepositional phrase, as in (104) (headed by the deictic *sidi*), but with a locative voice construction as the pivot. It thus seems to display the same predicate-pivot structure of a typical Philippine clause without the oblique focus fronting seen in (104).

Kankana-ey

(104) *sidi ay duntug di pan-ʔanup-an=jo!*
DIST LNK mountain NOM TR-hunt-LV=2p.GEN
'There in the mountain is your hunting place!'

7.4 Referential expressions

7.4.1 Pronouns

Philippine languages have distinct pronominal paradigms for at least nominative, genitive and oblique cases.[38] A typical example in this respect can be seen in the Maranao pronouns in Table 9 (McKaughan 1958, Kaufman 2010b). As is typical, clusivity is distinguished in the first person plural although several languages distinguish a dual (1+2 person) form from a plural inclusive (Liao 2008, Reid 2009), as in Maranao.

Table 9. Maranao pronoun paradigm

	Nom (bound)	**Nom (free)**	**Gen (bound)**	**Obl (free)**
1sg	*(a)ko*	*sakən*	*akən ~ ko*	*rakən*
1pl.ex	*kami*	*səkami*	*(a)mi*	*rəkami*
1+2 dual	*ta*	*səkta*	*ta*	*rəkta*
1pl.in	*tano*	*səktano*	*tano*	*rəktano*
2sg	*ka*	*səka*	*(ŋ)ka*	*rəka*
2pl	*kano*	*səkano*	*(n)iyo*	*rəkano*
3sg	*səkaniyan*	*səkaniyan*	*(n)iyan*	*rəkaniyan*
3pl	*siran*	*siran*	*(i)ran*	*kiran*

In the central and southern languages, an independent set of genitive or oblique pronouns are used to indicate possessor predicates but in the northern languages, such predicates tend to be expressed with a dummy noun and a genitive clitic, as seen in the contrast in (105).

	Ilokano		Tagalog	
(105)a.	*kukua=da*	*ti balay*	b. *kanila*	*aŋ bahay*
	thing=3p.GEN	ART house	3p.GEN	ART house
	‘The house is theirs.’		‘The house is theirs.’	

The set of case markers used with personal names are typically reflected in the nominative and genitive pronouns, a pattern that can be reconstructed to at least PMP if not PAN (Blust 1977). For instance, a name in the nominative and genitive cases would be expressed as in *si Ondoy* (P.NOM Ondoy) and *ni Ondoy* (P.GEN Ondoy), respectively, whereas the corresponding pronouns would be *siya* (3sg.NOM) and *niya* (3sg.GEN). Oblique pronouns are formed with a range of formants, typically a reflex of the prepositions/case markers *di, *sa, *ka or some combination thereof. Typically, the formants found in participant pronouns (first and second person) differ

[38] Pronouns are not necessarily a unique part of speech in Philippine languages, as they can occasionally form a base for voice marking, e.g. Tagalog *<um>a:min* (<AV>1pl.ex) ‘to admit’, which can be compared with Malay *məŋ-aku* (AV-1sg) ‘to admit’.

from those in non-participant pronouns (third person) although it is difficult to generalize due to the vast amount of analogical change and recombination.[39]

Indefinite pronouns are most often constructed with an interrogative pronoun combined with a frozen clitic adverb, for instance, *=man* 'even' in Tagalog *sinu-man* (who-even) 'whoever' or Ilokano *ania man* (what even) 'whatever' (Rubino 1997:71).

Reflexive pronouns are generally constructed with a noun such as 'self' or 'body' and a genitive pronoun, not unlike English 'herself', 'myself', as in (106). It appears that no Philippine language has a dedicated set of reflexive pronouns, as commonly found in Indo-European languages.

Kankana-ey (Allen 2014:230)

(106) *I-saad=na din awak=na ay pangolo.*
CV-establish=3s.GEN NOM body=3s.GEN LNK leader
'He sets himself up as leader.'

7.4.2 Demonstratives and deictics

Demonstratives and deictics in Philippine languages often distinguish three types of proximity: speaker proximate, hearer proximate and distal, although languages may conflate this to a two-way split or, conversely, elaborate it with finer distinctions. The Philippine language with the most deictic distinctions appears to be Western Subanon (Bulalang 2025:96), shown in Table 10.

Table 10. Western Subanon deictics

Decitic	**Meaning**
koni	The speaker is holding or touching the entity
konia	The speaker is near the entity
koyon	The addressee is holding the entity or it is near to both the speaker and addressee
kio/koyo	The entity is close but not too close to the speaker and the addressee
kioyo/koyoyo	The entity is far from both the speaker and addressee
kitu'/kotu'/kutu'/kuitu'	The entity is very far from both the speaker and addressee or referring to an entity common to both the speaker and the addressee but not physically present

Deictics are in most languages derived transparently from demonstratives with one of the PAn locative/directional markers *sa, *ka, *di mentioned above (Ross 2006 and Blust et al. 2023). Many North Luzon languages imbue deictic meanings with a temporal dimension and may even

[39] For instance, in both Tagalog and Maranao, the formant deriving from *ka is found only in the non-participant pronouns, e.g. *rek**an**iyan*, ***k**iran* in Table 9 and Tagalog ***ka**niya* (3sg.OBL), ***ka**nila* (3pl.OBL) although there is considerable variation even within Central Philippine languages, some of which show a total lack of *ka in the oblique pronouns (e.g. Naga Bikol) and others which show the use of *ka throughout the entire oblique paradigm (e.g. Cebuano).

possess deictics that are used specifically with deceased persons or things no longer in existence. Rubino (1997:45) notes such use with the Ilokano remote past demonstrative *daydi,* in contrast to the recent past demonstrative *daytay*.

In addition to demonstratives, many Philippine languages have specificity marking enclitics independent of case markers, deictics and demonstratives, although often in agreement with them, especially when the case marking system encodes definiteness. For instance, Arta obligatorily distinguishes indefinite and definite referents in its case marking system but also has optional specificity markers that encliticize to the first word in the argument and indicate that the referent is a concrete entity known to the speaker. This is seen in the minimal pair in (107) where an indefinite agent is introduced with the case marker *na* but a definite agent is introduced with *ni* and contains the *=i* enclitic.

Arta (Kimoto 2017:191)

(107)a. *s<in>a:-sa:ngor=tid na dapug*
<PST>RDP-horn=3P.NOM GEN.INDEF water.buffalo
'They were attacked by water buffalo(s) with their horns.'

b. *s<in>a:-sa:ngor=tid ni dapug=i*
<PST>RDP-horn=3P.NOM GEN.DEF water.buffalo=SPEC
'They were attacked by the water buffalo(s) with their horns.'

A parallel development can be seen in the Bisayan language, Cebuano, where the *=a* enclitic appears on both definite (108a) and indefinite (108b) arguments with a specific reference.

Cebuano (Wolff 1972)

(108)a. *agi, ka-daku? a:na=ŋ isda?=a!*
oh STA-big that=LNK fish=SPEC
'Oh! How big that fish is!'

b. *unsa=y uras=a?*
what=INDEF hour=SPEC
'What is the time?'

7.4.3 Case markers

Austronesian case markers have received ample attention from a historical perspective (Reid 2002, Blust 2005b, Reid 2007, Ross 2006, Blust 2015) and Philippine languages offer the best evidence for their reconstruction to PMP, as case marking is often lost (and possibly reinnovated) south of the Philippines. In nearly all Philippine languages, however, they form an integral cue to argument structure and grammatical functions. As seen in Table 11, Philippine case marking systems typically distinguish three cases, labelled here nominative, genitive and oblique, although some have conflated certain cases or have lost case marking on noun phrases altogether.[40] A handful of languages, including the Batanic languages and some Central Luzon languages like Kapampangan, mark non-nominative objects uniquely. Most languages, however, use either the genitive case or oblique case for non-nominative objects.

[40] 'Nominative' and 'genitive' are labelled 'absolutive' and 'ergative' under ergative analyses (De Guzman 1988, Payne 1982, Aldridge 2004 inter alia) while nominative is often labeled 'pivot' under symmetrical and case agreement approaches (Chen 2017, Reisberg 2014, inter alia). 'Nominative' is used here without any implication that Philippine-type languages display a nominative-accusative alignment pattern. The label 'genitive' is preferred because this case does double duty in all Philippine languages (except for Tawbuid, see Fleming 2022), marking possessors in addition to non-actor voice agents.

Table 11. Common noun case markers

	Tagalog	**Aklanon**	**Western Subanon**	**Maranao**	**Tboli**	**Ilokano**	**Dupaningan**	**Kapampangan**
NOM	*aŋ*	*ro*	*og*	*so*	∅	*ti*	∅	*iŋ* (sg) / *riŋ* (pl)
GEN	*naŋ*	*it* (indef) *ku* (def)	*nog*	*o*	∅	*ti*	∅ / *na* / *di*	*niŋ* (sg) / *riŋ* (pl)
OBL	*sa*	*sa*	*sog*	*sa* (indef) *ko* (def)	*beʔ*	*(i)ti*	*ha*	*kiŋ* (sg) / *kariŋ* (pl)
Non-NOM OBJECT	GEN/OBL	GEN	GEN	OBL	∅	OBL	OBL	*-ŋ* (sg/pl)

Almost all Philippine languages have a distinct set of case markers that are used with personal names (glossed here as PN), the few exceptions to this being Manide and related languages, which have apparently merged the personal case markers to their common counterparts (Lobel 2010). These often mark a singular vs plural distinction idiosyncratically, the plural also being used with a singular name to indicate the referent and their associates. A sample of the singular forms of the personal case markers is shown in Table 12.

Table 12. Personal case markers (singular)

	Tagalog/ Aklanon	**Northern Subanen**	**Maranao**	**Tboli**	**Ilokano**	**Dupaningan**	**Kapampangan**
NOM/ABS	*si*	*si*	*si*	∅	*ni*	*ni*	*i*
GEN/ERG	*ni*	*ni*	*i*	∅	*ni*	*ni*	*ri*
OBL	*kay*	*ni*	*ki*	∅	*ken ni*	*ha ni*	*kari*

In those languages that have conflated or lost case marking on full NPs, like Tboli, case distinctions are still made on pronominals.

