DRAFT (comments welcome)

Typological overview of the languages of central and southern Philippines

Daniel Kaufman

1.0 Introduction

This chapter provides a typological overview of the languages of the Central and Southern Philippines (henceforth, CSP languages). Despite not forming a discrete phylogenetic group, the CSP languages share certain morphosyntactic retentions from Malayo-Polynesian (henceforth MP) which make them a useful unit for typological generalizations. Like other Philippine languages, almost all the CSP languages maintain the full MP voice system. However, unlike many languages of the northern Philippines, many CSP languages still maintain a distinction between the independent and dependent paradigms of the voice system (Wolff 1973). They may also reflect the Proto-Malayo-Polynesian (PMP) voice system more directly in disallowing innovative combinations of voice morphology commonly found in the north (e.g. the combination of actor voice and conveyance voice morphology). In one case, at least, there is a syntactic retention from Proto-Austronesian (PAn) that is only found in tiny handful of MP languages, discussed in §4.5.4. On the other hand, the voice system mutates in interesting ways in the southernmost CSP languages, specifically, in the Bilic and Sama languages.

The larger question of a Philippine subgroup has been recently addressed by Blust (2019), who argues for its existence on the basis of a large set of proposed lexical innovations. Others have argued that the Philippine languages are best understood as a collection of primary subgroups of MP, rather than the daughters of a single MP proto-language. Phonologically, the only potential innovation that would characterize Proto-Philippines is the merger of PMP *z and *d, which is so common outside Philippine languages that it is of little probative value. Similarly, there are no identifiable morphosyntactic innovations that characterize the putative Proto-Philippines. At this point in our understanding, a reconstruction of Proto-Philippine morphosyntax would not appear any different from a reconstruction of PMP morphosyntax. The question of whether there existed a Proto-Philippine language distinct from PMP will likely remain under debate for some time to come.

1.1 Overview of languages

The languages examined in this chapter are given below, following the family tree argued for by Blust (1991), who posits a large Greater Central Philippine subgroup. These languages comprise a subset of the Philippine language group argued for by Blust (2019) with the addition of the Sama-Bajaw languages.

Greater Central Philippines
- Umiray Dumaget Agta (1)
- South Mangyan (3)
- Central Philippine subgroup (40)

Palawanic subgroup (3)
Danaw subgroup (3)
Manobo subgroup (15)
Subanen subgroup (6)
[Gorontalic (9)]

Kalamian

Agutaynen
Calamian Tagbanwa

Bilic

Tboli
Blaan
Tiruray
Giangan / Bagobo

Sama-Bajaw

Central Sama
Southern Sama
Balangigi
Pangutaran Sama
Yakan
Mapun
Abaknon

The Sama-Bajaw languages (henceforth Sama), while located in part within the Philippines, are now understood to be more recent arrivals originating from Borneo and belonging to the Barito subgroup (Blust 2007). The morphosyntax of these languages differs from other CSP languages in interesting ways, although a good deal of typological convergence has taken place (Pallesen 1985), and so it is not out of place to discuss Sama languages together with Philippine languages in this overview.

It should also be noted that members of the putative Philippine group are located outside of the Philippines. Namely, the Sangiric and Minahasan languages, located in North Sulawesi, are argued by Blust (2019) to belong to the Philippine subgroup and the Gorontalic languages are argued by Blust (1991) to belong to the Greater Central Philippines subgroup of Philippine languages. We exclude here Minahasan, Sangiric and Gorontalic languages on purely geographical grounds, although it should be noted that these languages have been influenced by distinct contact scenarios (most notably by local Malay varieties) over the last several centuries, which have made them diverge morphosyntactically from their more northern relatives.

1.2 A note on sources and glossing

As might be expected, only a small share of the above languages have been fully described, although at least some information is available for almost all of them. Among the Central Philippine languages, I rely most heavily upon Tagalog, the basis for the national language of the Philippines, which has a rich history of description and whose grammar continues to provide a lively arena for theoretical linguists. David Zorc's (1977) seminal dissertation on the structure and subgrouping of the Bisayan languages of the Central Philippine group has also provided many of the examples cited herein. I also draw heavily upon

McKaughan's (1957) description of Maranao morphosyntax, Daguman's (2004) description of Northern Subanen, and Wang, Hunt, McGriff and Elkins' (2006) grammar of Matigsalug Manobo, to represent the Danaw, Subanen and Manobo subgroups, respectively. Outside the Greater Central Philippine group, I draw mostly from Forsberg's (2002) description of Tboli. The almost total lack of grammatical description for other Bilic languages is especially unfortunate considering how divergent this subgroup is from all other Philippine languages. For the Sama languages, I rely mostly on Brainard & Behrens (2002) description of Yakan, Walton's (1986) description of Pangutaran Sama, and Akamine's (2005) description of Southern Sama. For Kalamianic, a distinct subgroup found in the northern Palawan province (Himes 2006), I rely on Quakenbush et al (2010) for Agutaynen and Scebold's (2003) description of Central Tagbanwa. Previous typological overviews of Philippine languages include Reid and Liao (2004) and Himmelmann (2005), both considered key works in Philippine linguistics.

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. Rather, I have glossed functional/historical cognates uniformly only to facilitate cross-linguistic comparison. 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].

2.0 Phonology

2.1 Segment inventories

Vowel inventories in the CSP zone are relatively simple. On the simplest side, 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, as well. In rare cases, the inherited vowel inventory has been expanded in complex ways. In Tboli, for instance, the four vowel system has been expanded into a seven vowel system (Porter 1977, Forsberg 1992).

Several languages have developed an allophonic relationship between the high vowels and their mid counterparts. In Tagalog, a generally word-final process of vowel lowering turns *i* and *u* into *e* and *o*, respectively. Kapampangan of the Central Luzon group (outside the purview of this chapter) has innovated a new set of mid vowels not from lowering of high vowels but rather through monophthongization of **aj* > *e* and **aw* > *o*, but this is vanishingly rare in the CSP zone. A large monophthongization zone begins just southeast of the CSP languages in Sulawesi and includes the Sangiric languages.

Table 1. Typical Central Philippine vowel inventory

i	ɨ	u
(e)		(o)
	a	

Consonant inventories are also relatively simple and do not vary much across the area surveyed here. A typical inventory of phonemic consonants for the Central Philippine group is shown in Table 2. The tap ɾ can have several historical sources. Most typically, it is an intervocalic allophone of /d/.

We also find palatals at various stages of phonemicization, typically resulting from the combination of alveolars preceding /ij/, e.g. Tagalog 3s.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).

Table 2. Typical Central Philippine consonant inventory

	labial	**alveolar**	**palatal**	**velar**	**laryngeal**
voiceless stop	p	t		k	ʔ
voiced stop	b	d		g	
nasal	m	n		ŋ	
fricative		s			h
lateral		l			
tap/trill		ɾ			
glide	w		j		

We can compare this picture with Blust's reconstruction of the Proto-Philippine inventory shown below. Proto-Philippine *q (inherited from PAn *q) is reflected as ʔ in the vast majority of Philippine languages but as *k* in Bilic and Kalamianic. The palatal consonants, *z, *ñ and *j, are generally merged with other consonants in the Philippine languages (although Kapampangan appears to exceptionally preserve a distinct reflex of *ñ).[1]

[1] Proto-Philippine *r is a relatively marginal phoneme and argued by Wolff (1974) to be an artifact of loan vocabulary rather than a distinct reconstructable phoneme at PMP or PAn.

Table 3. Blust's (1991:88) Proto-Philippine inventory

	labial	**alveolar**	**palatal**	**palatalized velar**	**velar**	**post-velar**
voiceless stop	p	t			k	q
voiced stop	b	d	z [dʒ]	j [kʲ]	g	
nasal	m	n	ñ [ɲ]		ŋ	
fricative		s				h
lateral		l				
tap/trill		r			R [ɣ]	
glide	w		y			

Unusual segments in CSP languages include the fortis/heavy stops of Maranao, described by Lobel and Riwarung (2009), and the aspirated stops of Subanen, described by Lobel and Hall (2010). As Lobel & Hall (2010:336-337) note, these form part of a larger set of unusual reflexes of consonant clusters in the languages of Mindanao and northern Borneo, a fact that they tentatively attribute to language contact. There is a velar fricative or approximant that has been described for Aklanon and which also exists in dialects of Bikol, although these developed independently.

2.2 Phonotactics

The canonical lexical root in Philippine languages is a disyllable with the following template: CV(C).CV(C). On one 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.[2] Relatedly, there is a general lack of vowel hiatus in most CSP languages, as vowel hiatus relies on the possibility of onsetless syllables.[3] Root initial glottal stops, whether they are underlying or epenthetic, surface predictably with prefixation, as

[2] I have found only one language in the CSP zone that is described as contrasting vowel initial syllables with glottal initial syllables. Scebold (2003:30) offers the following contrast, in which the glottal stop of *ʔinɨsɨl* 'repent' disappears with prefixation but that of *ʔɨm* does not.

root-initial /ʔ/: [mag]+[ʔɨm] → [magʔɨm] 'to soak'
root-initial V: [mag]+[ʔinɨsɨl] → [maginɨsɨl] 'to repent'

He takes this to imply that the first root is underlying glottal initial, i.e. /ʔɨm/, while the second root obtains a glottal via epenthesis when it is word-initial but is underlyingly vowel initial, i.e. /inɨsɨl/. Note, though, that even in Tagbanwa the putative contrast is neutralized in isolation, where both roots would be pronounced with a glottal stop.

[3] Words that are written with two vowels orthographically in languages like Tagalog, e.g. *bait* 'goodness', are pronounced with an intervening glottal stop, e.g. [baʔit]. Zorc (1977:54) mentions Cuyunon and certain dialects of Tausug as exceptional in allowing vowel hiatus.

in /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.

For the vast majority of the CSP languages, lexical roots are disyllabic and occasionally trisyllabic whereas function words are either monosyllabic or disyllabic but very rarely trisyllabic. Monosyllabic lexical roots are a relatively recent innovation in Central Philippine languages, having entered through loans and various processes of reduction, but such roots are rare even now. In several languages of the Sulu archipelago, the deletion of intervocalic /l/ has created monosyllables with long vowels, e.g. Basilan Yakan *ulu* > *u:* 'handle of a knife' (Brainard and Behrens 2002:6). Tboli shows another pattern of historically truncated monosyllabic roots, e.g. PMP **epat* > *fat* 'four', PAN **kaen* > *ken* 'eat'.

Affixes do not have the same constraints as lexical roots; they are typically monosyllabic and need not contain onsets. Onsetless affixes are typically provided with an onset either through epenthesis or infixation. Infixation only applies at the left edge of the base; there are no attested cases of right-aligned infixation in Austronesian languages. 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 Tagkaolo 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 Tagkaolo.[4]

(1)	a.	bagu-hin	b.	bagu-wun	
		new-PV		new-PV	(Burton 2018)

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

(2)	a.	s<um>agot	b.	t<um>ubag	
		<AV>answer		<AV>answer	(Burton 2018)

Gemination is relatively rare in the CSP zone but is attested in Bagobo and the Sama languages.[5] In Central Sama **(**Reid 1971, Blust & Trussel ongoing), consonant gemination arises historically from a preceding schwa, e.g. PMP *qahelu > *hallu* 'pestle', PMP *qateluR > *ɨntɨllo* 'egg', PWMP *qalesem > *lessom* 'sour'. Gemination seems persistent in those subgroups where it has developed. Abaknon, an endangered Sama language spoken in the Bisayan region, maintains geminates from Proto-Sama-Bajaw, e.g. PMP *qahelu > Abaknon *allo* 'rice pestle', PMP *beli > Abaknon *balli* 'to buy', even while having reduced the Proto-Sama-Bajaw seven vowel system to a three vowel system under Bisayan influence.