In a number of languages, case is often expressed syncretically with other referential, temporal or syntactic features. Certain North Luzon languages such as Pangasinan and Ibaloy have two nominative markers for common NPs, one which is used in post-predicate position and another which is used in clause initial position (Pangasinan *so* and *say*, respectively, Orpilla & Kaufman 2024, Reid 2006). With regard to definiteness and referentiality, Waray employs three types of nominative and genitive case markers for full noun phrases: *ʔin* NOM indefinite, *ʔan* NOM past definite and *ʔit* NOM non-past definite, with genitive counterparts *hin, han, hit*, respectively (Zorc 1977:85). McFarland (1974) discusses similar specific/non-specific distinctions in the Legazpi Bikol case markers shown in (109) and (110). The (a) examples show that indefinite possessors and genitive agents are introduced by *ki* while definite ones are introduced by *kan.*

Legazpi Bikol (McFarland 1974:161)

(109)a. *aruŋ ki lala:ki*
house GEN.INDEF man
'a man's house'

b. *aruŋ kan lala:ki*
house GEN.DEF man
'the man's house'

Legazpi Bikol (McFarland 1974:161)

(110)a. *pig-bakal ki lala:ki.*
PV.BEG-buy GEN.INDEF man
'bought by a man'

b. *pig-bakal kan lala:ki*
PV.BEG-buy GEN.DEF man
'bought by the man'

Certain Bikol languages make even finer distinctions in specificity and definiteness in their case marking system (McFarland 1974:165) while other Philippine languages do not express referentiality distinctions at all in this manner. In these latter languages, the definiteness or referentiality of an argument may be largely predictable on the basis of grammatical function, with the pivot obtaining a definite reading and the actor voice object obtaining an indefinite one. The Sama languages (Kaufman 2024b) provide several examples of this type, one of which is seen in (111), where the undergoer voice clause assigns a definite interpretation to the undergoer while the actor voice clause assigns it an indefinite interpretation (among other semantic effects), despite the lack of case markers.[41]

Pangutaran Sama (Walton 1986:120)

(111)a. *∅-bonoʔ sultan bantaʔ-na*
UV-kill king enemy-3s.GEN
'The killed his enemy.'

b. *monoʔ sultan bantaʔ-na*
AV:kill king enemy-3s.GEN
'The kills/fights some of his enemies.'

Case markers are generally proclitic on the following NP, which they are syntactically associated with. In many languages, however, case markers display a ditropic pattern (Cysouw 2005), in which they *encliticize* to the preceding word, against their syntactic association. This can be seen in (112), where the case marker *din* takes on a reduced enclitic allomorph *=n*, when following a vowel final word.

Kankanaey (Allen 2014:93)

(112) *man-ʔoga=n anak*
man-ʔoga [din anak]
AV-cry NOM child
'The child cries.'

7.4.4 The positioning of enclitics

All Philippine languages possess enclitics whose positioning differs from full phrases. Pronominal and adverbial enclitics are typically placed in second position, following the first word and occasionally the first phrase within a clause-like syntactic domain (Kaufman 2010a). In languages like Maranao and Tagalog, pronominal arguments are in complementary distribution with full phrasal arguments, as seen in (113). When a potential clitic host precedes the predicate (in this case the progressive marker *diʔi*), a bound pronoun must typically attach to it, as shown in (113a), but this position is not available for full noun phrases, as shown in (113b).

[41] Himmelmann (2016) and Reid & Liao (2002:466) treat the Tagalog phrase marker *aŋ*, glossed NOMINATIVE here, as a definiteness marker of sorts without any inherent case features. Collins (2018), on the other hand, treats the same morpheme as a case marker without any inherent semantics at all. The fact that NP fragments with the nominative case marker always receive a referential interpretation (e.g. *dagaʔ!* 'a rat!' versus *aŋ dagaʔ!* 'the rat!') favors an analysis in which the case markers at least have some semantic features (Kaufman 2026).

Maranao (Kaufman 2010b: 136)

(113)a. *diʔi*[*=ako*] *ma-matiya*[*=*ako*] *sa kitab*
PROG=1SG.NOM AV-read=1SG.NOM OBL book
'I'm reading a book.'

b. *diʔi* [**so wataʔ*] *ma-matiya* [*so wataʔ*] *sa kitab.*
PROG NOM child AV-read NOM child OBL book
'The child is reading a book.'

In languages of this type, free pronouns are only used in predicate position, as independent fragments, or as fronted topics, all contexts in which the pronominal is initial in its clause or intonational unit. Interestingly, it seems that in most Philippine languages, there is no option to use free pronominals in the regular position of post-predicate arguments. In other words, when enclitic pronominals can be used, they must be used. The complementary distribution between free and clitic pronominals is exemplified by Tagalog below, where (114a) shows the pronominal in predicate position and (114b) shows the pronominal as an argument.

Tagalog

(114)a. *{ikaw / *ka} aŋ taːmaʔ* b. *taːmaʔ {ka / *ikaw}*
2S.NOM NOM correct correct 2S.NOM
'You're the correct one.' 'You're correct.'

Nonetheless, some languages, such as Cebuano, the long forms of the genitive and nominative pronouns show more syntactic freedom than in languages like Tagalog (Wolff 1966).

In the unique case of Iraya, a language of northern Mindoro, most pronominal arguments must appear clause-initially, as seen in (115).

Iraya (Reid 2017:34,27)

(115)a. *Nay ʔinəm-en ʔag sapaʔ ŋuna*
1SG.GEN drink-PV DEF water now
'I'm drinking the water now.'

b. *kawu nay malyag.*
2SG.NOM 1SG.GEN like
'I like you.'

This most likely comes about as a result of what Starosta, Pawley and Reid (1982) refer to as the process of "aux-axing" wherein clitics attach to a clause-initial auxiliary only for that auxiliary to be historically lost, thus stranding the clitics in clause-initial position where they were formerly unattested. This can be seen in several South Cordilleran languages where pronominal clitics alternate between first and second position in their syntactic domain. However, when they appear in first position, they carry the meaning of their former host, the now phantom auxiliary. For example, in Ibaloy, the position of the pronominals in initial position in (116) signals 'go and…', a meaning that can still be signalled with an auxiliary /ʔan/ that serves as an overt host for the enclitics when present.

Ibaloy (Ruffolo 2004:430)

(116) *moak ial'an ni inapoy*
mo=ak ʔi-ʔala-an ni ʔinapoj
2S.GEN=1S.NOM CV-get-BNF GEN rice
'you go and get me some cooked rice'

There are many co-occurrence constraints on pronominal clitics in Philippine languages with a fascinating variety of repair mechanisms and ordering patterns which cannot be discussed fully here (see Kaufman 2010a and references therein). Most Philippine languages avoid the simple combination of a first person singular genitive clitic with a second person singular nominative one. Other languages have more wide-ranging constraints that cover all types of first person genitive clitics with second person nominative ones. One repair strategy is to use a suppletive portmanteau clitic that expresses both the features of the genitive and nominative argument. This clitic often takes the form of *taka* or *takaw* in the Central Philippine languages, but is *kita* in Tagalog. A second strategy is to remove the offending first person features from the genitive pronoun, leaving a default third person clitic. Thus, in Kapampangan, where we expect the unattested *=ku=ka* (=1s.GEN=2s.NOM) we find *=na=ka* (=3s.GEN=2s.NOM) in the same meaning. Then there are languages like Ilokano, which deletes the genitive pronoun altogether, so that instead of the unattested clitic combination in **ay-ayat-en=ku=ka* (IMPRF~love-PV=1s.GEN=2s.NOM) 'I love you', we find *ay-ayat-en=ka* (IMPRF~love-PV=2s.NOM) in the same meaning. Finally, there are languages like Maranao, where the nominative must be expressed as a free pronoun when it co-occurs with a first person genitive clitic. Interestingly, even though the nominative argument takes the form of a free pronoun, *seka*, rather than the enclitic *ka*, the pronoun still appears within the second-position clitic cluster, as seen in (117), rather than appearing in the post-predicate position.

Maranao (Alonto et al. 2009:8)

(117) *oman=aken=**seka**=kiran* *ma-pan-[t]otol*
when=1s.GEN=2s.NOM=3p.OBL PV.POT-DIST-tell
'when I talk about you to them'

The relative ordering of clitics within the clitic cluster is determined by up to three factors: prosody (shorter precedes longer), case (genitive precedes nominative) and person (1st person precedes 2nd person precedes 3rd person). Different constraints are active in different languages, but if a particular domain is active, it will always follow the above scales. The relative order between pronominal clitics and adverbial clitics is highly variable across Philippine languages. In the North Luzon languages, pronominals tend strongly to precede adverbials while in Cebuano we find the opposite pattern. In Tagalog and several other languages in which syllable count determines overall order, we find adverbials sandwiched by pronominals, specifically, monosyllabic pronominals come first, then monosyllabic adverbs, then disyllabic adverbs and then disyllabic pronominals. The details of these patterns, especially how pronominals combine with adverbials, have been studied in depth for relatively few languages of the Philippines (Schachter 1973, McFarland 2001, Billings and Kaufman 2004, Peng and Billings 2008, Lee and Billings 2008, Liao 2004, Kaufman 2010a) and remain a rich topic of study for the interface between morphology, phonology and syntax.

Clitic doubling, where a pronominal clitic corefers to a full NP in the same clause, is found commonly in the Central Luzon and North Luzon subgroups but is very rare elsewhere in the Philippines. Tboli, an exception to this pattern on the southern edge of Mindanao, shows clitic doubling with certain preverbal elements, as seen in (118), where the second position clitic *le* doubles the nominative argument *kem dumu*.

Tboli (Forsberg 1992:63)

(118) *deŋ=le ma koyu kem dumu*
already=3PL.NOM AV.fetch wood PL companion
'The others already fetched some wood.'

In the North Luzon languages, clitic doubling is most often restricted to the (genitive) agent argument of non-actor voice clauses, as seen below, where *=na* doubles the full NP argument 'Maria' in (119) and *=to* doubles the full argument 'Margie' in (119).

(119) Ilokano (Rubino 1997:32)
na-kita=na=k ni maria
PRF-see=3S.GEN=1S.NOM GEN Maria
'Maria saw me.'

Kalanguya (Santiago 2015:4)

(120) *ʔayag-an=to=ak ni=hi=margie*
call-LV=3S.GEN=1S.NOM GEN=PERS=Margie
'Margie called me.'

When clitic doubling becomes obligatory in languages such as the ones above it takes on the appearance of verbal agreement. Liao (2005) examines Central Cagayan Agta with regard to the clitic vs. agreement question in such "intermediate" systems. Further along the cline of person marking grammaticalization we find Kapampangan, which requires clitic doubling/agreement for both genitive and nominative case participants, as in (121).

Kapampangan (Roswell 1983:21)

(121) *s<in>ese-∅=na=la niŋ=babai riŋ=pusa*
<PRF>feed-PV=3S.GEN=3P.NOM GEN=babai NOM.PL=pusa
'The woman fed the cats.'

Kitano (2008) observes that doubling the nominative/absolutive argument in Kapampangan correlates with a specific, individuated interpretation. Thus, a mass noun, like 'rice', might not be doubled even if it is definite.

9. Complex constructions

9.1 Finite complement clauses

All Philippine languages allow for finite clause complements, as in (122). The embedded clause is generally introduced with the linker and has all the hallmarks of a main clause predicate.

Tagalog

(122) *s<in>a:bi-Ø ko sa iyo na ga:~gaw-in niya bu:kas.*
<BEG>say-PV 1SG.GEN OBL 2SG LNK IPFV~do-PV 3SG.GEN tomorrow
'I told you that s/he will do (it) tomorrow.'

Existentials and negative existentials play a very important role in Philippine languages as they express propositions which would typically employ a variety of indefinite pronouns and quantificational expressions in non-Philippine languages (e.g. English *someone, anyone, nobody*). With an entity denoting complement, existentials and their negative counterparts give the expected meaning, as exemplified by a North Luzon and Central Philippine language in (123) and (124), respectively.

Ilokano

(123)a. *adda aso=k*
EXT dog=1s.GEN
'I have a dog.'

b. *awan ti aso=k*
NEG.EXT CORE dog=1s.GEN
'I have no dog.'

Tagalog

(124)a. *may a:so ako*
EXT dog 1s.NOM
'I have a dog.'

b. *wala? ako=ŋ a:so*
NEG.EXT 1s.NOM=LNK dog
'I have no dog.'

However, existentials can take verbal complements as well and here we find that the existential quantifies over what would be the nominative argument. Thus, with a simple intransitive predicate, the existential yields an interpretation that would require an indefinite pronoun subject in English, as seen in (125) and (126).

Ilokano

(125)a. *adda s<imm>aŋpet*
EXT <AV.PRF>arrive
'Someone arrived.'

b. *awan ti s<imm>aŋpet*
EXT CORE <AV.PRF>arrive
'Nobody arrived.'

Tagalog

(126)a. *may d<um>atiŋ*
EXT <AV.PRF>arrive
'Someone arrived.'

b. *wala? d<um>atiŋ*
NEG.EXT <AV.PRF>arrive
'Nobody arrived.'

Similarly, with non-actor voice predicates, the existential quantifies over whatever the nominative argument would be, as seen in the following examples.[42]

Ilokano

(127)a. *adda na-kita=k*
NEG POT.PRF-see=1s.GEN
'I saw someone.'

b. *awan ti na-kita=k*
NEG.EXT CORE POT.PRF-see=1s.GEN
'I didn't see anyone.'

Tagalog

(128)a. *may ni-lu:to?-∅ ako*
EXT BEG-cook-PV 1s.NOM
'I cooked something.'

b. *wala? ako=ŋ ni-lu:to?-∅*
NEG.EXT 1s.NOM=LNK BEG-cook-PV
'I didn't cook anything.'

[42] Note that Ilokano and Tagalog differ in that the existential assigns nominative case to the possessor in Tagalog but does not assign any case in Ilokano. The North Luzon languages in general function like Ilokano in this respect while most Central Philippine languages function like Tagalog.

9.2 Questions and interrogative complements

When the interrogative phrase is a noun phrase, a cleft-like construction is required where the interrogative is in the predicate position and the remainder of the clause is embedded in a nominative phrase, as shown in (129).

Tagalog

(129) *ano aŋ s<in>a:bi-Ø niya?*
what NOM <BEG>say-PV 3SG.GEN
'What did s/he say?'

The position of the interrogative in questions like (129) is generally indistinguishable from the position of the predicate, suggesting a lack of real interrogative movement in questions that target participants (specifically, 'who', 'what' and 'which' questions). Rather, the predicate argument structure is simply manipulated with the aid of the voice morphology to map the interrogative phrase to the predicate position, the natural home of the pragmatic focus (cf. Keenan 2008, Potsdam 2009 for deriving Philippine-type participant questions without movement).

As mentioned earlier, Central Philippine languages possess a position in the left periphery of the clause dedicated to focusing oblique phrases, which can include prepositional phrases as well as adjunct interrogatives such as 'when', 'where', 'how'. Second position clitics follow the fronted oblique phrase and the remainder of the clause typically employs the dependent form of the verb, if this paradigm happens to be preserved. Otherwise, the independent form is used, just as in a canonical clause without oblique fronting. The North Luzon languages, among other subgroups, lack such a position. In its place, these languages employ the locative voice in an extended sense. For instance, in (130), a 'when' question takes a nominalized complement to form a structure that would be translated literally as 'When is your giving to him?'.

Kankanaey (Allen 2014: 328)

(130) *pig?an di paŋ-i-dawt-a(n)=m en sisya?*
when NOM NMLZ-CV-give-LV=2S.GEN OBL 3S.NOM
'When are you going to give it to him?'

Ilokano (Rubino 1997:61)

(131) *sadino ti pag-adal-an=na?*
where ART GER-study-LV=3S.GEN
'Where does he study?' = 'Where is his place of study?'

In languages that employ this question formation strategy it is often ungrammatical to formulate a question such as (131) above in the actor voice. Ilokano often does away altogether with the 'where' interrogative, which is understood in common questions such as (132) purely on the basis of the locative voice.

Ilokano (Rubino 1997:6)

(132) *na-pan-an=na ti adiŋ=mo*
PRF.STA-go-LV=3S.GEN ART younger.sibling=2S.GEN
'Where did your younger sibling go?'

Interrogative complements are used in subordinate clauses as complements to matrix predicates of cognition as well as subjunctive type complements. These complements are typically identical to questions except that the interrogative phrase is introduced by a conditional marker, as in (133)-(135).

Central Tagbanwa (Scebold 2003:73)

(133) *pog-tuʔma iŋ kali ka nag-giʔit.*
IPFV.AV-ask HYP where 2SG.NOM PFV.AV-depart
'He is asking where you came from.'

Matigsalug Manobo (Wang et al. 2006:112)

(134) *Su mig-inse sikandan ke hendei key eg-pa-bulus*
so AV.PFV-ask 3PL.NOM if where 1PL.EX.NOM AV.PROG-CAUS-continue
'So they asked where we were going.'

Hanunoo (Epo 2014:22)

(135) *sabi-hun nimu sa kaŋku nu hayga*
tell-PV 2SG.GEN OBL 1SG.GEN COND why
'Tell me why (it's) that way.'

9.3 Nonfinite complement clauses

Clausal complementation with verbs of wanting, trying, and certain non-verbal predicates are typically non-finite and appear in a neutral form that does not indicate aspect, as shown in (136) and (137).

Cebuano

(136) *kinaháŋlan ni tibúʔ ŋa* ***táwg-un*** *aŋ páriʔ*
need GEN Tibo LNK call-PV NOM priest
'It is necessary for Tibo that a priest be called.' / 'Tibo needs to call a priest.'

Agutaynen (Quakenbush et al. 2010:13)

(137) *mambeŋ aŋ* ***mag-pa-layog*** *ta boradol*
fun LNK AV-CAUS-fly OBL kite
'It's fun to fly a kite.'

Note that many Philippine languages do not have a dedicated infinitive form. It seems that in most languages, the aspectually unmarked form may indicate a prospective or irrealis meaning in a finite main-clause predication. In Tagalog, and a number of other languages, the aspectually unmarked form is only found in non-finite contexts such as embedded clauses, imperatives and subjunctives.

Voice and the functions subsumed under 'mode' (e.g. potentive, causative, reflexive, etc.) are still present in most non-finite subordinate clauses. In some languages, aspect in the subordinate clause agrees with the matrix predicate in what are typically non-finite contexts for other languages, as shown in (138) and (139)

Agutaynen (Quakenbush et al. 2010:20)

(138) a. ***mam****-[p]ag-t<ar>abaŋ-an* *tanira=ŋ* ***maŋ****-ayeg*
PFV.PL-TR-<PL>help-LV 3PL.NOM=LNK AV.DIST-harvest
'They will help one another to harvest.'

b. ***nam****-[p]ag-t<ar>abaŋ-an* *tanira=ŋ* ***naŋ****-ayeg*
AV.PFV-PL-TR-<PL>help-LV 3PL.NOM=LNK AV.PFV.DIST-harvest
'They helped one another to harvest.'

Batad Ifugao (Newell 2005:21)

(139) a. ***um****-uy* ***maŋ****-anup*
AV-go AV-hunt
'will go hunting / to go hunting'

b. ***imm****-uy* ***naŋ****-anup*
AV.BEG.go AV.BEG-hunt
'went hunting'

Another type of complementation pattern, employed for various types of biclausal constructions, treats the subordinate predicate as a case marked complement, as shown in (140)-(142).

Tagalog

(140) *b<in>ilis-an ko aŋ pag-ka:ʔin*
<BEG>fast-LV 1SG.GEN NOM GER-eat
'I sped up my eating.'

Cebuano (Tanangkingsing 2009:36)

(141) *Si Junjun mi-sulay=pa sa pag-sagaŋ sa kutsilyo…*
P.NOM Junjun AV-try=still OBL GER-shield OBL knife
'Junjun tried to shield himself from the knife…'

Ilokano (Bloomfield 1942:197)

(142) *kayat=ku ti ag-laŋuy*
want=1s.GEN CORE AV-swim
'I want to swim.'

Various types of clause union or "restructuring" phenomena are attested throughout Philippine languages, especially with "pseudo-verbs" that take on an auxiliary function. For instance, pseudo-verbs in Batad Ifugao don't require subjects and can host genitive clitics associated with the main verb, as seen in (143). This appears to contrast with Ilokano, where the same type of auxiliary assigns nominative case like a regular actor voice verb, as in (144).

Batad Ifugao (Newell 2005:21)

(143) *um-uy=ʔu ala-n nan ba:ŋa*
AV-go=1s.GEN get-PV NOM pot
'I will go to get the pot.'

Ilokano (Rubino 1997:57)

(144) *in=ka=man* *pa-kan-en* *dagidiay* *baboy*
go=2s.NOM=please CAU-eat-PV those pig
'Please go feed those pigs.'

The syntax of restructuring and auxiliaries in Philippine languages has only been studied sporadically (see Kroeger 1993 and Mercado 2002 for Tagalog) and remains a rich topic for further research.

9.3.1 Control patterns

Control refers to coreference between an argument in a matrix clause and a missing argument in a (typically non-finite) subordinate clause. Many Philippine languages pattern as in (145) and (146), where an embedded agent co-referring with a matrix argument must be null.

Tagalog

(145) *gusto ko=ŋ* *tawa:g-an (*ko)* *si* *boboy*
want 1SG.GEN=LNK call-LV 1SG.GEN NOM Boboy
'I want to call Boboy.'

Bugkalot

(146) *ɣamak nima oɣta=n* *on-kadɨŋ*
want GEN deer=LNK AV-cross
'The deer wants to cross.'

Conversely, the agent of a subordinate non-finite clause must be overt when it does not co-refer with a matrix clause argument, as in (147).

Tagalog

(147) *gusto ko=ŋ* *tawa:g-an* *mo* *ako*
want 1SG.GEN=LNK call-LV 2SG.GEN 1SG.NOM
'I want you to call me.'

The volitionality of the subordinate predicate determines which argument can be controlled, as seen in the minimal pair in (89) (cf. Ceña 1977; Kroeger 1993; Schachter 1976).

Tagalog

(148)a. *gusto ko=ŋ* *tawa:g-an*
want 1SG.GEN=LNK call-LV
'I want to call (someone).'

b. *gusto ko=ŋ* *ma-tawa:g-an*
want 1SG.GEN=LNK STA-call-LV
'I want to be called.'

This seems to hold true for at least the Central Philippine subgroup although this type of data is generally lacking for other subgroups.

Ilokano (Bloomfield 1942:197-198)

(149)a. *kayat=ku* *ti* *ag-laŋuy*
want=1s.GEN CORE AV-swim
'I want to swim.'

b. *kayat=na=k ŋa ag-gubernadur*
want=3s.GEN=1s.NOM LNK AV-governor
'He wanted me to be a governor.'

Overall, the use of dependent forms in control constructions is rare although it is attested. In Bugkalot, shown in (150), replacing the dependent form of the verb *iɣap-a* with its independent counterpart, *iɣap-en*, is judged ungrammatical. Conversely, only the independent form is allowed in a plain declarative main clause, as in (150b).