Glottal stop arises from the historical change PMP *q > ʔ which took place widely throughout the Philippines but the synchronic distribution of the glottal stop varies by language and region.[6] For instance, PMP *baqeRu 'new', reduced historically to a disyllable, yields (Naga) Bikol *baʔgo*, Cebuano *bagʔo* (with metathesis), and Tagalog *ba:go* (with deletion). These

[4] In many languages of the northern Philippines, where root-final glottal stops are lost historically, the glottal stop serves as the epenthetic consonant providing an onset to onsetless suffixes.

[5] Among non-Sama languages, Blust (2013:229) also includes Kagayanen, Mansakan languages and Rinconada Bikol as allowing at least some geminates.

[6] See Zorc (1996) for the reconstruction of a glottal stop unrelated to *q. The existence of an independent glottal stop phoneme in PMP is not widely accepted.

changes follow a general pattern as Cebuano does not allow ʔC clusters and Standard Tagalog does not allow either Cʔ or ʔC clusters. 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/.

There appears to be a gradated loss of root final glottal stop from south to north. In southern CSP languages, glottal stop is highly salient phonetically and never undergoes deletion. In many languages of the northern Philippines, root final glottal stop has been lost completely. In Tagalog, which lies on the border, word final glottal stop is less phonetically salient than in the south and it is often lost in phrase medial position, occasionally with compensatory lengthening, e.g. /walaʔ=na/ NEG.EXT=already [walaːna]. But even within a single subgroup, 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 Iranon and Magindanao have both eliminated it in this position.

To summarize the status of the glottal stop in CSP languages: (i) there is only one language that possibly shows a contrast between V and ʔV at the beginning of roots (Central Tagbanwa); (ii) root internally, some language allow ʔC, others Cʔ, while others allow neither; (iii) most but not all CSP languages contrast root-final ʔ with root-final V.

In most CSP languages, glides pattern like any other consonant in the native vocabulary, but in some languages, glides can be the internal segment in a tautosyllabic cluster.[7] For instance, in the Jolo dialect of Tausug, we find monosyllables such as *awn* EXIST and *lawŋ* 'inside'. In onset position, we find languages such as Maranao and Tagkaolo where the historical perfective infix *<in> has been reduced to a single glide <y>. In these languages, onset clusters with *y* as a second member are commonly derived through infixation. Tagalog shows a historical pattern of intervocalic *l* deletion which occasionally gives rise to similar clusters, e.g. PMP *bulan > Earlier Tag. *buwan* > Tag. *bwan*.

The Bilic languages of Southern Mindanao, Tboli and Blaan, are exceptional with regard to the typically simple syllable margins of Philippine languages. Tboli allows for a large number of typologically rare onset clusters that violate the principle of sonority sequencing with regard to manner and voicing, e.g. /btaŋ/ 'fall', /tboli/ 'Tboli'.[8]

Complex tautosyllabic clusters have also entered CSP 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.

As noted by Blust (2013:62), Philippine languages overall show far more heterorganic clusters across non-reduplicated syllables than found outside the Philippines, where these are rare. All Philippine languages allow heterosyllabic clusters although each language exhibits its own constraints and preferences. Interestingly, such clusters may be innovative and do not generally reconstruct to PMP. The only clusters found at the PMP level as reconstructed in Blust & Trussel (ongoing) are either nasal+stop sequences, e.g. *simbuR 'to sprinkle', or the result of

[7] Alternatively, diphthongs can be seen to occupy the nucleus of the syllable.

[8] 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.

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. For these cases, Blust reconstructs their Proto-Philippine forms with an intermediate schwa in parentheses, e.g. Proto-Philippine *lib(e)tuŋ ‘deep place in river’. However, there is no direct evidence for these intermediate vowels which serve only to maintain PMP phonotactics for what would otherwise be a phonologically conservative proto-language.[9] It should be noted that gradient phonotactic patterns have not been examined systematically for languages of the CSP area and this could be a rich area for further study.[10]

2.3 Phonological and morphophonological processes

The phonology of most 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. We exemplify these processes with Tagalog and a handful of other languages below.

2.3.1 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, although other languages, show tapping in these contexts, too, as shown in (Q) for Matigsalug Manobo.

(P) LENITION/TAPPING

a. /da:~datiŋ/ → [da:ɾatiŋ] IMPRF~arrive

b. /aku=din/ → [ako ɾin] 1s.NOM=also

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

(Q) LENITION/TAPPING (Wang et al 2006:3)

/me=datuʔ/ → [me ɾatuʔ]
PL=chief

Other types of lenition can be found in Western Bukidnon Manobo (Blust 2013:236), where it applies productively with affixation, e.g. *balәy* ‘house’, *bә-valәy* ‘build a house’, *guraŋ* ‘old’, *mә-ɣuraŋ* ‘old person; old’. Deletion of intervocalic /l/ is also common across the area and was clearly a historical process in Tagalog, as well, although it was not carried out to completion.

[9] Languages such as Western Bukidnon Manobo, which often maintain historical trisyllablic stems as such (e.g. PPh *baketin ‘piglet, suckling pig’ > *beketin* versus Bikol, Hanunóo, Waray-Waray, Cebuano, Mansaka *baktin* and Maranao *baktiŋ*) still do not offer evidence for Blust’s parenthetical schwa (e.g. PPh *sab(e)láy > *sebley* ‘to hang something over something else’, *sub(e)líq > *subliʔ* ‘to repeat an action’).

[10] Zorc (1977:53) notes the existence of phonotactic constraints in heterosyllabic clusters but laments the lack of data to address its nature. For Austronesian languages outside the CSP area, see Coetzee and Pater (2006) for Muna (Southeast Sulawesi) and Benton (1971) for Pangasinan (Northern Luzon) for examples.

2.3.2 Palatalization

A palatalization processes takes place in Tagalog with the alveolar obstruents /t/, /s/ and /d/ before /j/, as shown in (S). A phonetically less natural palatalization process also takes place with the sequence /ts/, transforming it to [tʃ].

(S) PALATALIZATION

a.	/sija/ → [ʃja] ~ [ʃa]	b.	/tijan/ → [tʃjan] ~ [tʃan]
	3s.NOM		stomach

In Central Tagbanwa, we find a similar but more circumscribed palatalization rule: /t/ → [tʃ] / __i. A similar pattern, although less advanced, is found in other Central Philippine languages like Cebuano, although the rule /ts/ → [tʃ] appears somewhat unique to Tagalog. Despite allophonic rules that create palatal or alveopalatal segments, no CSP language has fully phonemicized a palatal series of consonants.

2.3.3 Syncope and metathesis

The canonical Austronesian root is disyllabic and trisyllabic stems are often reduced to disyllables through an active rule of syncope in many CSP languages, exemplified by Agutaynen in (U). In other languages, including Tagalog, syncope is not productive but attested in stem allomorphy.

(U) SYNCOPE (Quakenbush et al 2013:41)

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

In rarer cases, syncope has been attested across clitic boundaries, as described by Lobel and Riwarung (2009, 2011) for Maranao clitics, such as *səka* 2s.NOM and *səkano* 2s.NOM, shown in (G).

(G) [dɤ.ʔɤ.mɪs.ka.no.ma.ɪ.lay]
/daʔ ami səkano ma-ilay/
NEG 1PEX.GEN 2P.NOM PV.POT-see
'We didn't see you (pl.)' (Lobel and Riwarung 2011:41)

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 (W).[11] While in other cases, Agutaynen employs metathesis, as in (T).

(W) SYNCOPE + DEBUCCALIZATION (Quakenbush et al 2002:42)
/te~teled/ → tetled → [teʔled]

[11] Quakenbush (1991) discusses a more general neutralization rule that changes the first consonant of a consonant cluster to a glottal stop. It is described as obligatory when the first consonant is voiceless but variable in other conditions.

PROG~enter

(T) SYNCOPE + METATHESIS (Quakenbush et al 2002:41)

a. /pa-belag/ → pablag → [palbag]
CAU-separate

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

Neither syncope nor metathesis are productive in Tagalog but both processes are richly attested in allomorphy, as seen in (K).

(K) METATHESIS (Bloomfield 1933:391)

a. /atip-an/ → atpan → [aptan]
roof-LV

b. /silid-an/ → sildan → [sidlan]
room-LV

Metathesis can also be triggered by non-contiguous segments. In Tagalog, the aspectual infix <*in*> undergoes metathesis to become a prefix *ni-* before /l/, /r/ and between the conveyance voice prefix ʔ*i-* and a stem beginning with a glottal consonant, as exemplified in (L). While the prefix *ʔi-* can attach to a glottal-initial stem, as in (La), and <*in*> can infix into such a stem, as in (Lb), the infix metathesizes when both elements co-occur, as seen in (Lc) and (d), possibly as a repair mechanism that avoids glottal consonants in consecutive syllables.[12] The default exponence of these co-occurring morphemes is shown in (Le).

(L) a. ʔi-ha:nap
CV-search

b. h<in>a:nap-Ø
<BEG>search-PV

c. ʔi-ni-ha:nap
CV-BEG-search

d. *ʔi-h<in>a:nap
CV-<BEG>search

e. ʔi-b<in>igay
CV-<BEG>give

Sequencing of liquids is also constrained in some languages and "repaired" by metathesis. Matthes (2014:62) notes that /r/ always precedes /l/ in the native Bikol lexicon and that metathesis of the plural infix <Vr> takes place to make derived forms also conform to this pattern, as shown in (O).

(O) /l<Vr>u:tuʔ/ → [rulu:toʔ]
<PL>cook

[12] Similar effects in a wide range of languages have been attributed to the Obligatory Contour Principle (OCP) (Leben 1973), whereby marked segments (e.g. high tone) cannot appear adjacent to each other at some underlying level of representation.

Tagalog has a process of glottal deletion with compensatory lengthening, shown in (L), but this has not been reported for many other languages of the CSP zone. This process takes place in most phrase internal contexts in connected speech and is obligatory before enclitics. This can be viewed as the first step in the loss of final glottal stops, a process which has already taken place in many languages to the north of Tagalog.

(L) COMPENSATORY LENGTHENING (Bloomfield 1933:391)
/walaʔ=na/ → [walaːna]
NEG.EXT=already

Vowel reduction is not common in Central Philippine languages but found in several languages of Mindanao and Sulu, e.g. Sindangan Subanen (Arms 1996:5), as well as Bornean languages south of the Philippines. One example from a CSP language comes from Yakan (H), although here it is not reduction to schwa but to [ɛ] in pre-stress (pre-penultimate) positions. Note that a productive rule of intervocalic /l/ deletion can also be seen here after prefixation of *pa-*.