(150) Bugkalot
a. *ɣamak=ko=n iɣap-a ima tawen*
want=1s.GEN=LNK see-PV.DEP NOM sky
'I want to see the sky.'
b. *iɣap-ən=ko ima tawen*
see-PV.INDEP=1s.GEN NOM sky
'I will see the sky.'

9.4 Temporal adjuncts

Temporal adjuncts are often formed via nominalization in Philippine languages (Kaufman 2011). A typical structure is shown in Sarangani Manobo (151), which displays the combination of the gerundive *peg-* with the lack of a nominative case on either of the arguments.

Sarangani Manobo (DuBois 1976:94)
(151) *peg-dineg te amay din kenyan*
GER-hear GEN father 3SG.GEN that.OBL
'When his father heard that.'

It appears that an earlier gerundive marker *paR- has degrammaticalized in a number of languages as an independent temporal subordinator in its own right. The prefixal gerundive and independent uses can be seen side by side in (152).

Tagalog
(152)a. *pag-alis=nila*
GER-leave=3s.GEN
'when they leave'
b. *pag hindiː=na sariːwaʔ*
when NEG=already fresh
'when it is no longer fresh'

Fully finite temporal adjuncts, as exemplified in (153), are also found in most languages. Here, the subordinator *idi* introduces a verbal clause in the perfective aspect unlike the above cases where there is no overt subordinator and the verb lacks aspectual morphology.

Ilokano (Clausen 1995:36)
(153) *maŋ~maŋan=kami idi s<imm>aŋpet ni Mel*
AV.PROG~eating=1p.EX.NOM PT <AV.PRF>arrive P.NOM Mel
'We (excl) were eating when Mel arrived.'

10. Conclusion

This chapter has given a broad overview of the phonology, morphology and syntax of the Philippine languages while focusing on several phenomena of interest that are characteristic of the region. I have also attempted to highlight areas in need of further research. The Philippines is host to roughly 177 languages only some of which have been fully described. While most languages share certain typological similarities (Himmelmann 2005), it may be the case that these similarities were also characteristic of PMP and that languages outside the region underwent typological shifts due to contact with non-Austronesian languages. The other possibility is that the Philippines represents a contact zone in which certain grammatical and lexical elements have spread over centuries, despite the rather clear evidence for subgroups at various levels among Philippine languages. As Blust (2019, 2022) presents a new argument for the unity of a Philippine subgroup based on shared lexical innovations, it becomes even more urgent to understand the dual roles of contact and inheritance in the historical formation of Philippine languages, as internal levelling and contact effects continue apace.

References

Abrams, Norman. 1963. Historical development of Bilaan vowels and some consonant reflexes in Bilaan and related dialects. *Philippine Sociological Review* 11: 147-54.

Abreu, Marvin M. 2018. A Reference Grammar of Southern Alta. Ph.D. diss. De La Salle University.

Adams, Karen L. and Alexis Manaster-Ramer. 1988. 'Some questions of topic/focus choice in Tagalog', *Oceanic Linguistics* 27(1/2): 79–101.

Aldridge, Edith. 2004. 'Ergativity and Word Order in Austronesian Languages', PhD thesis, Cornell University.

Allen, Janet L. 2014. Kankanaey: A Role and Reference Grammar Analysis. Dallas: SIL International, Global Publishing.

Alonto, Almahdi G., Abdullah B. Adam, R. David Zorc. 2009. *Maranao Dialogs and Drills*. Hyattsville, MD: Dunwoody Press.

Anderson, Barbara B. 1974. Linguistic differentiation in central Bicolandia. *Philippine Journal of Linguistics*. 4-5(1-2): 55-71.

Antworth, Evan L. 1979. *A Grammatical Sketch of Botolan Sambal*. Philippine Journal of Linguistics Special Monograph Issue Number 8 (LSP 8). Manila: Linguistic Society of the Philippines.

Arms, D. G. 1996. 'Categories of the Sindangan Subanen verb', in Hein Steinhauer (ed.), *Papers in Austronesian Linguistics No. 3*, Canberra: Pacific Linguistics, 1-32.

Awed, Silin A., Lillian Underwood, and Vivian M. van Wynen. 2004. *Tboli-English dictionary*. Manila: SIL.

Bardají i Farré, Maria, Sonja Riesberg & Nikolaus P. Himmelmann. 2022. Limited control predicates in western Austronesia: stative, dynamic, or none of the above? *Oceanic Linguistics* 61(1): 118-165.

Bautista, Ma. Lourdes. 1975. A model for bilingual competence based on an analysis of Tagalog-English code switching. *Philippine Journal of Linguistics*, 8 (1), 57-89. Retrieved from https://animorepository.dlsu.edu.ph/faculty_research/10209

Bautista, Ma. Lourdes, ed. 1996. *Readings in Philippine sociolinguistics*. Manila: De La Salle University Press.

Bautista, Ma. Lourdes, 1998. Another look at Tagalog-English code-switching. *Pagtanaw: Essays on language in honor of Teodoro A Llamzon*, pp.121-146.

Bell, Sarah J. 1978. 'Two Differences in Definiteness in Cebuano and Tagalog', *Oceanic Linguistics* 17(1): 1-9.

Bennásar, Guillermo. 1892. *Observaciones gramaticales sobre la lengua Tiruray*. Manila: Imp. y litog. de M. Pérez.

Benton, Richard A. 1971. 'Phonotactics of Pangasinan', PhD thesis, Department of Linguistics, University of Hawaii, Mānoa.

Billings, Loren and Daniel Kaufman. 2004. "Towards a typology of Austronesian pronominal clisis." In Paul Law (ed.), *Proceedings of AFLA* [Austronesian Formal Linguistics Association] *11* (ZAS Papers in Linguistics 34). Berlin: Zentrum für Allgemeine Sprachwissenschaft. 15–29.

Bloomfield, Leonard. 1917. *Tagalog texts with grammatical analysis*. Urbana, Illinois: University of Illinois.

Blust, Robert. 1971. 'A Tagalog consonant cluster conspiracy', *Philippine Journal of Linguistics* 2: 85-91.

Blust, Robert. 1972. Note on PAN *qa(R)(CtT)a "outsiders, alien people". *Oceanic Linguistics* 11: 166-171.

Blust, Robert. 1979. Coronal-noncoronal consonant clusters: New evidence for markedness. *Lingua* 47: 101–17.

Blust, Robert. 1991. 'The Greater Central Philippines hypothesis'. *Oceanic Linguistics* 30: 73-129.

Blust, Robert. 1992. 'On speech strata in Tiruray' in Malcolm Ross (ed.), *Papers in Austronesian Linguistics, No. 1*. Canberra: Pacific Linguistics, 1-52.

Blust, Robert. 2002. 'Notes on the history of focus in Austronesian languages', in Fay Wouk and Malcolm Ross (eds.), *The History and Typology of Western Austronesian Voice Systems*. Canberra: Pacific Linguistics, 63–80.

Blust, Robert. 2004. 'Austronesian nasal substitution: a survey', *Oceanic Linguistics* 43:73–148.

Blust, Robert A. 2005. The linguistic macrohistory of the Philippines: Some speculations. In Liao, H.-C. and Rubino, C. R. G., (eds.), *Current issues in Philippine linguistics and anthropology: Parangal kay Lawrence A. Reid*, pp. 31–68. Linguistic Society of the Philippines and SIL, Manila, Philippines.

Blust, Robert. 2005. 'A note on the history of genitive marking in Austronesian languages', *Oceanic Linguistics* 44: 215-22.

Blust, Robert. 2013. *The Austronesian Languages,* revised edition. Canberra: Asia-Pacific Linguistics.
Blust, Robert. 2015. 'The case-markers of Proto-Austronesian', *Oceanic Linguistics*. 54(2): 436-491.
Blust, Robert. 2018. 'The "Mystery Aspirates" in Philippine Languages', *Oceanic Linguistics* 57(1):221–247.
Blust, Robert. 2019. 'The resurrection of Proto-Philippines', *Oceanic Linguistics* 58(2): 153-256.
Blust, Robert, Stephen Trussel, & Alexander D. Smith. 2023. *CLDF dataset derived from Blust's "Austronesian Comparative Dictionary"* (v1.2) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.7741197
Blust, Robert. 2022. Proto-Philippine Addenda: Theory, Method and Data. *Oceanic Linguistics*, 61(1): 322–404.
Brainard, Sherri. 1997. 'Ergativity and Grammatical Relations in Karao', in T. Givón (ed), *Grammatical Relations: A Functional Perspective*. Amsterdam and Philadelphia: John Benjamins, 85-154.
Bulalang, Sharon. 2025. *Western Subanon Grammar*. Leiden: Brill.
Burton, Scott L. 1996. 'A case study of lexical borrowing between language families: the East Mindanao and Manobo languages', PhD thesis, University of Texas at Arlington.
Burton, Scott L. 2018. *Tagakaulo-English Dictionary*. Manila: SIL Philippines.
Ceña, Resty M. 1977. 'Patient primacy in Tagalog', Paper presented at the LSA Annual Meeting, Chicago.
Ceña, Resty M. 2012. *Sintaks ng Filipino*. Manila: National Commission for Culture and the Arts.
Ceña, Resty M. and Ricardo M. D. Nolasco. 2011. *Gramatikang Filipino: Balangkasan*. Diliman: University of the Philippines Press.
Charles, Mathew. 1974. Problems in the reconstruction of Proto-Philippine phonology and the subgrouping of the Philippine languages. *Oceanic Linguistics* 13(1/2):457–509.
Chen, Victoria, Maria Kristina Gallego, Jonathan Kuo, Isaac Stead and Benjamin van der Voorn. 2026. "Contact or inheritance? New evidence on the Proto-Philippine hypothesis." *Oceanic Linguistics* 65(1): 58-93.
Chen, Victoria. 2017. A reexamination of the Philippine-Type voice system and its implications for Austronesian primary level subgrouping. PhD Diss., University of Hawai'i.
Collins, James N. 2018. 'Definiteness determined by syntax: A case study in Tagalog', *Natural Language and Linguistic Theory*.
Constantino, Ernesto. 1977. *Selected Writings of Cecilio Lopez in Philippine Linguistics*. Diliman, Quezon City: University of the Philippines Press.
Cubar, Nelly I., & Cubar, Ernesto H. 1994. *Writing Filipino Grammar: Traditions & Trends*. New Day Publishers.

Cysouw, Michael. 2005. “Morphology in the Wrong Place: A Survey of Preposed Enclitics” in *Morphology and Its Demarcations*, Dressler, Wolfgang U. and Kastovsky, Dieter and Pfeiffer, Oskar E. and Rainer, Franz (eds.), pp. 17-37. Amsterdam: John Benjamins.

Daguman, Josephina S. 2004. ‘A grammar of Northern Subanen’, PhD thesis, La Trobe University.

De Guzman, Videa P. 1978. ‘Syntactic derivation of Tagalog verbs’, *Oceanic Linguistics Special Publication No. 16*. Honolulu: University of Hawaii Press.

De Guzman, Videa P. 1988. ‘Ergative analysis for Philippine languages: An analysis’, in R. McGinn (ed.), *Studies in Austronesian Linguistics*. Athens, OH: Ohio University Center for Southeast Asia Studies, 323-345.

Dell, François. 1983. An aspectual distinction in Tagalog. *Oceanic Linguistics* 22–23:175–206.

Dempwolff, Otto. 1934-38. *Vergleichende Lautlehre des Austronesischen Wortschatzes*. Berlin: Deitrich Reimer.

Dita, Shirley. 2010. A Reference Grammar of Ibanag: Phonology, Morphology, & Syntax. Saarbrücken: Lambert Academic Publishing.

Donohue, Mark 2007. ‘Word order in Austronesian from north to south and west to east’, *Linguistic Typology* 11: 349–391.

Dubois, Carl .D. 1976. *Sarangani Manobo: an introductory guide*. Manila: Linguistic Society of the Philippines.

Eberhard, David M., Gary F. Simons, and Charles D. Fennig (eds.). 2025. Ethnologue: Languages of the World. Twenty-eighth edition. Dallas, Texas: SIL International. Online version: http://www.ethnologue.com.

Elkins, Richard. 1986. *Manobo-English Dictionary*. Honolulu: University of Hawai'i Press.

Erlewine, M. Yoshitaka., Levin, Theodore, and Coppe van Urk. 2017. Ergativity and Austronesian-type voice systems. In Lisa Travis, Jessica Coon, and Diane Massam, editors, *Oxford Handbook of Ergativity*. Oxford University Press.

Epo, Yrrah Jane S. 2014. ‘Discourse analysis of suyot: a Hanunuo-Mangyan folk narrative.’ MA thesis. Payap University.

Estrera, Edward G. 2024. Bagobo-Klata: grammar and vocabulary. MA thesis. University of the Philippines-Diliman.

Fleischman, Eric. 1981. ‘The Danao Languages: Magindanaon, Iranun, Maranao, and Illanun’, *Philippine Journal of Linguistics* 12: 57-78.

Fleming, Hannah M. 2022. *A Grammar of Eastern Tawbuid*. SIL e-Books 79.

Foley, William A. 2008. ‘The place of Philippine languages in a typology of voice systems’, in Peter K. Austin and Simon Musgrave (eds.), *Voice and Grammatical Relation in Austronesian Languages*. Stanford: CSLI Publications, 22-44.

Foley, William A. 2017. “Structural and semantic dependencies in word class” in *Dependencies in Language: On the Causal Ontology of Linguistic Systems*, N.J. Enfield (ed.). pp.179-194. Berlin: Language Science Press.

Foley, William A. 2023. Word-Class Systems and Other Grammatical Properties. In *The Oxford Handbook of Word Classes*, ed. by Eva van Lier, pp. 109–134. Oxford: Oxford University Press.

Forman, Michael L. 1971. *Kapampangan Grammar Notes*. PALI Language Texts: Philippines. Honolulu: University of Hawaii Press.

Forman, Michael L. 1972. Zamboangueño texts with grammatical analysis: a study of Philippine Creole Spanish. Ithaca, NY: Ithaca: Cornell University dissertation.

Forsberg, Vivian M. 1992. 'A Pedagogical Grammar of Tboli', *Studies in Philippine Lingiustics* 9: 1-110.

Fukuda, Takashi. 1997. 'A Discourse-oriented Grammar of Eastern Bontoc', *Studies in Philippine Linguistics* 10(1). Manila: Linguistic Society of Manila, Summer Institute of Linguistics.

Gallman, II, Andrew Franklin. 1983. 'Proto East Mindanao and Its Internal Relationships', PhD thesis, University of Texas at Arlington.

Gallego, Maria Kristina. 2022. The stratigraphy of a community: 150 years of language contact and change in Babuyan Claro, Philippines. Ph.D. dissertation, Australian National University.

Garcia, Lydia Gonzales. 1992. *Mga Gramatikang Tagalog/Pilipino 1893-1977*. Diliman: Sentro ng Wikang Filipino.

Garcia Laguia, Alexandro-Xavier. Documentation of Northern Alta: grammar, texts and glossary. Ph.D. diss. Universitat de Barcelona.

Gerdts, Donna B. 1988. 'Antipassives and causatives in Ilokano: Evidence for an ergative analysis', in R. McGinn (ed.), *Studies in Austronesian Linguistics*. Athens, OH: Ohio University Center for Southeast Asia Studies, 295-321.

Gieser, C. Richard. 1963. *A grammatical sketch of Kalinga*. Nasuli, Malaybalay, Bukidnon: Summer Institute of Linguistics.

Gonzalez, Andrew. 1981. *Pampangan: Towards a Meaning-based Description*. Pacific Linguistics C-48. Canberra: Australian National University.

Gonzalez, Andrew. 1985. Sociolinguistics in the Philippines. *Southeast Asian Journal of Social Science* 13(1): 52-60.

Hauk, Bryn. 2019. Functions of the Subanon mo-prefix: Evidence from paradigms and argument structure. *Oceanic Linguistics* 58(2): 257–91.

Headland, Thomas N. 2003. "Thirty endangered languages in the Philippines," Work Papers of the Summer Institute of Linguistics, University of North Dakota Session: Vol. 47, Article 1.DOI: 10.31356/silwp.vol47.01

Headland, Thomas N. 2010. Why the Philippine negrito languages are endangered. In *Endangered Languages of Austronesia*, M. Florey, ed. Oxford University Press, 110–118.

Headland, Thomas N. and Elmer P. Wolfenden. 1967. The Vowels of Casiguran Dumagat. pp. 592-596 in Studies in Philippine Anthropology (Mario D. Zamora, ed.) Alemar-Phoenix, Quezon City.

Headland, Thomas N. and Healey, Alan. 1974. Grammatical Sketch of Casiguran Dumagat. In *Papers in Philippine Linguistics* 6, 1-54. Canberra: Pacific Linguistics.

Himes, Ronald S. 1998. 'The Southern Cordilleran Group of Philippine Languages', *Oceanic Linguistics* 37(1):120–177.

Himes, Ronald S. 2012. 'The Central Luzon Group of Languages', *Oceanic Linguistics* 51(2):490-537.

Himmelmann, Nikolaus P. 1991. 'The Philippine challenge to Universal Grammar', Arbeitspapier Nr. 15. Köln: Institut für Sprachwissenschaft, Universität zu Köln.

Himmelmann, Nikolaus P. 2005. 'The Austronesian languages of Asia and Madagascar: Typological characteristics', in K. Alexander Adelaar and Nikolaus P. Himmelmann (eds.), *The Austronesian languages of Asia and Madagascar*. London: Routledge, 110-181.

Himmelmann, Nikolaus P. 2006. How to miss a paradigm or two: Multifunctional ma- in Tagalog. In *Catching Language*, ed. by Felix Ameka, Alan Dench, and Nicholas Evans, 487–526. Berlin: Mouton de Gruyter

Himmelmann, Nikolaus P. 2008. "Lexical categories and voice in Tagalog". In *Voice and Grammatical Functions in Austronesian Languages*, Peter K. Austin and Simon Musgrave (eds.), pp.247–293. Stanford: CSLI Publications.

Himmelmann, Nikolaus P. 2016. 'Notes on 'noun phrase structure' in Tagalog', in Jens Fleischhauer, Anja Latrouite, and Rainer Osswald (eds.), *Explorations of the Syntax-Semantics Interface*. Düsseldorf: Dup, 319-341.

Himmelmann, Nikolaus P. and John Ulrich Wolff. 1999. *Toratán (Ratahan)*. Languages of the World/Materials 130. Munich: Lincom Europa.

Hohulin, E. Lou, and Donald Burquest. 2014. *Tuwali Ifugao Dictionary and Grammar Sketch*. Linguistic Society of the Philippines Special Monograph Issue No.57 (LSP 57). Manila: Linguistic Society of the Philippines.

Hohulin, Richard M., B. Lou Hohulin, Alberto K. Maddawat. 2018. Keley-i Dictionary and Grammar Sketch. Webonary.org. SIL International. Retrieved 4/6/25.

Hopper, Paul J. and Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 56(2):251-299.

Humboldt, Wilhelm von. 1836. Über die Kawi-Sprache auf der Insel Java, nebst einer Einleitung über die Verschiedenheit des menschlichen Sprachbaues und ihren Einfluß aud die geistige Entwickelung des Menschengeschelchtes. Erster Band. Berlin: Königliche Akademie der Wissenschaften.

Hsieh, Henrison. 2016. "Prosodic indicators of phrase structure in Tagalog transitive sentences". In AFLA 23: *The Proceedings of 23rd Meeting of the Austronesian Formal Linguistics Association*, Hiroki Nomoto, Takuya Miyauchi, Asako Shiohara (eds.), pp. 111-122. Canberra: Asia-Pacific Linguistics.

Hsieh, Henrison. 2019. Distinguishing nouns and verbs: A Tagalog case study. *Natural Language and Linguistic Theory* 37:523-569.

Hsieh, Henrison. 2023. Observations on Tagalog Genitive Inversion. *Oceanic Linguistics* 62(1):242-266.
Johnson, Rex E. 1996. *A Bibliography of Philippine Linguistics*. Manila: Linguistic Society of the Philippines.
Clausen, Josie Paz. 1995. The taxonomy, semantics, and syntax of Ilokano adverbial clauses. Ph.D. diss. University of Hawai'i at Mānoa.
Kaufman, Daniel. 2005. 'Aspects of pragmatic focus in Tagalog', in I Wayan Arka and Malcolm Ross (eds.), *The many faces of Austronesian voice systems: Some new empirical studies*. Canberra: Pacific Linguistics, 175–196.
Kaufman, Daniel. 2009a. Inner causatives and the reconstruction of the Proto-Austronesian affixes *pa-, *pa and *pa<ŋ>-. International Conference on Austronesian Linguistics XI, Aussois, France.
Kaufman, Daniel. 2009b. "Austronesian nominalism and its consequences: A Tagalog case study". *Theoretical Linguistics* 35(1):1-49.
Kaufman, Daniel. 2009c. Austronesian typology and the nominalist hypothesis. In Adelaar, A. and Pawley, A., editors, Austronesian historical linguistics and culture history: A festschrift for Robert Blust, Pacific Linguistics, pp. 179–226. Canberra: ANU Press.
Kaufman, Daniel. 2010a. 'The morphosyntax of Tagalog clitics: A typological approach', PhD thesis, Cornell University.
Kaufman, Daniel. 2010b. 'The grammar of clitics in Maranao', in Loren Billings and Nellke Goudswaard (eds.), *Piakandatu ami Dr. Howard P. McKaughan*. Manila: Linguistic Society of the Philippines and SIL, 132-157.
Kaufman, Daniel. 2011a. 'Exclamatives and temporal nominalizations in Austronesian', in Foong Ha Yap, Karen Grunow-Hårsta, and Janick Wrona (eds.), *Nominalization in Asian Languages: Diachronic and typological perspectives*. Amsterdam: John Benjamins, 721-754.
Kaufman, Daniel. 2011b. Interpreting the geography of Tense, Aspect Mood marking across Austronesian. Invited talk. International Workshop on TAM and Evidentiality in Indonesian Languages. Tokyo University of Foreign Studies. URL: https://bahasawan.com/wp-content/uploads/2020/01/The-geography-of-TAM-marking-H.pdf
Kaufman, Daniel. 2012. Predicate classes and Proto-Austronesian *ka-. Keynote talk. Austronesian Formal Linguistics Association 19 (AFLA 19), Academia Sinica, Taiwan. URL: https://bahasawan.com/wp-content/uploads/2020/01/Predicate-classes-and-PAn-ka-H.pdf
Kaufman, Daniel. 2017. 'Lexical category and alignment in Austronesian', in Lisa Travis, Jessica Coon, and Diane Massam (eds.), *Oxford Handbook of Ergativity*. Oxford: Oxford University Press.
Kaufman, Daniel. 2018a. 'Austronesian predication and the emergence of biclausal clefts in Indonesian languages', in Sonja Riesberg, Asako Shiohara, and Atsuko Utsumi (eds.), *Perspectives on information structure in Austronesian languages*. Berlin: Language Science Press, 207-245.
Kaufman, Daniel. 2018b. 'Inner and outer causatives in Austronesian: A diachronic perspective'. *McGill Working Papers in Linguistics* 25(1): 201–215.