(H) VOWEL REDUCTION (Brainard and Behrens 2002:7)
/mag-pa-ˈlaboʔ-an/ → mɛg-pɛːˈboʔan
AV-CAU-drop-LOC ‘repeatedly drop something’

Vowel harmony does not play a prominent role in Philippine languages but Lobel and Riwarung (2009, 2011) describe an intriguing case 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, ʊ]. 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. Furthermore, Lobel and Riwarung (2011:39) show that tensing spreads from left to right with /l/ and /ʔ/ being transparent but with other consonants blocking harmony. Because the heavy/light distinction on consonants plays an important role in the morphology, there are minimal pairs for every verb, as exemplified in (X). The “future” (most likely, prospective aspect) is signaled by the change of a light stem initial consonant to its heavy counterpart, and the consequent vowel harmony.

(X)a. [t̪a.ʔa.man] b. [t̪’ʊ.ʔʊ.man]
/taʔam-an/ /t’aʔam-an/
taste-LV FUT/taste-LV (Lobel and Riwarung 2011:40)

Central Tagbanwa shows a rightwards vowel harmony process with prefixes, as shown in (Y). Unlike Maranao, this process does not affect lexical stems and is restricted to the change /a/→[u] immediately following a syllable bearing /u/.

(Y)a. [pupuŋaralan] b. [pugputabas]
/pu-paŋ-aral-an/ /pug-pa-tabas/
INCMP-DIST-study-LV AV.INCMP-CAU-prune (Scebold 2003:35)

Philippine morphophonology has played a prominent role in the theoretical literature, especially within Optimality Theory and McCarthy and Prince's (1993) theory of Prosodic Morphology, in which morphology can only treat prosodic units such as the syllable and mora as anchors for concatenation and reduplication, in contrast to purely segmental structure. Theoretical investigations have looked at how the position of segmental and reduplicant infixes are determined, with some opting to analyze infixes as displaced prefixes in a framework of violable constraints (Prince and Smolensky 1993, McCarthy and Prince 1993), others following allowing for a more permissive approach to affixation constrained by diachrony (Blevins 1999, Yu 2002), and yet others taking different approaches (Ryan 2010). The variation in how Tagalog infixes are positioned in onset clusters (in loan words) has been used by Zuraw (2007) to explore potentially innate phonotactic preferences. Nasal substitution has been studied with regard to how phonologically grounded patterns interact with lexically specific phonology (Zuraw 2010). We do not review here the implications of these phenomena but only note that the patterns and variation found in CSP morphophonology have been and continue to be a hotbed of theoretical interest.

2.3.4 Infixation

Two productive infixes inherited from PAn, *<um> ACTOR VOICE and *<in> PERFECTIVE/BEGUN ASPECT, continue to play an important role in CSP languages and Philippine languages more generally (Reid 1992). They are positioned after the first consonant of the stem, as shown in (X) for Tagalog. Historically, both of these infixes could co-occur as shown in Bikolano (Y), although this is only found in a small number of living languages (see Lobel 2004 for discussion).

(X)a.	k<in>u:ha-Ø	b.	k<um>u:ha	(Y)	k<um><in>u:ha
	<in>ku:ha-PV		<um>ku:ha		<um><in>ku:ha
	<BEG>take		<AV>take		<AV><BEG>take
	'taken'		'take'		'consequently took'

Infixation has been externalized altogether in a number of CSP languages in a process which turns *<um> into *mu-* and *<in> into *ni-*, as found in modern Cebuano. In many cases, metathesis operates on infixes under particular phonological conditions, as in (L), above. Reflexes of *<in> have been reduced to a single segment in Danao languages (e.g. Maranao *t<i>abas* <PRF>cut), Tboli, Mansakan, and elsewhere in Mindanao.

Other more recently developed infixes occur, as well, in a large number of CSP languages. For instance, in Bikolano and several Bisayan languages we find a plural infix <Vr>, whose vowel harmonizes with the first vowel of the stem, shown above in (O). Another widepsread *<aŋ>* prefix marks plurality in adjectives and, in some languages, event denoting predicates.

2.3.5 Reduplication

Philippine languages tend to make heavy use of various types of reduplication for a vast number of purposes, as first noted in the English literature by Blake (1917). 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 (Xa), intensive formation, and elsewhere. CV reduplication with vowel length, shown in (Xb), is used chiefly in the aspectual paradigm for the imperfective or progressive.

(X)a.	mag-na~naːkaw	b.	mag-naː~naːkaw
	AV-AGT.NMLZ-steal		AV-IMPRF-steal
	‘thief’		‘will steal’

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

(Z)a.	ma-ganda~ganda=sila	b.	baliː~baliːtaʔ
	ADJ-MODER~beauty=3p.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 (O).

(O)a.	naka-tohod	b.	naka-tohod~naka-tohod
	LOC-forest		LOC-forest~LOC-forest
	‘in the forest’		‘deep in the forest’ (Scebold 2003:42)

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

(W)a.	mag-haː~hanap~hanap	b.	p<in>ag-saː~sa~sabi
	AV-IMPRF~ITER~search		<BEG>TR-IMPRF~INTNS~search
	‘will keep searching’		‘what is being said (intensively)’

Whereas Tagalog reduplication simply truncates a base that is more than two syllables, Cebuano and Bikol and use an entirely different strategy: reduplication with fixed segmentalism. Thus, for a trisyllabic Cebuano stem like *padala* ‘send’ we find *pulupadala*, where the first consonant of the stem has been copied and the following *ulu* is infixed, instead of **padala~padala* or **pada~padala*. In Bikolano, we find a similar strategy employing *Curu-* (where C represents the first consonant of the stem). In Bikol, this allomorph is also employed (with some exceptions) when the stem contains consonant clusters (contains tautosyllabic or heterosyllabic) and when the stem consists of two identical syllables, e.g. *rara* ‘poison’ → *ruru~rara* (Matthes 2014:76).

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

(P) t<um>umpok a t<um>umpok
<AV>pile LNK <AV>pile
'kept piling up' (Scebold 2003:57)

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

Finally, as noted by Reid (1992), the order of operations with reduplication and infixation is unexpected. Whereas there is some disagreement on whether to treat Philippine voice as derivational or inflectional, it should be clear that aspect is more on the inflectional end of the spectrum when compared to voice as the latter is more idiosyncratic and can transform the core meaning of a word. Nonetheless, aspect reduplication must take place before infixation to produce the attested forms in most Philippine languages. In (Da), we see the attested Tagalog form, where imperfective CV reduplication has applied first, yielding *gaːganda* after which the actor voice marker has infixed, yielding *gumaːganda*, and the unattested form in (Db), where aspect applies to the voice derived form as might have been expected.

(D)a. g<um>aː~ganda b. *guː~g<um>anda
<AV>IMPRF~beauty IMPRF~<AV>beauty
'becoming beautiful'

2.3.6 Nasal substitution

Languages of the CSP zone, 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 (§3.5.2). 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 (Q).[13]

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

Patterns of nasal substitution in Austronesian languages have been surveyed in detail by Blust (2004), who shows that several factors, including voicing, place and manner features, come into play in determining when deletion takes place and when the consonant remains or is preceded by an epenthetic vowel.

In comparison to Malay and other languages of Indonesia, 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 CSP languages, nasal substitution patterns are completely

[13] The nasal coda of the prefixes that trigger nasal substitution are often represented by N, a placeless nasal with special morphophonological properties.

predictable on the basis of phonology alone, typically with stem-initial voiceless segments undergoing deletion and voiced segments being maintained.

2.3.7 Morphological use of vowel length

The morphology of CSP languages commonly involve vowel length, generally referred to as "contrastive stress" or "accent" in the literature. 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-raːyuʔ/). 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.

The morphological use of vowel length and stress in the Central Philippine languages is still largely uncharted territory. Even for Tagalog, the best studied language of the CSP region, the facts remain elusive and not well understood. Little progress has been made since Zorc 1977 and some following work may have obscured these matters by ignoring the crucial role of vowel length in favor of a purely stress based analysis (French 1988, 1991).

2.4 Stress and prosody

Languages of the Philippines are unusual when compared to Austronesian languages outside the Philippines in having a phonemic stress/prominence distinction on roots. As discussed in Chapter X, the feature underlying this is best described as vowel length for at least a portion of these languages if not all of them and, in at least the native stratum, is only contrastive in the penultimate syllable. As noted in the previous subsection, vowel length plays an important morphological role in many CSP languages, as well. For instance, Tagalog /baːyad/ 'payment' contains vowel length on the penultimate syllable which is removed by a process of resultative formation yielding /bayad/ 'paid for'.[14] In isolation, this appears as a difference between penultimate and final stress. However, what appears to be word final stress in isolation can be shown to originate on the phrasal level, in contradistinction to apparent penultimate stress, which is clearly a word-based phenomenon. This becomes clear in the presence of second position clitics, as in (X), where the apparent final stress shifts to the last syllable of the prosodic phrase.

(x)a.	[da'la]		b.	[dalako'na]
	/dala/			/dala=ko=na/
	carry			carry=1s.GEN=ALRD

[14] This is rendered *bayád* in the official orthography as words without long vowels give the impression of final stress in isolation.

Central Philippine languages differ in whether closed penultimate syllables attract pitch prominence in the same way that penultimate long vowels do. 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 do attract pitch prominence on par with syllables containing a long vowel. Thus, a root like /dakdak/, in isolation, would surface as [dakˈdak] in Tagalog but [ˈdakdak] in Cebuano.

As noted by Blust (2013:251) and in Chapter X, 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. Pallesen also suggests that it is easier to lose this distinction than to recoup it through contact. Abaknon, a Sama language that has undergone heavy contact with Samareño, a Bisayan language with contrastive accent, has not (re-)developed the contrast but rather patterns like its Sama relatives in displaying predictable penultimate stress.

Other languages of the CSP zone without contrastive accent include Central Tagbanwa, which shows variable stress (Scebold 2003:27), Agutaynen, described by Quakenbush et al (2010:40) as having penultimate phrase-based stress, Matigsalug Manobo, which shows regular penultimate word based stress (Wang et al 2006:3), Maranao (Lobel and Riwarung 2011), and Tboli, which shows regular word final stress (Forsberg 1992).

3.0 Morphology

The morphology of most Philippine languages is highly complex along several dimensions: (i) a large proportion of morphs are multifunctional and take on distinct meanings in different morphological contexts; (ii) much of the morphology is portmanteau, yielding a highly "fusional" language in Sapir's classic typology; (iii) 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. In this section, we examine root classes (§3.1), aspect morphology (§3.2), voice morphology (§3.3), case (§3.4), and a variety of common derivational functions that typically fall under the heading of "mode" (§3.5).

3.1 Root classes and derivation

3.1.1 Pronouns, demonstratives and deictics

There are almost always distinct pronominal paradigms for the nominative, genitive and oblique cases. A typical example in this respect can be seen in the Maranao pronouns in Table 4 (McKaughan 1958, Kaufman 2007). The person and number features distinguished in Maranao are also typical for the CSP zone. All languages distinguish the inclusive and exclusive first person plural but not all languages have a distinct dual form, as Maranao does.

Table 4. Maranao pronoun paradigm

	Nom (bound)	**Nom (free)**	**Gen (bound)**	**Obl (free)**
1sg	(a)ko	sakən	akən ~ ko	rakən
1+2 dual	kami	səkami	(a)mi	rəkami
1pl.ex	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

Demonstratives and deictics in CSP languages typically distinguish three types of proximity: speaker proximate, hearer proximate and distal. Some languages, such as Matigsalug Manobo, shown in Table 5, distinguish four grades of proximity in deixis, although even in this language, the demonstratives only show the canonical three-way distinction.