Kaufman, Daniel and Nikolaus P. Himmelmann (2024). 'Supersegmental Phonology'. In *The Oxford Guide to the Malayo-Polynesian Languages of Southeast Asia*, Alexander Adelaar and Antoinette Schapper (eds.), pp. 703-718. Oxford: OUP.

Kaufman, Daniel. 2024a. 'The languages of the southern and central Philippines' In *The Oxford Guide to the Malayo-Polynesian Languages of Southeast Asia*, Alexander Adelaar and Antoinette Schapper (eds.), pp. 347-373. Oxford: OUP.

Kaufman, Daniel. 2024b. 'Sama-Bajaw languages'. In *The Oxford Guide to the Malayo-Polynesian Languages of Southeast Asia*, Alexander Adelaar and Antoinette Schapper (eds.), pp. 375-387. Oxford: OUP.

Kaufman, Daniel. 2024c. Resegmentation and the fecundity of Philippine morphology. International Conference on Austronesian Linguistics XVI, De La Salle University, Manila.

Kaufman, Daniel. 2025. Austronesian prosody and contrast preservation. In Alessa Farinella, Kristine Yu and Jed Pizarro-Guevara (eds.). Proceedings of the Austronesian Formal Linguistics Association (AFLA 31). University of Massachusetts, Amherst.

Kaufman, Daniel. 2026a. Case marking determiners and the role of register in Philippine NP-interpretation. In *Philippine Languages and Education Reform: Papers in Honor of Ricardo Ma Duran Nolasco*, Mercedes E. Arzadon, Antonio D. Igcalinos and Thomas E. Payne (eds.), pp.125-154. Sorsogon City: Sorsogon University Press.

Kaufman, Daniel. 2026b. 'Subgrouping'. In: *The Wiley Blackwell Companion to Diachronic Linguistics*, Adam Ledgeway, Alexandra Simonenko, Anne Breitbarth, Edith Aldridge, Joseph Salmons, Katalin É. Kiss (eds.), pp.1-36. Malden, MA: Wiley-Blackwell.

Kaufman, Daniel. 2026c. A global view on the Ausrichtungsaffix and the deep differences between nominalization-derived versus applicative-derived diathesis in Austronesian. In: *The Documentarist Turn: From observable linguistic behaviour to typological generalizations*, Sonja Riesberg, Uta Reinöhl and Birgit Hellwig (eds.), pp. 518–571. Amsterdam: John Benjamins.

Kaufman, Daniel, Nathan Adamson, Victoria Chen, Bradley McDonnell, and Olivia Waring. 2026. Austronesian: Tagalog and Indonesian. In *Conversion in Morphology*, Laurie Bauer and Salvador Valera (eds.), pp. 193-211. Oxford: Oxford University Press.

Keenan, Edward L. 2008. Voice and relativization without movement in Malagasy. *Natural Language and Linguistic Theory* 26, 467–497. https://doi.org/10.1007/s11049-008-9049-x

Kerr, Harland. 1988. 'Cotabato Manobo Grammar', *Studies in Philippine Linguistics* 7: 1–123.

Kimoto, Yukinori. 2017. A Grammar of Arta: A Philippine Negrito Language. (Doctoral dissertation, Kyoto.

King, Benedict, Simon J. Greenhill, Lawrence A. Reid, Malcolm Ross, MaryWalworth and Russell D. Gray. 2024. Bayesian phylogenetic analysis of Philippine languages supports a rapid migration of Malayo-Polynesian languages. *Scientific Reports* **14.**

Kitano, Hiroaki. 2008. Transitivity and Pronominal Clitic Order in Kapampangan. *Studies in Philippine Languages and Cultures* 17:88–97.

Kroeger, Paul R. 1990. Stative aspect and unaccusativity in Kimaragang Dusun. *Oceanic Linguistics* 29:110–31.
Kroeger, Paul R. 1993. *Phrase structure and grammatical relations in Tagalog*. Stanford: CSLI Publications.
Kroeger, Paul R. 2005. 'Kimaragang', in Alexander Adelaar and Nikolaus Himmelmann (eds.), *The Austronesian Languages of Asia and Madagascar*. London & New York: Routledge, 397-428.
Latrouite, Anja. 2011. 'Voice and Case in Tagalog: The coding of prominence and orientation', PhD thesis, University of Dusseldorf.
Lee, Celeste [Chia Yen], & Loren Billings. 2008. "Clitic-pronoun clusters in Central Philippine." Wilaiwan Khanittanan & Paul Sidwell (eds.), *SEALS XIV: Papers from the 14th meeting of the Southeast Asian Linguistic Society*, pp.193–203. Canberra: RSPAS, ANU.
Lendoyro, Constantino. 1909. *The Tagalog Language*. Manila: Juan Fajardo.
Liao, Hsiu-chuan. 2004. 'Transitivity and ergativity in Formosan and Philippine languages', PhD thesis, Department of Linguistics, University of Hawai'i at Mānoa.
Liao, Hsiu-chuan. 2005. 'Pronominal Forms in Central Cagayan Agta: Clitics or Agreement Features', in Hsiu-chuan Liao and Carl R. Galvez Rubino (eds.), *Current Issues in Philippine Linguistics and Anthropology: Parangal kay Lawrence A. Reid*. Manila: Linguistic Society of the Philippines and Summer Institute of Linguistics, 346-363.
Liao, Hsiu-chuan. 2008. 'A Typology of First Person Dual Pronouns and Their Reconstructibility in Philippine Languages', *Oceanic Linguistics* 47(1):1-29.
Liao, Hsiu-chuan. 2011a. 'On the Development of Comitative Verbs in Philippine Languages', *Language and Linguistics* 12(1):205-237.
Liao, Hsiu-chuan. 2011b. 'Some morphosyntactic differences between Formosan and Philippine languages', *Language and Linguistics* 12: 845-876.
Liao, Hsiu-chuan. 2011c. "On the reconstruction of Proto-Malayo-Polynesian *si- 'simultaneous/ concomitant aspect'", paper read at the 20th International Conference on Historical Linguistics (ICHL2011), National Museum of Ethnology (MINPAKU), Osaka, Japan, July 25-30, 2011.
Liao, Hsiu-chuan. 2020. A Reply to Blust (2019) "The Resurrection of Proto-Philippines." *Oceanic Linguistics* 59(1/2): 426–449.
Liao, Hsiu-chuan and Lawrence A. Reid. 2024. Languages of the Northern Philippines. In *The Oxford Guide to the Malayo-Polynesian Languages of Island Southeast Asia*, Alexander Adelaar and Antoinette Schapper (eds.), pp.325-346. Oxford: OUP.
Lobel, Jason W. 2004. 'Old Bikol -um- vs. mag- and the loss of a morphological paradigm', *Oceanic Linguistics* 43: 469-497.
Lobel, Jason W. 2010. 'Manide: An undescribed Philippine language'. *Oceanic Linguistics*. 49:478–510.

Lobel, Jason W. 2013. ‘Philippine and North Bornean languages: issues in description, subgrouping, and reconstruction’, PhD thesis, Department of Linguistics, University of Hawai’i at Mānoa.
Lobel, Jason W. 2026. Bantayanon: The complex history of an underdocumented Central Philippine language. In Arzadon, Mercedes E., Antonio D. Igcalinos, and Thomas E. Payne (eds). 2026. *Philippine languages and education reform: Papers in honor of Ricardo Ma Duran Nolasco*, PAGES. Sorsogon City: Sorsogon State University Press.
Lobel, Jason W. and Labi Hadji Sarip Riwarung. 2009. ‘Maranao revisited: An overlooked consonant contrast and its implications for lexicography and grammar’, *Oceanic Linguistics* 48(2):403–438.
Lobel, Jason W. and Labi Hadji Sarip Riwarung. 2011. ‘Maranao: a preliminary phonological sketch with supporting audio’, *Language Documentation and Conservation* 5: 31-59. Available online at http://hdl.handle.net/10125/4487/.
Lobel, Jason W. and William C. Hall. 2010. ‘Southern Subanen Aspiration’, *Oceanic Linguistics* 49(2):319–338.
Lobel, Jason William. 2013. *Philippine and North Bornean Languages: Issues in Description, Subgrouping, and Reconstruction*. Ph.D. dissertation, University of Hawai‘i at Mānoa.
Lobel, Jason William, and Orlando Vertudez Surbano (2019). ‘Notes from the Field: Remontado (Hatang-Kayi): A Moribund Language of the Philippines’, *Language Documentation & Conservation* 13:1-35.
Lobel, Jason, Wilmer Tria, and Jose Carpio. 2000. *An satuyang tataramon / A study of the Bikol language*. Holy Rosary Minor Seminary: Naga City, Philippines.
Lynch, John, Malcolm Ross, and Terry Crowley (2002). *The Oceanic Languages*. London/New York: Routledge.
Maree, Rundell D. 2007. *Ibatan: A Grammatical Sketch of the Language of Babuyan Claro Island*. LSP Special Monograph No.53. Manila: Linguistic Society of the Philippines.
Mattes, Veronika. 2014. *Types of reduplication: A case study of Bikol*. Berlin: Walter de Gruyter.
McFarland, Curtis D. 1974. ‘The Dialects of the Bikol Area’, PhD thesis, Yale University.
McFarland, Curtis D. 1978. ‘Definite objects and subject selection in Philippine languages’, in Casilda Edrial-Luzares and Austin Hale (eds.), *Studies in Philippine Linguistics, Vol. 2*. Manila: Linguistic Society of the Philippines, 139–182.
McFarland, Curtis D. 1977. Northern Philippine linguistic geography (Study of Languages and Cultures of Asia & Africa Monograph Series 9). Institute for the Study of Languages and Cultures of Asia and Africa.
McFarland, Curtis D. 1981. A Linguistic Atlas of the Philippines. Manila: Linguistic Society of the Philippines.
McFarland, Curtis D. 2001. “Filipino enclitics.” *Philippine Journal of Linguistics* 32(1). 1–16.
McKaughan, Howard P. 1958. *The inflection and syntax of Maranao verbs*. Manila: Institute of National Language.