Table 5. Matigsalug Manobo locative pronouns (Wang et al 2006:28)

kayi, dini	here
due	there (within reach)
dutu	there (beyond reach but within sight)
diyeʔ	way over there (nonspecific/out of sight)

Deictics are in most languages derived transparently from demonstratives with one of the PAn locative/directional markers *sa, *ka, *di (see Ross 2006 and Blust & Trussel ongoing for the reconstruction of these morphemes). The Sama languages, which have a richer inventory of prepositions beyond those based on *sa, *ka, *di, make use of these in their deictics, as well. Akamine (2005:388) lists four demonstratives for Southern Sama: *itu* (speaker proximate), *ilu* (hearer proximate), *inaʔan* (medial "location away but not far from both speaker and hearer"), *ili* (distal) and four deictic counterparts with an *m-* formative (*mitu, milu, minaʔan, mili*), a reduction of Sama locative *ma*.

Demonstratives and deictics can serve as unmarked predicates and deictics can typically form voice marked predicates with the addition of a reflex of the PMP causative *pa-, as in Tagalog (N).

(N) p<um>a-riːto
<AV>CAU-here
'to come here'

3.1.2 Nouns and verbs

The categorization of lexical roots in CSP languages has been a topic of interest beginning with Bloomfield's study of Tagalog and continuing until the present and is taken up more fully in Chapter XX. Lexical roots in CSP languages can almost always be used in isolation but typically have a patient-oriented interpretation in such contexts, as exemplified in (A) and (B).

(A) magkano aŋ na:kaw mo?
how.much NOM steal 2s.GEN
'How much was your heist?'

(B) a:wit ko iyan
sing 1s.GEN that.NOM
'That's my song.'

It does not seem to be the case that Central Philippine roots are completely acategorial nor does zero conversion seem a particularly apt way to describe the flexibility. Zero morphology fills a very particular role in the historical Austronesian voice paradigm (Wolff 1973, Ross 2002). Namely, it signals the dependent form of actor voice verbs, which in the Central Philippine languages are used primarily in intransitive imperatives. This can be seen in the contrast shown in (C), where a bivalent predicate cannot be used in its bare form as a transitive imperative.

(C) a. la:kad na!
walk already
'Walk!'

b. *na:kaw na!
steal already
(for 'Steal (it)!')

Outside these clearly verbal contexts, bare roots have a surprisingly strong entity-denoting interpretation (Kaufman 2009a, 2017). It is still unclear at what point this feature of Philippine languages developed, as the use of bare roots has not been explored very systematically in languages other than Tagalog. Judging from examples such as (X) and (Y), it would seem that Tboli, a CSP outlier, also gives a canonical nominal interpretation to what would typically be considered event-denoting roots. This is especially apparent in (Y), where the root appears to denote the result of an action, much like what we find in Tagalog and other Central Philippine languages. Awed et al (2004:78) also give examples of instrument interpretations of bare roots.

(X) tey nagaw boŋ nagaw-en ŋaʔ tu.
INTNS steal big steal-3S.GEN child that
'It was a big theft, that child's stealing.' (Awed et al 2004:428)

(Y) hilu lasaʔ deŋ t<n>asaʔ ye beŋ?
how.many lay.out PRF flatten 2PL wall
'How many pieces laid out have you flattened (for) walling?' (Awed et al 2004:363)

Conversely, voice and aspect morphology can convert nearly all lexical roots into event denoting predicates. Central Philippine languages do not require an extra step of verbalization beyond voice marking to turn notionally nominal roots into event denoting predicates, nor do CSP languages possess anything that can be called a copula. This is apparently true for the Bilic and Sama languages on the southern periphery, as well, as shown in Tboli (X) and Yakan (Y). In (X), we see canonical entity denoting roots like *gunù* 'house' and *ówóng* 'boat' functioning as event denoting predicates with the addition of the actor voice marker *<m>/m-*. Similarly, in (Z),

luma? ‘house’ is converted into an event denoting predicate with the general actor voice prefix *mag-*.

(X) g<m>unu?=le te=bulul le=ma?
<AV>house=3p.NOM OBL=mountain PL=father
‘Father and his household have a house on top of the mountain.’ (Awed et al 2004:231)

(Y) m-ówóŋ=me te=sbu?
AV-boat=1p.ex.NOM OBL=Sebu
‘We rode in a boat on Lake Sebu.’ (Awed et al 2004:450)

(Z) sinna=ku mag-luma? dem puweblo
like=1s.GEN AV-house in town
‘I like to live in town.’ (Behrens 2002:236)

The vexed distinction between verbs and nouns has been discussed by a host of authors and its significance both to the synchronic analysis of Philippine languages as well as the historical development of Austronesian continues to be debated.[15] Most descriptive works have been satisfied to call words with aspect inflection “verbs”, those prefixed with a reflex of adjectival **ma-* or serving as modifiers as “adjectives”, and those entity denoting roots without such markers as “nouns”. It has been widely acknowledged, however, that these distinctions play little role in the higher, clause level syntax of Philippine languages. All three major word classes, verb, noun and adjective, are equally (un)marked as predicates, modifiers and arguments. Moreover, there is great flexibility at the root level; almost any lexical root can enter into any paradigm in a typical Central Philippine language. This has not been fully appreciated in the literature, possibly due to an over-reliance on dictionaries, which cannot possibly provide all potential derivations of all roots. To illustrate with a simple example, none of the major Tagalog dictionaries (English 1986, Panganiban 1972) appears to list an actor voice form for Tagalog *la:ŋit* ‘sky’, and yet we find attestations such as (X).[16]

(X) l<um>a:~la:ŋit na:wa aŋ iyoŋ kaluluwa?
<AV.BEG>IMPRF~sky OPT NOM 2s.OBL:LNK soul
‘May your soul be “heavening”.’

Panganiban (1972:610), but not English (1986), furthermore notes the existence of *ni-la~la:ŋit-∅* BEG-IMPRF~sky-PV, a patient voice form with a somewhat idiosyncratic definition, ‘the thing one desires the most’, but attestations of this voice marked form are also found in other aspects and so it cannot be considered frozen. Similarly, with the widespread descendants of PMP *ma-, which indicates possessing a property denoted by the root, we find a great deal of flexibility, in addition to some categorial grammatical distinctions. Central Philippine languages can attach a reflex of *ma- to roots which we might *prima facie* consider action denoting, like *la:kad* ‘walk’,

[15] For a sampling of different approaches to the problem see Capell (1964), Starosta, Pawley, and Reid (1982), De Guzman (1996), Foley (1998), Gil (1995, 2000), Himmelmann (2008), Kaufman (2009a, 2009b), Ross (2009), Hsieh (2018).

[16] The more usual form for this usage contains the oblique marker as part of the stem to which the actor voice attaches, i.e. *s<um>a~sa-laŋit* <AV.BEG>IMPRF~OBL-sky ‘to be in heaven’.

as shown in (Ta). But for roots that are hard to construe as a property, as in (Tb), uninflected *ma-* forms may be judged unacceptable.

(T)a. ma-laːkad siya
STA-walk 3s.NOM
'S/he has a lot of errands.'

b. ??ma-sampal siya
STA-slap 3s.NOM
'S/he has a lot of slaps.'

A more categorical distinction exists between static and dynamic property denoting roots in CSP languages (Wolff 1993, Himmelmann 2006, 2008). This is best illustrated by two Tagalog roots with nearly identical meanings *tuwaʔ* and *saya*. The former is treated as dynamic and requires aspect marking in finite contexts whereas the latter is uninflectable for aspect, as shown in (Y).

(Y)a. *ma-tuwaʔ siya
STA-happy 3s.NOM

b. ma-saya siya
STA-happy 3s.NOM
'S/he is happy.'

c. na-tuː~tuwaʔ siya
STA.BEG-IMPRF~happy 3s.NOM
'S/he is happy.'

d. *na-saː~saya siya
STA.BEG-IMPRF~happy 3s.NOM

Similarly, Daguman (2004:242) describes two sets of property denoting words in Northern Subanen. One set, which she terms "adjectives", exemplified in (Za), cannot be inflected for aspect while another set, which she terms "stative verbs", exemplified in (Zb-c), must be inflected for aspect in finite environments. Subanen seems unique among Philippine languages in distinguishing nouns and adjectives with an obligatory proclitic (G=, whose surface realization is dependent on the following segment), as seen in (Za).

(Z)	a.	ŋməlәŋkaʔ	b.	mətunag	c.	mitunag
		G=mə-lәŋkaʔ		mə-tunag		mi-tunag
		SCM=ADJ-lazy		STA.PAT.IRR-melt		STA.PAT.RL-melt
		'lazy'		'will melt'		'melted'

Other hard constraints exist at the innermost morphological level, too. For instance, some event denoting roots are specified as being non-volitional and thus must occur with non-volitional *ka-/*ma- in their most basic forms (e.g. Tagalog *kiːtaʔ* 'see') while others event denoting roots appear not to license an affected patient and thus cannot occur with a reflex of patient voice *-en.[17] For the most part, though, it seems that inner morphology in at last the Central Philippine languages is constrained more by pragmatics and convention than by any categorial grammatical principles. Much the same can be said for verbalization in English; while 'to tree' and 'to sky'

[17] Wolff's (1972:xvii-xx) Cebuano dictionary most likely represents the most sophisticated and complete root classification of any Philippine language to date. All roots are classified with respect to which modes and voices they do *not* appear with. In the actor voice, this relates primarily to the markers *<um>, mag-* and *maka-*, whose use bears some relation to transitivity and volitionality but in an unpredictable way. For non-actor voice morphology, each root is classified by whether it can appear with the patient voice, locative voice, or conveyance voice, among other variables.

may appear odd to the average English speaker, they are not ungrammatical and indeed, they have arisen with specialized meanings in certain English speech communities.

In contrast to the general freedom we find at the innermost level, categorizing a root with a particular morpheme at this level will sharply delimit the further morphological potential of that stem. For instance, voice morphology can apply to practically any lexical root in CSP languages and once voice morphology applies, aspect morphology can always follow. However, outside of a few well defined cases, aspect morphology can only apply to voice marked stems. In the same vein, a vast number of roots in any given CSP language can take a reflex of adjectival (uninflected) *ma-, as mentioned above, and can then take a special plural marker used with such formations (typically <*aŋ*> or CV reduplication) but this plural marking is highly dependent on the category of its host. CSP morphosyntax, and perhaps Philippine morphosyntax more generally, thus shows a surprising degree of freedom both at the root level, where almost any root can take event, property or entity related morphology without requiring category changing derivation, and at the highest phrasal level, where almost any word can play the role of argument, predicate or modifier in an equally unmarked manner. However, at the intermediate, word-internal level, categorization and morphological potential is strictly delimited, as seen in the fact that many CSP languages have three different morphological means of marking plurality depending on what categorizing morphology the stem has already taken.

3.1.3 Prepositions

Moving beyond the "cardinal" categories of adjective, verb and noun, we find that prepositions are not a highly developed category in CSP languages and have generally been grammaticalized from other categories in the relatively recent history of most languages. The oblique case typically covers the functions of English *to, towards, from, on, in, for*, as seen in the various interpretations of Tagalog (Z).