Mercado, Raphael. 2002. Verb Raising and Restructuring in Tagalog. M.A. forum paper, University of Toronto.
Nagaya, Naonori. 2007. 'Information Structure and Constituent Order in Tagalog'. *Language and Linguistics* 8(1): 343–372.
Naylor, Paz Buenaventura. 1975. 'Topic, focus and emphasis in the Tagalog verbal clause', *Oceanic Linguistics* 14: 12–79.
Naylor, Paz Buenaventura. 1986. 'On the pragmatics of focus', in Paul Geraghty, Lois Carrington, and S. A. Wurm (eds.) *FOCAL 1: Papers from the Fourth International Conference on Austronesian Linguistics*, 43–57.
de Nebrija, Antonio. 1492. *Grammar of the Castillan Language*. Salamanca: Juan de Porras.
Newell, Leonard E., and Francis Bon'og Poligon (comps.). 1993. *Batad Ifugao Dictionary with Ethnographic Notes*. Linguistic Society of the Philippines Special Monograph Issue, Number 33. Manila: Linguistic Society of the Philippines.
Nolasco, Ricardo Ma. 2003. 'Ang pagkaergatibo at pagkatransitibo ng mga wikang Pilipino: Isang Pagsusuri sa Sistemang Bose', PhD thesis, University of the Philippines, Diliman.
Olson, Kenneth S., Jeff Mielke, Josephine Sanicas-Daguman, and Hugh J. Patterson III. 2010. 'The Phonetic Status of the (Inter)dental Approximant', *Journal of the International Phonetic Association* 40(2):199–215.
Or, Elsie Marie T. 2020. Focus and Aspect in Iraya Mangyan Verbs. *Archive: A Journal Dedicated to the Study of Philippine Languages and Dialects* 1: 35-74.
Orpilla, Melchor and Daniel Kaufman. 2024. Say at So: ang Dalawang Mukha ng Simuno sa Pangasinan. Presented at Talaytayan, July 20, 2024.https://bahasawan.com/wp-content/uploads/2026/04/Orpilla-and-Kaufman-2024-dalawang-mukha.pdf
Otanes, Fe T. 1966. 'A contrastive analysis of English and Tagalog verb complementation', PhD thesis, UCLA.
Pallesen, Kemp A. 1985. *Culture contact and linguistic convergence*. Manila: Linguistic Society of the Philippines.
Pangilinan, Michael R. M. 2012. *An Introduction to Kulitan, the Indigenous Kapampangan Script*. Angeles City: Center for Kapampangan Studies.
Pascasio, Emy M., ed. 1977. *The Filipino bilingual: Studies on Philippine bilingualism and bilingual education*. Quezon City: Ateneo de Manila University Press.
Pascasio, Emy M. 1984. Philippine bilingualism and code switching. *Panagani: Essays in honor of Bonifacio P. Sibayan*, ed. by Andrew Gonzalez, FSC, 122–133. Manila: Linguistic Society of the Philippines.
Payne, Thomas. 1982. Role and reference related subject properties and ergativity in Yupik Eskimo and Tagalog. *Studies in Language* (6)1:75–106.
Paz, Consuelo J. 1981. A reconstruction of Proto-Philippine phonemes and morphemes. Diliman, Quezon City: Cecilio Lopez Archives of the Philippine Languages and the Philippine Linguistics Circle.

Pebley, Carol J. and Thomas E. Payne. 2024. *A Grammar of Kagayanen*. Berlin: Language Science Press.
Peng, Adam and Loren Billings. 2008. Binukid pronominal clisis. *Studies in Philippine Languages and Cultures* 17:179-212.
Pennoyer, F. D. 1986–1987. Inati: The hidden negrito language of Panay, Philippines. Philippine Journal of Linguistics. 17(2)–18(1):1–36.
Pérez, Marilola. 2015. Cavite Chabacano Philippine Creole Spanish: Description and Typology. University of California at Berkeley dissertation.
Pittman, Richard. 1966. 'Tagalog -um- and mag-, an interim report'. Canberra: Linguistic Circle of Canberra Publications, ANU.
Porter, Doris. 1977. *A Tboli grammar*. Manila: Linguistic Society of the Philippines.
Post, Ursula, and Mary Jane Gardner. 1992. *Binukid dictionary*. Manila: Linguistic Society of the Philippines and Summer Institute of Linguistics.
Postma, Antoon. 1972.Treasure of a Minority: The Ambahan, a Poetic Expression of the Mangyans of Southern Mindoro, Philippines. Manila: Arnoldus Press.
Postma, Antoon. 1992. 'The Laguna Copper-Plate Inscription: Text and Commentary'. *Philippine Studies* 40(2):182–203
Potsdam, Eric. 2009. Austronesian verb-initial languages and wh-question strategies. Natural Language and Linguistic Theory, 27. 737-771.
Quakenbush, J. Stephen, Gail R. Hendrickson, and Josenita L. Edep (2010). *A brief overview of Agutaynen grammar*. Manila: SIL Publishing. Available at:
Rackowski, Andrea. 2002. The Structure of Tagalog: Specificity, Voice and the Distribution of Arguments. PhD thesis, MIT, Cambridge, MA.
Ramos, Teresita V. 1973. The case system of Tagalog verbs. Ph.D. diss. University of Hawai'I at Mānoa.
Reid, Lawrence A. 1973. 'Diachronic Typology of Philippine Vowel Systems', in Thomas A. Sebeok (ed.), *Current Trends in Linguistics 11: Diachronic, Areal, and Typological Linguistics*. The Hague and Paris: Mouton and Co., 485–506.
Reid, Lawrence A. 1974. 'The Central Cordilleran Subgroup of Philippine Languages', *Oceanic Linguistics* 13:511-560.
Reid, Lawrence A. 1987. The early switch hypothesis: linguistic evidence for contact between Negritos and Austronesians. *Man and Culture in Oceania* 3:41-59.
Reid, Lawrence A. 1989. 'Arta, Another Philippine Negrito Language', *Oceanic Linguistics* 28(1):47-74.
Reid, Lawrence A. 1992. 'On the Development of the Aspect System in Some Philippine Languages', *Oceanic Linguistics* 31(1):65-91.
Reid, Lawrence A. 1994a. 'Unravelling the Linguistic Histories of Philippine Negritos', in T.E. Dutton and D.T. Tryon (eds.), *Language Contact and Change in the Austronesian World*. Berlin: Mouton de Gruyter, 443-475.

Reid, Lawrence A. 1994b. ‘Possible non-Austronesian lexical elements in Philippine Negrito languages’. *Oceanic Linguistics* 33(1):37-72.
Reid, Lawrence A. 2002. ‘Determiners, Nouns or What? Problems in the Analysis of Some Commonly Occurring Forms in Philippine Languages’, *Oceanic Linguistics* 41(2):295- 309.
Reid, Lawrence A. 2005. ‘A cross-generational view of contact-related phenomena in a Philippine language: Phonology’. In *Sociolinguistics and language education in the Philippines and beyond: Festschrift in honor of Ma. Lourdes S. Bautista*, ed. by J. Stephen Quakenbush and Danilo T. Dayag. Manila: Linguistic Society of the Philippines and the Summer Institute of Linguistics.
Reid, Lawrence A. 2006. ‘On Reconstructing the Morphosyntax of Proto-Northern Luzon’, *Philippine Journal of Linguistics* 37(2):1–65.
Reid, Lawrence A. 2007. ‘Another look at the marking of plural personal noun constructions in Austronesian languages’, *Oceanic Linguistics* 46(1): 232-252.
Reid, Lawrence A. 2009. ‘The Reconstruction of a Dual Pronoun to Proto Malayo- Polynesian’, in Bethwyn Evans (ed.), *Discovering History through Language: Papers in Honour of Malcolm Ross*. Pacific Linguistics 605. Canberra: Pacific Linguistics, 461- 477.
Reid, Lawrence A. 2013. ‘Who are the Philippine Negritos? Evidence from language’. *Human Biology* 85 (1-3): 329-358.
Reid, Lawrence A. 2017. ‘Re-evaluating the position of Iraya among Philippine languages’, in Hsiu-chuan Liao (ed.), *Issues in Austronesian Historical Linguistics*. JSEALS Special Publication No. 1, 23–47.
Reid, Lawrence A. 2020. Calquing structure: How to say ‘the’ in Isinay, Philippines. *Language Documentation and Description* 17: 1-24.
Reid, Lawrence A. 2020. Response to Blust “The Resurrection of Proto-Philippines.” Oceanic Linguistics, 59(1/2): 374–393.
Reid, Lawrence A. and Hsiu-chuan Liao. 2004. ‘A brief syntactic typology of Philippine languages’, *Language and Linguistics* 5(2): 433-490.
Revel-Macdonald, Nicole. 1979. *Le Palawan (Philippines)*. Langues et Civilisations de l'Asie du Sud-Est et du Monde Insulindien Monograph #4. Paris: Centre National de la Recherche Scientifique.
Richards, Norvin. 2000. Another look at Tagalog subjects. In Paul, I., Phillips, V., and Travis, L., editors, Formal Issues in Austronesian Linguistics, pp. 105–116. Dordrecht: Kluwer.
Richards, Norvin. 2009a. Nouns, verbs, and hidden structure in Tagalog. *Theoretical Linguistics* 35:139-152.
Richards, Norvin. 2009b. The Tagalog copula. In *Proceedings of the Sixteenth Meeting of the Austronesian Formal Linguistics Association*, Sandra Chung, Daniel Finer, Ileana Paul, and Eric Potsdam (eds.), pp.181-195. London, Canada: AFLA Proceedings Project, University of Western Ontario.
Ridruejo, Emilio. 2011. The first grammar of a Philippine language? The anonymous Arte de la Lengua Sambala y Española [Grammar of the Sambal and Spanish languages] (1601). In

Lawrence Reid and Emilio Ridruejo and Thomas Stolz (eds.), Philippine and Chamorro Linguistics before the Advent of Structuralism, 11-32. Berlin: Akademie Verlag.
Riesberg, Sonja. 2014. *Symmetrical voice and Linking in Western Austronesian Languages*. Boston and Berlin: De Gruyter Mouton.
Robinson, Laura C. 2011. *Dupaningan Agta: Grammar, Vocabulary, and Text*. Canberra: Pacific Linguistics.
Robinson, Laura C., and Jason William Lobel. 2013. 'The Northeastern Luzon Subgroup of Philippine Languages', *Oceanic Linguistics* 52(1):125-168.
Ross, Malcolm. 1995. 'Reconstructing Proto Austronesian verbal morphology: evidence from Taiwan' in Paul Jen-kuei Li, Dah-an Ho, Ying-kuei Huang, Cheng-hwa Tsang, and Chiu-yu Tseng (eds.), *Austronesian studies relating to Taiwan*. Taipei: Institute of History and Philology, Academia Sinica, 727–791.
Ross, Malcolm. 2002. 'The history and transitivity of western Austronesian voice and voice marking', in Fay Wouk and Malcolm Ross (eds.), *The History and Typology of Western Austronesian Voice Systems*. Canberra: Pacific Linguistics, 17-62.
Ross, Malcolm. 2006. 'Reconstructing the case-marking and personal pronoun systems of Proto Austronesian', in Henry Yungli Chang, Lillian M. Huang, and Dah-an Ho (eds.), *Streams converging into an ocean: Festschrift in honor of Professor Paul Jen-kuei Li on his 70th birthday*. Taipei: Institute of Linguistics, Academia Sinica, 521–563.
Ross, Malcolm. 2009. Proto-Austronesian verbal morphology: A reappraisal. In Adelaar, A. and Pawley, A., editors, *Austronesian Historical Linguistics and Culture History: A festschrift for Robert Blust*, pp. 295–326. Pacific Linguistics, Canberra.
Ross, Malcolm. 2020. Comment on Blust "The Resurrection of Proto-Philippines." *Oceanic Linguistics* 59(1/2): 366–373.
Rubino, Carl R. Galvez. 2000. *Ilocano Dictionary and Grammar*. Honolulu: University of Hawai'i Press.
Rubino, Carl R. Galvez. 1997. A Reference Grammar of Ilocano. PhD thesis, University of California, Santa Barbara.
Ruffolo, Roberta. 2004. Topics in the Morpho‑syntax of Ibaloy, Northern Philippines. Ph.D. dissertation, ANU.
Sabbagh, Joseph. 2005. Non-verbal argument structure: evidence from Tagalog. PhD thesis, MIT, Cambridge.
Sabbagh, Joseph. 2011. Adjectival passives and the structure of VP in Tagalog. *Lingua* (121):1424–1452.
Salazar, Mariles S. 2012. *Perspectives on Philippine Languages: Five Centuries of European Scholarship*. Quezon City: Ateneo de Manila University Press.
Santiago, Paul Julian. 2015. Information Structure and Voice Alternation in Kalanguya. *Journal of the Southeast Asian Linguistics Society* 8:1-23.
Santos, Lope K. 1940. *Balarilà ng Wikang Pambansá*. Manila: Bureau of Printing.

Savage, Dale T. 1986. 'A Reconstruction of Proto-Southern Mindanaon', *Studies in Philippine Linguistics* 6: 181-223.
Scebold, Robert A. 2003. *Central Tagbanwa: a Philippine language on the brink of extinction: sociolinguistics, grammar, and lexicon*. Manila: Linguistic Society of the Philippines.
Schachter, Paul. 1973. 'Constraints on Clitic Order in Tagalog', in Andrew B. Gonzalez, F.S.C. (ed.), *Parangay kay Cecilio Lopez: Essays in Honor of Cecilio Lopez on His Seventy-fifth Birthday*. Quezon City: Linguistic Society of the Philippines, 214-231.
Schachter, Paul. 1976. 'The subject in Philippine languages: actor, topic, actor-topic, or none of the above. In Charles Li (ed.), *Subject and Topic*. New York: Academic Press, 491–518.
Schachter, Paul. 1996. The subject in Tagalog: still none of the above. UCLA occasional papers in linguistics, 15.
Schachter, Paul, and Fe T. Otanes. 1972. *Tagalog Reference Grammar*. Berkeley, Los Angeles, and London: University of California Press.
Smith, Alexander D. 2017. The Western Malayo-Polynesian problem. *Oceanic Linguistics* 56(2): 435–489.
Starosta, Stanley, Andrew K. Pawley, and Lawrence A. Reid. 1982. 'The evolution of focus in Austronesian' in Amran Halim, Lois Carrington, and S. A. Wurm (eds.), *Papers from the Third International Conference on Austronesian Linguistics, vol. 2, Tracking the travellers*. Canberra: Pacific Linguistics, 145-170.
Steinkrüger, Patrick O. 2008. The puzzling case of Chabacano: creolization, substrate, mixing and secondary contact. *Studies in Philippine Languages and Cultures* 19. 142-157.
Stevens, Alan M. 1969. 'Case grammar in Philippine languages', Paper read at the Fifth Meeting of the Linguistic Society of America, San Francisco.
Tanangkingsing, Michael. 2009. A functional reference grammar of Cebuano. Ph.D. dissertation, National Taiwan University.
Tharp, James A. 1974. The Northern Cordilleran Subgroup of Philippine Languages. *Working Papers in Linguistics*, University of Hawaii 6(6):53-114.
Thiessen, Henry Arnold 1981 Phonological reconstruction of Proto-Palawan. Anthropological Papers, no. 10. Manila: National Museum of the Philippines.
Totanes, Sebastian de. 1865 [1745]. *Arte de la Lengua Tagala y Manual Tagalog*. Binondo: Miguel Sanchez.
Wang, Peter, Robert Hunt, Jeff McGriff, and Richard E. Elkins. 2006. *The Grammar of Matigsalug Manobo*. Manila: SIL Philippines.
Werlen, Iwar. 2011.Wilhelm von Humboldt and the role of the verb in Tagalog. In Lawrence Reid and Emilio Ridruejo and Thomas Stolz (eds.), Philippine and Chamorro Linguistics before the Advent of Structuralism, 117-140. Berlin: Akademie Verlag.
Wolfenden, Elmer. 1961. *A Re-statement of Tagalog Grammar*. Manila: Summer Institute of Linguistics and the Institute of National Language.
Wolff, John U. 1966. *Beginning Cebuano I*. New Haven: Yale University Press.

Wolff, John U. 1967. History of the dialect of the Camote Islands, Philippines, and the spread of Cebuano Bisayan. *Oceanic Linguistics* 6: 63-79.
Wolff, John U. 1972. *A Dictionary of Cebuano Visayan*. Ithaca: Cornell University, Southeast Asia Program and Linguistic Society of the Philippines.
Wolff, John U. 1973. 'Verbal inflection in Proto-Austronesian', in Andrew B. Gonzalez (ed.), *Parangal Kay Cecilio Lopez*. Manila: Linguistic Society of the Philippines, 71-91
Wolff, John U. 1968. The historical development of the Samar-Leyte Bisayan vowel system. *Samar-Leyte Studies* 2:33-39.
Wolff, John U. 2005. The Philippine Languages and the Determination of PAN Syllable Structure. In Liao, H.-C. and Rubino, C. R. G., (eds.), *Current issues in Philippine linguistics and anthropology: Parangal kay Lawrence A. Reid*, pp. 69–80. Linguistic Society of the Philippines and SIL, Manila, Philippines.
Wu, Jiang. 2017. 'Plural Words in Austronesian Languages: Typology and History', MA thesis, Leiden University.
Yamada, Yukihiro. 2002. *Itbayat-English Dictionary.* Endangered Languages of the Pacific Rim (ELPR) Publications Series A3-006. Kyoto: Nakanishi Printing Co., Ltd.
Zamar, Sheila Maria. 2022. Apprehending Philippine Negrito Languages, 1890-1990: An Inquiry into Linguistic Ideology. Ph.D. dissertation. University of Wisconsin-Madison.
Zobel, Eric. 2005. 'Buol', in Alexander Adelaar and Nikolaus Himmelmann (eds.), *The Austronesian Languages of Asia and Madagascar*. London and New York: Routledge, 625-648.
Zorc, R. David. 1974. 'Internal and External Relationships of the Mangyan Languages', *Oceanic Linguistics* 13: 561–600.
Zorc, R. David. 1977. *The Bisayan Dialects of the Philippines: Subgrouping and Reconstruction*. Pacific Linguistics C-44. Canberra: Department of Linguistics, Research School of Pacific Studies, Australian National University.
Zorc, R. David. 1986. 'The Genetic Relationships of Philippine Languages', in Paul Geraghty, Lois Carrington, and S. A. Wurm (eds.), *FOCAL II: Papers from the Fourth International Conference on Austronesian Linguistics*. Pacific Linguistics C-94. Canberra: Department of Linguistics, Research School of Pacific Studies, Australian National University, 147-173.
Zorc, R. David. 1993. Overview of Austronesian and Philippine accent patterns, in Jerold A Edmondson and Kenneth J Gregerson (eds.), *Tonality in Austronesian Languages*. Honolulu: University of Hawai'i Press, 17-24.
Zorc, R. David. 1982. Where, O where have the laryngeals gone? Austronesian laryngeals re-examined. Papers from the Third International Conference on Austronesian Linguistics, Vol. 2: Tracking the travelers, ed. by Amran Halim, Lois Carrington, and S. A. Wurm, 111–144. Pacific Linguistics C–75. Canberra: Department of Linguistics, Research School of Pacific Studies, The Australian National University.

Zorc, R. David. 1996. The reconstruction and status of Austronesian glottal stop — chimera or chameleon. In Bernd Nothofer (ed.), *Reconstruction, classification, description: Festschrift in honor of Isidore Dyen*, 41–72. Abera Verlag.
Zorc, R. David. 2019. 'Klata/Giangan: A new Philippine subgroup', *The Archive: Special Publication* 16: 33-51.
Zorc, R. David. 2020. Reactions to Blust's "The Resurrection of Proto-Philippines." *Oceanic Linguistics* 59(1/2): 394–425.
Zorc, R. David, Jason Lobel, and William Hall. 2024. 'The historical linguistics of the Philippines.' In *The Oxford Guide to the Malayo-Polynesian Languages of Island Southeast Asia*, Alexander Adelaar and Antoinette Schapper (eds.), pp.79-94. Oxford: OUP.
Zuraw, Kie. 2000. 'Patterned Exceptions in Phonology', PhD thesis, UCLA.
Zwartjes, Otto. 2011. Oyanguren de Santa Inés's grammar of Tagalog (Tagalysmo Elucidado 1742): towards a reconstruction of 18th century reflections on comparative typology. In Reid, Lawrence A. and Ridruejo, Emilio and Stolz, Thomas (eds.), *Philippine and Chamorro Linguistics Before the Advent of Structuralism*, 63-85. Berlin: Akademie Verlag.

APPENDIX 1. ABBREVIATIONS

ABS – absolutive
ADJ – adjective
ASP – aspect marker
AV – actor voice
BEG – begun aspect
BV – benefactive voice
CAUS – causative
CO – "co-"/sharing relation
CV – conveyance voice
DEF – definite
DEP – dependent mood
DET – determiner
DIST – distributive
DUR – durative
ERG – ergative
EXT – existential
FUT – future
GEN – genitive case
GER – gerund
HYP – hypothetical
IMMD – immediate future
INCH – inchoative
INDEF – indefinite
INTNS – intensive

IPFV – imperfective
ITER – iterative
LNK – linker
LOC – locative
LV – locative voice
MODER – moderate degree
NEG – negative
NM – noun marker
NMLZ – nominalizer
NOM – nominative case
OBL – oblique case
PL – plural marker
POT – potentive
PRETEND – pretendative
PFV – perfective
PROG – progressive
PV – patient voice
RCT – recent perfective
REAS – reason voice
RECP – reciprocal
RL – realis
SBJV - subjunctive
SOC – sociative
SPEC – specific
STA – stative
TOP – topic marker
TR – transitivity related
UV – undergoer voice
VRB – verbalizer