(Z) l<um>undag sa duyan aŋ baːtaʔ
<AV.PRF>jump OBL hammock NOM child
'The boy jumped [from off/over/onto/on] the hammock.' (Cena 1971:134)

Most CSP languages do, however, possess a class of prepositions which typically take oblique case phrases as complements.[18] Such examples are shown in (Xa) and (b) for Bikol and Tagalog, respectively.[19]

(X)a. sagkod sa oːras na ini
until OBL time LNK this
'until this time'

b. tuŋkol sa naŋ-yaːri
about OBL AV.PRF.DIST-happen
'about what happened'

[18] Note that there is some debate whether to consider oblique case as a preposition (Himmelmann 2005:146-7, 2016) or a case marker (Kaufman 2009a).

[19] Syntactic tests in Tagalog show that the phrase headed by a preposition behave similar to adjuncts and oblique clauses rather than constituting a secondary predicate. For instance, the preposition is fronted together with the following oblique phrase the adjunct focus construction discussed in Schachter & Otanes (1973) and Kroeger (1993).

Many roots that play a prepositional role still have a full derivational range. For example, Tagalog *mulaʔ* 'from' combines with an oblique phrase in its prepositional function but it can also serve as a stem for voice morphology meaning 'to originate from'.

Despite the paucity of true adpositions in CSP languages, there is also very little that resembles true serial verb constructions. We do, however, find action denoting roots taking the multifunctional *pa-* prefix with a directional function in several Central Philippine languages. This can be seen in Tagalog (X) and Cebuano (Y). These *pa-* marked words cannot be inflected for voice or aspect in this function and they introduce their complements either with the linker, as in (X), or an oblique phrase, as in (Y).

(X) <um>alis sila pa-punta=ŋ manilaʔ
<AV.BEG>leave 3p.NOM DIR-go=LNK Manila
'They left heading towards Manila.'

(Y) mi-dagan siya pa-iŋun sa gawaŋ
AV.PRF-run 3s.NOM DIR-toward OBL door
'He ran toward the door.' (Wolff 1972:384)

A possible degrammaticalization of this directional *pa-* yields an independent directional marker in the Sama languages, as shown in Pangutaran Sama (Z).

(Z) t<um>udʒu aʔa pa lumaʔ saupak
<AV>toward person OBL house Saupak
'The man is headed for Saupak's house.' (Walton 1986:87)

The Sama languages are altogether richer in prepositions. Akamine (2005:385) enumerates five for Simunul Island Sama: *leʔ* agent/reason, *ma* location, *ni* goal, *min* source and *maka* instrument, comitative. Unlike the prepositions of other CSP languages, these take noun phrase complements directly and not oblique case phrases.

Every CSP language also possesses roots that are commonly used as locative nouns with genitively marked complements, as in (Wa). There is little that differentiates such nouns from non-locative nouns but in Tagalog several of them also appear with a frozen locative prefix *i-*, which typically gives a more generalized meaning, as shown in (Wb).

(W)a. sa taʔas naŋ baːhay
OBL top GEN house
'on top of the house'

b. sa i-taʔas
OBL LOC-top
'in the upper area'

Note that the same type of genitive construction is used to introduce instruments in Tagalog, as in (Y), among other languages. The word translated with a preposition in (y) is a complex gerund-like derivation based on the stem *pagitan* 'between' with the distributive prefix.

(Y) sa pamamagiːtan naŋ bato
OBL intervention GEN rock
'using/with a rock'

Finally, functions typically carried out by adpositions in other languages are also carried out by a type of reduced clause in at least some Central Philippine languages. Here, a bare entity denoting stem denotes the relation while its complement is expressed by a nominative phrase, as in Cebuano (X) and Tagalog (Y).

(X) iya=ŋ gi-pusil-Ø [gaːmit aŋ sumpak]
3s.OBL:LNK PRF-shoot-PV use NOM improvised_gun
'He shot (him) with an improvised gun.'

(X) <um>uwiʔ ako [kasaːma aŋ kapatid ko]
<AV.BEG>go_home 1s.NOM companion NOM sibling 1s.GEN
'I went home with my sibling.'

Such sentences can appear to have two subjects even though the nominative phrase in the reduced clause cannot be questioned or relativized as a normal nominative argument of a full main or subordinate clause can.

3.1.4 Adverbs

Manner adverbs are typically simple property denoting words although they are introduced either as genitive/oblique marked arguments, as in the Agutaynen example in (T), or as fronted phrases connected to an event denoting predicate via the linker, as in Matigsalug Manobo (W).

(T) nag-kanta tanandia ta ma-sinlo.
AV.PRF-sing 3s.NOM OBL ADJ-nice
'She sang nicely.' (Quakenbush et al 2010:32)

(W) manekal ne eg-basuk
strong LNK AV-till
'energetic at cultivating' (Wang et al 2006:89)

The categorial status of adverbials in constructions like (W) is open to debate, as the "adverbial" could also be interpreted as a simple adjective. Kaufman (2006) points out that there is at least a semantic argument for treating this as an adverbial expression rather than a plain property predication in that sentences such as Tagalog (S) do not entail that the subject is delicious.

(S) ma-sarap siya=ŋ mag-luːtoʔ
ADJ-delicious 3s.NOM=LNK AV-cook
'S/he cooks deliciously.' (Not: 'S/he is delicious at cooking.')

Topicalization can also introduce property denoting words in an adverbial function, but this position typically hosts higher adverbs, such as comparative, quantificational, subject-oriented, modal and speech act adverbs. Tagalog (U) shows an example of a frequency adverbial in a topic position.

(U) madalas ay nag-i~imbita naŋ panauhiŋ taga-pag-salitaʔ...
often TOP AV.BEG-IMPRF~invite GEN guest PROF-TR-word
'Often (they) invite a guest speaker…' (Casanova & Rubin 2001:180)

There are a few morphological formations that can be considered strictly adverbial. Property denoting roots affixed with *ma- -an* appear to modify aspect inflected forms or gerunds (although plain *ma-* forms can fill this role, as well) but cannot be used to modify or predicate non-event denoting arguments, as seen in (S) and (T). It is unclear to what extent dedicated manner adverb formations such as these exist in other languages of the region. Few such derivations have been described in the literature.

(S) t<um>akbo aŋ kabayo naŋ [ma-bilis / ma-bili:s-an]
<AV.BEG>run NOM horse GEN STA-speed STA-speed-ADV
'The horse ran fast.'

(S) [ma-bilis / *ma-bili:s-an] aŋ kabayo
STA-speed STA-speed-ADV NOM horse
'The horse is fast.'

Aspectual modifiers, most prominently, those meaning 'still', 'already' and 'again', as well as a large number of evidentials and mood markers, are expressed as second position clitics in all Philippine languages. Temporal and locative adverbs (deictics), such as are often expressed within a single word although these often behave syntactically as oblique phrases.

3.2 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 and this appears to be largely true for Austronesian languages as a whole, which do not mark tense with bound morphology. Voice and aspect are grammatically prominent and paradigmatically interconnected in most Philippine languages, despite being notionally independent.[20] These two categories are handled separately in this and the following subsection but it will be helpful to make a first approach to voice/aspect paradigms from a historical perspective. Building on Wolff (1973), Ross (2002) reconstructs the PAn voice/aspect paradigm roughly as shown in Table 6.[21] The first three rows, Ross terms "indicative" and the bottom two rows, he terms "non-indicative", corresponding to Wolff's (1973) "independent" and "dependent", respectively. I return to the usage of the indicative versus the non-indicative below in §3.3 and §4.5.

[20] See Reid 1992; Ross 1995, 2002; Himmelmann 2005; Zeitoun and Huang 1997; and Zeitoun et al. 1996.

[21] I have retained the simpler paradigm of Ross 2002 while updating the terminology ('imperfective' instead of 'durative') and the reduplicant (Ca~ instead of CV~) to fit with Ross 2009.

Table 6. Ross's (2002:33, 2009:296) reconstruction of the PAn voice paradigm

	Actor voice *<um> / *∅	**Patient voice** *-en / *-a, *-u	**Locative voice** *-an / *-i	**Conveyance voice** *Si- / *-áni, *-ánay
Neutral	<um>√	√-en	√-an	Si-√
Perfective *<in>	<um><in>√	<in>√	<in>√-an	Si-<in>√
Imperfective *Ca~	<um>Ca~√	Ca~√-en	Ca~√-an	Si-Ca~√
Atemporal	√	√-u, √-a	√-i	án-i+√, √-áni
Projective	<um>√-a	√-aw	√-ay	án-ay+√, √-ánay

The indicative mood morphemes for each voice are shown before the slash in the first row and those for the non-indicative mood following the slash. There is one irregularity in this paradigm, which is inherited in the vast majority of Austronesian languages: the patient voice suffix *-en, does not co-occur with the perfective *<in> infix. If the reconstruction in Table 6 is correct, one innovation which typifies most CSP languages is the co-occurrence of *<in> with CV reduplication and the consequent three-way aspect distinction between perfective, progressive and prospective rather than the two-way perfective/imperfective distinction reconstructed by Ross. This can be seen in the Tagalog voice/aspect paradigm in Table 7.

Table 7. Fragment of the Tagalog voice aspect paradigm

ba:sag 'break'	**Actor**	**Patient**	**Locative**	**Conveyance**
Neutral	b**<um>**a:sag	basa:g**-in**	basa:g**-an**	**i-**ba:sag
Perfective	b**<um>**a:sag	b<in>a:sag	b<in>a:sag**-an**	**i-**b<in>a:sag
Progressive	b**<um>**a:~ba:sag	b<in>a:~ba:sag	b<in>a:~ba:sag**-an**	**i-**b<in>a:~ba:sag
Prospective	ba:~ba:sag	ba:~basa:g**-in**	ba:~basa:g**-an**	**i-**ba:~ba:sag

We find additional irregularities in CSP languages, as well. The Tagalog <*um*> infix marks both the neutral and perfective of the actor voice (via the historical elimination of <*in*>) while plain CV reduplication marks the prospective of the actor voice without the help of <*um*>. Many CSP languages have externalized certain historical infixes as prefixes or have otherwise eliminated them, especially in the perfective actor voice where two infixes historically co-occurred. On the other hand, Subanen (X) shows how this combination of infixes continues to be productive in other CSP languages.

(X) s<in><um>ubaʔ=na su ŋa ŋ=maʔistra
<PRF><AV>descend=already ABS PL SCM=teacher
'The teachers already went up to the mountain.' (Daguman 2004:286)

Central Philippine languages generally display three primary aspects which can be termed perfective, progressive and prospective.[22] It seems possible to derive the three-way aspect distinction using two atomic features corresponding precisely to reflexes of *<in> (which may surfaces as *n-* or <*i*> in the actor voice) and *CV reduplication, as shown in (X).

(X)

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).[23]

Aspect marking is most often obligatorily on finite verbs in many languages of the CSP zone although in some languages, such as Cebuano and Agutaynen, a single form will be used for the imperfective/prospective and the infinitive and thus the common three-way aspect distinction is reduced to two. Such languages can be said to conflate the historical unmarked and prospective aspects into a general 'unrealized' inflection (Reid 1992:74). In Sama languages, aspect marking is not obligatory nor is it instantiated by bound morphology. Sama and Tboli express aspect with a class of free standing adverbs in pre-predicate position. This can be seen as part of a southern pattern that typifies the languages of Indonesia and includes the southern CSP zone. It is see in Tboli (U) and Simunul Sama (D).

(U) deŋ m-atù Mantil
already AV-win Mantil
'Mantil won.' (SOURCE)

(D) bey naŋis anak-anak
CPL AV-cry child
'The child cried.' (Akamine :383)

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

[22] The prospective 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.

[23] PAn *CV~, which may have originally marked the imperfective or durative 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].

(X) ka-ra:~ratiŋ ko laŋ
RCT$_1$-RCT$_2$~arrive 1s.GEN only
'I just arrived.'

(Y) pa-ratiŋ na ako
IMMD-arrive already 1s.NOM
'I'm about to arrive.'

Other languages of the CSP area appear to have expanded this system more dramatically using the PMP mode prefix *paR- as a durative, among other strategies. Zorc (1977:119) analyzes Aklanon as having six aspects, as shown in Table 8, although others have analyzed his AORIST as a different mood rather than simply another aspect.[24]

Table 8. Fragment of the actor voice aspectual paradigm in Aklanon (CITE)

	IMPERFECTIVE			PERFECTIVE		
	ACTUAL	CONTINGENT	AORIST	ACTUAL	CONTINGENT	AORIST
PUNCTUAL	ga-√	ma-√	<um>√	<um>√	<um>√	<um>√
DURATIVE	naga-√	maga-√	ga-√	nag-√	mag-√	mag-√

The use of a *paR- reflex as a durative can also be seen in Cotabato Manobo (Kerr 1988:8) (D), where *eg-* (< PMP *paR-) indicates the progressive and CV reduplication no longer plays any role in the aspect paradigm.

(D) Prospective √-en
Progressive eg-√-en
Perfective <in>√

The neighboring Danao languages also use a reflex of *paR- (*pe-*) for what is signaled by reduplication in Tagalog, as seen in Table 9.

Table 9. Maranao voice aspect paradigm

	Actor	**Patient**	**Locative**	**Conveyance**
Neutral	t<om>abas	tabas-en	tabas-an	i-tabas
Perfective	t<omi>abas	t<i>abas	t<i>abas-an	i-ni-tabas
Progressive	pe-tebas	pe-tebas-en	pe-tebas-an	i-pe-tebas
Immediate prospective	tebas	tebas-en	tebas-an	i-tebas

In many languages, disyllabic reduplication indicates repetitive action and can be considered an aspectual category as well, although it is rarely included as part of the basic aspect paradigm in the descriptive literature and perhaps rightly so; unlike CV imperfective or

[24] Although nothing in Table 8 distinguishes the contingent and aorist in the perfective, these are distinguished in other voices.

progressive reduplication, disyllabic repetitive reduplication is never seen to interact with mood or negation.

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.ABIL, typically yields *n-* initial forms without infixation (e.g. *naŋ-*, *naki-*, *nag-*, *naka-*). This "externalization" of *<in> postdates PMP, as we also find CSP languages that reflect *m<in>aR- as *mig-* rather than *nag-*, 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 (see §4.3). An example of this is seen in Sarangani Manobo, where aspect is marked on the verb in (Q) but through negation in (R). Similar examples could be produced for most Bisayan languages, as well. In Tagalog, negation does not interact at all with aspect marking, but this is atypical for Central Philippine languages.

(Q)a.	t<om>edogi	se bayi		b.	t<im>edogi se bayi
	<AV>sleep	NOM woman			<AV>sleep NOM woman
	'The woman will sleep.'				'The woman slept.'
(R)a.	edek tedogi	se bayi		b.	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.' (Dubois 1976:20)

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", as marked by enclitics, typically descendants of PMP *=dena 'already' and *=pa 'still'. Aspectual clitics in Philippine languages play a larger role than might be gleaned from their English glosses and are near obligatory in certain types of contexts. Outer aspect markers are both morphologically external to perfective, progressive and prospective morphology and also involve higher level pragmatics. Reflexes of PMP *=dena 'already' place a situation before an *expected* time while PMP *=pa 'still' places a situation after such a time.

3.3 Voice

Voice is a pivotal feature of the morphosyntax of all Philippine and Philippine-type languages.[25] The Philippine-type alignment system is generally understood to select a particular participant as the nominative argument (or absolutive, depending on the analysis) 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 subjects and topics remains debated.

[25] 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. Lobel (2013:152) further notes that the first step in the reduction of the voice system is typically the merger between conveyance voice and patient voice, which can be found across several languages of the Philippine periphery.

Agents of non-actor voice verbs are uniformly expressed in the genitive case in Philippine languages. Notional objects, when not selected by the voice morphology to become nominative arguments, are either expressed as genitives (as in Tagalog), as obliques (as in Cebuano), with something like a dedicated object marker (as in Maranao and Ivatan) or with the linker (as in the Bikol example below and more generally in Kapampangan). Case is examined further below in §3.4.

The four primary voices are the actor voice, patient voice, locative voice and conveyance voice, as seen earlier in Table 10.[26] The exponence of these voice markers in CSP languages do not differ drastically from their PMP reconstructions, as can be seen below:

Table 10. Common CSP reflexes of PMP voice markers

Voice	**PMP reconstruction**	**Common CSP reflexes**
actor voice	*<um>	<um>, m-, mu-
patient voice	*-en	-in, -un, -ɨn
locative voice	*-an	-an
conveyance voice	*(h)i-[27]	ʔi-, hi-, ∅

The basic use of the voice markers is very consistent across the CSP range with the exception of the Sama languages and, to a lesser extent, the Bilic languages. The system can be illustrated with the Naga Bikol examples in (E).

(E)a. nag-bakal aku=ŋ bagas
AV.PRF-buy 1s.NOM=LNK rice
'I bought rice.'

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

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

d. i-b<in>akal ko si hwan ki bagas
CV-<PRF>buy 1s.GEN NOM Juan OBL rice
'I bought some rice for Juan.' (McFarland 1974:104-105)

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 the "proto-

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

[27] 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.

agent" as the nominative argument; the patient voice typically selects an affected 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. While this collection of functions for the conveyance voice may appear quirky, they can all be traced to the earliest stage of this paradigm in Austronesian family.

The proper treatment of these voice markers remains an area of endless theorization and major debate in Austronesian linguistics. The earliest published analyses carried out by Spanish colonial linguists and inherited by Bloomfield 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 were not equivalent to the Indo-European passive, the latter which was a marked construction used primarily to background the agent and the latter which are fully transitive. In the symmetrical analysis of Philippine-type voice (Foley 1998, 2008; Himmelmann 1996; Riesberg 2014), the system represents a unique type of alignment in which no voice is more marked than any other, thus standing natural in opposition to accusative and ergative languages which often possess unmarked transitive clauses and marked passives and anti-passives. For the vast majority of CSP languages, it also holds true that there is no morphologically unmarked voice, just as in the Tagalog paradigm seen earlier.[28] Proponents of an ergative analysis of the Philippine voice system (Starosta et al 1982, De Guzman 1988, Gerdts 1988, Aldridge 2002, Liao 2004) have long argued that the actor voice appears less transitive than its non-actor voice counterparts. Although this is not the place to review the arguments for one analysis over another (but see Kaufman 2017), the principles of voice selection require basic explication.

There is widespread agreement that definiteness, specificity or some related notion of referentiality largely determines voice selection (see Wolff 1966, Wolfenden 1971, Schachter 1976, 1977, McFarland 1978 for early treatments). The chart in (O) abstracts away from many complications, additional factors and cross-linguistic variation (Schachter 1976, Naylor 1986, Adams and Manaster-Ramer 1988, McFarland 1978, Latrouite 2011, Nolasco 2003) but captures 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.

(O)

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

The pattern above conspires to avoid the expression of a definite argument as a non-directional, non-nominative object. Definite arguments are completely felicitous as genitive agents and nominatives, while directional arguments are felicitously expressed in the oblique case regardless

[28] In languages like Tagalog, the perfective aspect of the patient voice and the prospective aspect of the actor voice is unmarked, but this does not tilt the analysis of the language towards an accusative or ergative pattern.

of their definiteness. 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). There are cases when this configuration is unavoidable, as when the agent itself is in the predicate position in an apparent cleft, such as (U). Here, even though the patient is definite and highly affected by the action, it cannot be expressed as the nominative argument because the actor is in the predicate position and the verb must thus select it as the nominative argument. The patient is thus exceptionally expressed as an oblique argument in such a case. Note that this is otherwise impossible, as shown by (Ub), where the same patient is expressed in the oblique but without the extenuating circumstance of the cleft-like structure in (Ua).

(U)a. ako aŋ p<um>atay sa kanya
1s.NOM NOM <AV.BEG>kill OBL OBL.3s
'I'm the one who killed him/her.'

b. *p<um>atay ako sa kanya
<AV.BEG>kill 1s.NOM OBL OBL.3s

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

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

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

It should not be assumed that the patient voice is restricted to semantically bivalent predicates. Well known examples of the type in Tagalog (X) show that patient voice also selects affected subjects of monadic and even entity denoting predicates.

(X)a. la~laŋgam-in aŋ=asukal
IMPRF~ant-PV NOM=sugar
'The sugar will be "anted".'

b. s<in>i~sipon-Ø ako
<BEG>IMPRF~flu-PV 1s.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 (V), but it can just as easily "promote" an adjunct to become the nominative case argument, as in (W).

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

(W) 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 by assumed, as seen in the minimal pair in (Q).

(Q) a. k<in>ain-Ø ni maria aŋ isdaʔ
<BEG>eat-PV GEN Maria NOM=fish
‘Maria ate the fish.’

b. k<in>ain-an ni maria aŋ isdaʔ
<BEG>eat-LV GEN Maria NOM fish
‘Maria ate from/at the fish.’

The conveyance voice (PAn *Si-) is difficult to characterize semantically in a unified manner. It selects as the nominative argument benefactees, instrumentals and objects conveyed away from the agent. These seemingly disparate functions can be disambiguated in a number of CSP languages with an emergent marker for each function, e.g. Tagalog *i-pag-* BENEFACTIVE, and *i-paŋ-* INSTRUMENTAL, although the bare *i-* prefix in Tagalog is still as polysemous as its historical source.

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. The CSP languages are crucial in understanding the role of the non-indicative forms in PMP, as they are preserved more faithfully here than in languages of the northern Philippines. In most MP languages outside the Philippines, the distinction between the indicative and non-indicative forms are also merged. In the northern and central Philippines, the paradigms are generally merged in favor of the indicative paradigm and are reduced in various ways south of the CSP zone.[29] Wherever the non-indicative paradigm is preserved, it is used in the imperative. This is seen in Batangas Tagalog (D) and Maranao (E). 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 seen in (D). But as seen in Maranao (E), this does not hold for all languages.

(D) buks-i (*mo) aŋ pintuʔan
open-LV.DEP 2s.GEN NOM door
‘Open the door!’

(E) tabas-a ŋka so dinis
cut-PV.DEP 2s.GEN NOM cloth
‘Cut the cloth!’ (McKaughan 1958:25)

[29] 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/active voice. On the other hand, the independent patient voice *-en and conveyance voice *Si- are almost entirely lost south of the Philippine languages, although the distinction may be carried out through different morphological means.

In many Central Philippine languages, the dependent paradigm is also used in the negated perfective, as shown by Wolff (1973) for Samareño (I). 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).

(I)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 1s.GEN return-PV.DEP NOM beer
'I did not go back after the beer.'

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

d. waraʔ niya pilak-an an basuːra
NEG.EXT 3s.GEN throw.away-CV.DEP NOM garbage
'He did not throw the garbage away.' (Wolff 1973)

The dependent forms are also employed in temporal adjuncts, as seen in (I) and (J) (Stevens 1969, Zorc 1977). These contexts are particularly interesting as the voice morphology selects a particular argument to promote, *kanya suwildu* in (I) and *baŋku* in (J), but no argument actually surfaces with nominative case.

(I) pag-ta-tág-an=niya [sa kanya suwildu] [kanya nanay]...
SUB-ASP-give-IV.DEP=3s.GEN OBL 3s.GEN earning 3s.GEN mother
'When he gives all of his earnings to his mother…' (Zorc 1977:139)

(J) pag-liŋkur-i=niya han baŋku, na-rubaʔ
SUB-sit-LV.DEP=3s.GEN GEN.DEF bench STA.PRF-break
'When he sat on the bench, it broke.' (Zorc 1977:139)

The sentences in (5) show the use of the dependent conjugation in the recent perfective aspect Bikolano. Here again we find the locative voice in (5a) and the patient voice in (5b) but no nominative case where it would be expected, as the recent perfective behaves similar to the adjunct forms above.[30]

(5)a. ka-ʔiː~ʔinum-i pa sana nya kani=ŋ buːti=ŋ ini
RCNT-RCNT~drink-LV.DEP still only 3s.GEN OBL=LNK bottle=LNK this
'He just drank from this bottle.' (Stevens 1969:13)

[30] Lee (1964) discusses a Maguindanao paradigm with the *ka-* prefix having a seemingly identical function. But in Maguindanao, the *ka-* marked forms take aspect morphology but not voice morphology and are thus termed "non-focus verbs" by Lee. They are probably best analyzed as aspectually marked gerunds.

b. ka-ba~bakal-a ku pa sana ka-ini sa iya
RCNT-RCNT~buy-PV.DEP 1s.GEN still only OBL-this OBL 3s
'I just bought this from him.' (Stevens 1969:13)

The basic four-way voice distinction is maintained in almost all the CSP languages. The Sama languages, however differ from their Philippine neighbors in following a more canonical ergative pattern. In these languages, an unmarked transitive clause has supplanted the patient voice, as shown in Yakan (Z).

(Z) pogpog aʔa sawe-hin
hit person snake-DEF
'A person hit the snake.' (Brainard and Behrens 2002:113)

In this type of clause, the agent must be adjacent to the verb, which remains unmarked for voice. A second type of bivalent clause in Yakan introduces the agent with a marker *we'* (glossed ERGATIVE by Brainard and Behrens 2002) and marks the verb with a transitive infix <*in*>, as indicated in (R). Unlike in the previous clause type, here the order of the *we'* phrase is flexible.

(R) p<in>ogpog (weʔ aʔa) sawe-hin (weʔ aʔa)
<TR>hit ERG person snake-DEF ERG person
'A person hit the snake.' (Brainard and Behrens 2002:113)

Bivalent verbs in Yakan take an antipassive prefix (glossed by Brainard and Behrens 2002 as INTRANSITIVE) to map a proto-agent to the absolutive argument, as in (Y).

(Y) ŋ-uruŋ siːn iye para si iskulan
INTR-give money 3SG.ABS for OBL school
'He gave money for the school.' (Brainard and Behrens 2002:49)

A similar situation holds for Pangutaran Sama (Walton 1986), as seen in (H), where an unmarked bivalent predicate is interpreted as fully transitive and its actor focus/antipassive counterpart is interpreted less transitively (in the sense of Hopper and Thompson 1980).[31]

(H)a. Ø-bonoʔ sultan bantaʔ na
UF-kill king enemy 3s.GEN
'The king killed his enemy.'
(Walton 1986:120)

b. N-bonoʔ sultan bantaʔ na
AF-kill king enemy 3s.GEN
'The king kills/fights some of his enemies.'

This pattern differs from other CSP languages, in which both arguments of a transitive clause are typically preceded by case markers and the genitive/ergative marked agent tends to follow the verb immediately across the board. Thus, while Yakan and other Sama languages maintain voice distinctions to some extent, they have also developed unmarked transitive and intransitive verbs, unlike other CSP languages.

[31] I maintain Walton's glossing of the null prefix as a marker of undergoer voice and the nasal prefix for actor voice although the facts shown in (H) for this variety do not differ from Yakan.

The Tboli voice system has been reshaped by the general loss of suffixes and, to a lesser extent, the loss of 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 historical conveyance voice is left unmarked morphologically but still considered a distinct voice, as in other languages of Mindanao. Tboli agent voice, undergoer voice and instrumental voice clauses are exemplified in (X)-(Z). Note that only pronouns are case marked but a rigid word order, in which non-nominative agents must follow the verb immediately, serves to distinguish grammatical roles.

(X)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
IV-chop_down Walan it axe
'Walan chopped it down with an axe.' (Awed et al 2004:79, 25)

Just as there are minor aspects, we can also speak of minor voices. Minor voices appear to have been innovated more recently, often from combinations of inherited morphemes, and target adjuncts such as purposive clauses for promotion to the nominative argument. The complex Tagalog prefix *ika-* (historically composed of the conveyance voice *i-* plus the *ka-* prefix), selects a reason or purpose, as shown in (D). An etymologically and semantically equivalent form is seen in Sarangani Manobo, suggesting some antiquity to this minor voice.[32]

(D) ano aŋ ik<in>a-puː~punta niya doʔon?
what NOM REAS<BEG>-IMPRF~go 3s.GEN there
'What's his reason for going there?'

(E) yan se iŋke-opal ko
that NOM REAS-anger 1s.GEN
'That's why I became angry.' (Dubois 1976:67)

The CSP languages typically allow only one voice marker per word (Wolff 2002:439), but this is not the case in the languages of the northern Philippines. In the Cordilleran languages of North Luzon, conveyance voice **i-* combines with locative **-an* to form an unambiguous benefactive voice (Reid & Liao 2002:460) and predicates that employ the conveyance voice to select their object as the nominative argument, often maintain the conveyance voice prefix in the actor voice, yielding combinations such as *maŋ-i-* AV.DIST-CV-. Such combinations are vanishingly rare in the CSP languages.[33]

[32] Dubois (1976:67) terms this "Accessory Focus" and notes "The only verbs which can occur in Accessory focus are the ones which express emotion; the focused phrase indicates the reason or stimulus for the emotional response."

[33] Apparent combinations of voice markers do occur in the CSP languages when one voice marker derives the stem for the true voice marker. For instance, a Tagalog stem can be formed with locative nominalizer/voice marker *-an* and then go on to take the *mag-* actor voice prefix. Combinations of voice markers can also take on seemingly non-compositional functions, such as Tagalog *mag-tulug-tulug-an* AV-PRETEND~sleep-PRETEND 'to pretend to sleep', where both the reduplication and the *-an* suffix constitute multiple exponence of the 'pretendative'. But here there is no clear link between the pretendative function of *-an* and its more common locative voice function. Such cases are

3.4 A further look at case

As justice cannot be done here to the impressive variation found in CSP case markers and the issues in their reconstruction, I focus here only on a few salient points of typological interest and refer the reader to the considerable literature that investigates these forms from a comparative historical perspective (Reid 2002, Blust 2005b, Reid 2007, Ross 2006, Blust 2015).

In the Central Philippine languages, case is often expressed syncretically with other referential and even temporal features. For instance, 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 (E) and (F). The (a) examples show that indefinite possessors and genitive agents are introduced by *ki* while definite ones are introduced by *kan.*

(E)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' (McFarland 1974:161)

(F)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' (McFarland 1974:161)

Other varieties of Bikol make a subtle three-way distinction in referentiality, as seen for the Buhi dialect in Table 11. From the object marking in the examples in (O), we see that a generic object is marked by *nin*; a definite, but not yet "realized" object is marked by *nya*; and a definite, identifiable or "realized" object, is marked by *nyu.* As in Tagalog, the nominative phrase does not lend itself to an indefinite interpretation but still distinguishes what McFarland calls "definite" from "specific" arguments. In (Ob), because the action has not yet been realized, the subject receives the *a* marker. In (Oa) and (c), because the action has been realized, the subject receives the specific *yu* marker.

Table 11. Buhi Bikol case markers (McFarland 1974:164)

	Nominative	**Genitive**	**Oblique**
Indefinite	–	nin	sa
Definite	a	nya	
Specific	yu	nyu	

markedly different from *maŋ-i-* AV-CV- in Cordilleran languages, in which both the actor voice markers and the conveyance voice marker are playing a voice related role, the first determining the voice of the entire predicate and the second functioning as an applicative for objects moving away from the agent.

(O)a. aku yu nag-kaʔɨn nin aduːbu
1s.NOM NOM.SPEC AV.PRF-eat GEN.INDEF adobo
'I'm the one who ate adobo.'

b. aku a nag-kaʔɨn nya aduːbu
1s.NOM NOM.DEF AV.PRF-eat GEN.DEF adobo
'I'm the one who will eat the adobo.'

c. aku yu nag-kaʔɨn nyu aduːbu
1s.NOM NOM.SPEC AV.PRF-eat GEN.SPEC adobo
'I'm the one who ate the adobo.' (McFarland 1974:165)

On this understanding of "specific", it refers to a level of referentiality above and beyond uniqueness. In other languages which do not mark definiteness or specificity explicitly via case marking, the basic referentiality of an argument is largely predictable on the basis of grammatical function. There is some debate about whether these morphemes are inherently case markers or whether they have inherent referentiality related functions.[34] Yakan is interesting in this regard as it has an independent definite marker that is constrained but not fully determined by case. According to Behrens and Brainard (2002), the definite marking *-in* suffix is obligatory on absolutive arguments, optional on ergative arguments, and avoided on objects of AV/antipassive verbs in matrix clauses.

In Table 12, we see case markers for common nouns (all nouns but personal names) in six CSP languages and in Table 13 we see their counterparts for personal names. It is immediately clear that the Bilic language, Tboli and the Sama language, Yakan, diverge from the others in their reduced case system. All other languages make at least a three-way distinction between nominative case, genitive/ergative case, and an oblique case.

Table 12. Common noun case markers in six CSP languages

	Tagalog	**Aklanon**	**Subanon**	**Maranao**	**Tboli**	**Yakan**
NOM/ABS	aŋ	ro	su	so	∅	∅
GEN/ERG	naŋ	it (indef) ku (def)	nə nu (anaphoric)	o	∅	weʔ (erg)
OBL	sa	sa	sə (local) nə (non-local)	sa (indef) ko (def)	beʔ	si

[34] 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.

Table 13. Personal case markers in six CSP languages

	Tagalog	**Aklanon**	**Subanon**	**Maranao**	**Tboli**	**Yakan**
NOM/ABS	si	si	si	si	∅	si
GEN/ERG	ni	ni	ni	i	∅	
OBL	kay	kay	ni	ki	∅	

The Sama ergative markers are unrelated to the genitive/ergative markers of other Philippine languages. Rather, these appear to be cognate with Malay *oleh*, a preposition which introduces passive agents. Among the CSP languages, only the Sama group shows a unique ergative marker that does not also serve to introduce possessors.

On the southern periphery of the CSP zone, grammatical relations are indicated more through word order than case marking. The Cotabato Manobo example in (W) shows how actor voice objects and obliques may remain completely unmarked despite the existence of case markers in the language.

(W) h<um>ated a sagiŋ kaut ta
<AV>take 1s.NOM banana Kaut DET
'I will take some bananas to Kaut.' (Kerr 1988:13)

In Tboli, case is only distinguished on pronouns. As a predictable result, the order of arguments in multi-argument clauses such as (Q) is rigid.

(Q) ∅-oguh-en tum libun tum kun namak
CV-hand.to-3s.GEN that girl that 3s.OBL betel.nut.quid
'He hands his own quid of betel nut to the girl.' (Forsberg 1992:78)

While the rich case marking system of Bikol languages shows that subtle referentiality distinctions can be made in the markers themselves, the basic definiteness distinction typical to Philippine type voice systems remains even in languages that have lost their case markers. For instance, Walton (1986) describes a difference in definiteness as well as aspect between the undergoer voice and actor (antipassive) voice in Pangutaran Sama, as shown in (R) (repeated from H, above), despite the lack of any marking on the undergoer.

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

b. m-bonoʔ sultan bantaʔ na
AV-kill Sultan enemy 3s.GEN
'The king kills/fights some of his enemies.'
(Walton 1986:120)

We can make the following generalizations about case marking in CSP languages:

i. There is a common three-way case system involving NOMINATIVE/ABSOLUTIVE, GENITIVE/ERGATIVE and OBLIQUE cases.

ii. The OBLIQUE case is employed for a wide range of directional/locative functions, as well as for marking definite objects of actor voice clauses, when this is allowed.
iii. With the exception of the Sama languages, the case of non-actor voice agents is always the same as that of possessors, hence labelled GENITIVE/ERGATIVE.
iv. Common noun phrases and personal names have distinct but morphologically related case markers.
v. Case marking is typically obligatory on all arguments.
vi. Case marking persists in accordance with the following hierarchy: pronouns > personal names > common nouns, such that it is lost first on common nouns and last on pronouns.

3.5 Mode

Besides voice and aspect, there are several common verbal morphemes in CSP languages that are often treated under the header of "mode", a practice I continue here. These include the potentive (which subsumes both accidental and abilitative meanings), sociative and pluractional. Reflexives and reciprocals, as valency changing operations, are treated separately in §3.6.

3.5.1 Potentive *ka-

Nearly all CSP languages have a potentive paradigm, which is used to indicate both possible and unintentional action.[35] This paradigm, which is contrasted with the unmarked "dynamic" voice paradigm in Table 14 for Tagalog, has a very distinct history involving the PAn prefix *ka-, whose original function may have involved possession (Kaufman 2011).

Table 14. 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 (Xa) and a passive-like stative clause, as in (Xb),

[35] This paradigm is variously called abilitative, accidental, stative, among other terms. Himmelmann (2006) argues convincingly that *ma-* forms are involved in two distinct inflectional paradigms in Tagalog, which he terms the potentive and the stative.

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.

(X) a. ʔ<um>u:bos
<AV>finish
'to finish'

b. k<um>a-ʔu:bos → ma-ʔu:bos
<AV>STA-finish
'to get finished'

In all CSP languages that show a reflex of this *ma-, an agent can be introduced just as it is in a regular dynamic transitive clause yielding oppositions as in (Y).

(Y) na-ʔu:bos ni bo:boy aŋ pagka:ʔin
STA.BEG-finish GEN Boboy NOM food
'Boboy finished the food (accidentally)'

<in>u:bos-Ø ni Bo:boy aŋ pag-ka:ʔin
<BEG>-finish-PV GEN Boboy NOM food
'Boboy finished the food (purposefully)'

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), Bikol Naga, and Tagalog (both Central Philippine), shown in Table 15. Note that *ma-* is labelled as a patient potentive in Table 15 due to sharing a case frame with the patient voice in CSP languages, as seen above in (Y), but it is historically an actor voice form and is considered intransitive by Reid and Liao (2002:462).

Table 15. Potentive paradigms for three Philippine-type languages

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

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.[36]

[36] 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.

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 but one of causation.

The PAn *taR- prefix as reconstructed by Blust & Trussell (ongoing) overlaps semantically with the potentive paradigm in Saaroa, a Formosan language, and numerous languages south of the Philippines but is rarely found with this function in the Philippines, possibly having been replaced by the *ka- paradigm discussed here. It is however found prominently in the Sama languages. We can see in (Wa) how *maka-* functions as the actor voice potentive and, in (Wb), how the *ta-* functions as the undergoer voice potentive in Sama Bangingi.

(W)a. insaʔ aku maka-billi kuhapu; insaʔ niyaʔ
NEG 1s.NOM AV.ABL-buy grouper NEG EXT
'I was not able to buy grouper; there wasn't any.'

b saʔ ta-billi-ku taŋiliʔ itu
but UV-buy-1s Span.Mackerel this
'But I was able to buy this Spanish mackerel.'' (Gault 1999:15)

The Tboli sentences in (I) and (L) show the potentive *g(e)-* prefix, presumably a reflex of PMP *ka-, in both accidental and abilitative meanings. It is unclear how this prefix interacts with voice.

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

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

The potentive is CSP languages, as in many other Austronesian languages, does not simply provide a way of emphasizing the accidental or unintentional nature of an action. It is obligatory in such contexts and as a corollary, the unmarked (non-potentive) form unambiguously denotes intentional action with an animate agent.

Finally, it must be noted that PAn *ka-, which is the main ingredient of the potentive paradigm appears to define a class of predicates which could be called statives, as discussed in detail for Tagalog by Himmelman (2006) and reconstructed historically by Ross (2005). Stative predicates give rise to a different argument structure in which genitive agents are not licensed in the way typical of transitive predicates in Philippine languages. The distinction between potentives and statives is a complex and understudied area in Philippine languages.

3.5.2 Distributive *paŋ-/maŋ-

Many CSP 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 is 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 for forming the predicate ‘to fish’ in several CSP languages, including Tagalog. For other predicates, such as Tagalog *kuha* ‘take’, shown in (Y), it is optional and adds a meaning ranging from repeated action, action on plural generic objects and unwanted persistence.

(Y)	a.	k<um>uːha	b.	maŋ-[k]uːha
		<AV>take		AV.DIST-take
		‘to 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 (Z) and Sarangani Manobo (O), the distributive can co-occur with any voice in most CSP languages.

(Z)	iːlog na laːbis na p<in>aŋ-isdaʔ-an
	river LNK overly LNK <BEG>DIST-fish-LV
	‘an over-fished river’

(O)	i-m-pem-[b]egay dan se libro
	CV-PRF-DIST-give 3p.GEN NOM book
	‘They gave out books.’ (Dubois 1976:76)

South of the Philippines, the distributive takes on new functions, such as that of a dedicated anti-passive in certain South Sulawesi languages (Kaufman 2016), as well as the default marker of actor voice, as in Malayic languages. The combination of the distributive and the conveyance voice prefix, i.e. *Si-paŋ-, is used in some CSP languages as an unambiguous instrumental voice, which is reduced to a plain reflex of *paŋ in some Sama languages.

3.5.3 Sociative *paki-/maki-

A morphological category that is somewhat peculiar to Philippine languages is the so-called ‘sociative’, expressed with a reflex of PMP *paki- or its actor voice counterpart, *maki-. In most cases, this morpheme can be translated into English as ‘with others’, as in Tagalog (X), although this often does not capture the relation between the agent and the others.

(X)	a.	maki-hiŋiʔ	b.	maki-taːwa	c.	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 (Xb), than for others. The predicate *makitaːwa* is typically interpreted as laughing because other people are laughing whereas the predicate *makiupoʔ* 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 in CSP languages, e.g. Tagalog *maki-pag-ʔaːway* AV.SOC-TR-fight, Cebuano *makig-ʔaːway* AV.SOC-fight. The

difference between the sociative mode versus the unmarked mode in such cases is subtle but the sociative appears to foreground an aspect of social exchange, even with predicates like 'fight'.[37]

3.5.4 Plural agent marking

It appears possible to reconstruct a PMP marker *si- which necessitated a plural subject.[38] In Central Philippine languages, we find a reflex in such forms as Tagalog *mag-si-takbo* AV-PL-run, where it serves plainly 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). Unlike *paŋ-, this marker appears to be thoroughly agent oriented in Central Philippine languages but in at least some Sama languages, *si- can also indicate a plural patient rather than agent, as in Mapun (E).

(E) Subay bataŋ suwat ituʔu p<in>ag-si-tapit ka inaʔa
must trunk write this <PV>TR-PL-near OBL that
'This letter should be placed close together with that one.' (Collins et al 2001:36)

Conversely, reflexes of *paŋ-, which originally indicated a pluractional or plural object, occasionally develop an agent oriented plural meaning.

Although it is rare for CSP languages to show obligatory number agreement with any argument, plural marking can be indicated simultaneously by several morphemes for emphasis, as in Tagalog (V), where the matrix clause predicate takes both the *si-* prefix as well as the <*aŋ*> infix, both independently indicating agent plurality.[39] The subordinate verb again takes the plural marker *si-*, in addition to the pluractional marker *paŋ-*.

(V) n<aŋ>ag-si-handa=ŋ mag-si-pam[b]aril
AV.BEG<PL>-PL-prepare=LNK AV.BEG-PL-DIST/shoot
'they prepared to go shooting' (Venago 1929:62)

Similarly, in Agutaynen, we find 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.

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

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

[37] It appears that the imperative of the sociative, *paki-, has developed in another direction, now signalling a polite request in a number of Philippine languages. Liao (2011) argues that there need not be a derivational relationship between *paki- and *maki- although the pragmatic link between the sociative function and polite requests is unlikely to be accidental.

[38] Kitada (2019) reconstructs this as a sociative while Liao (2011) reconstructs this as a simultaneous aspect.

[39] While the <*aŋ*> plural infix is an archaism in modern Tagalog, we find that it continues to mark plurality for property denoting words in a number of CSP languages.

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, a <*g*> infix carries out the same function on dimension adjectives, e.g. *mu<g>bo?* <PL>short, *da<g>ko?* <PL>large. Plural marking of this type appears somewhat historically unstable, both in meaning and form, and therefore difficult to reconstruct. For instance, within the recent history of Tagalog, we find a plural infix <*aŋ*>, as in Maranao, which appeared in aspect inflected forms, e.g. *m<aŋ>ag-aral* <PL>TR-study, *m<aŋ>a-tuːlog* <PL>STA-sleep but this has all but completely disappeared in the modern language.

3.5.5 Multifunctional *paR-/maR-

Reflexes of *paR- (*maR-, in the actor voice) can be found in almost all CSP languages although the range of functions associated with these morphemes differ from language to language and have been the subject of much study beginning with Pittman (1966) onwards. 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 (2009, 2018) 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, all cross-linguistically common uses for middle voice. A typically mixed paradigm showing both the putative *R- function and *pa- function of *paR- is found in Palawano (Zorc 1971), shown in Table 16. Here, a reflex of *maR- is found in the progressive of both intransitive and transitive actor voice paradigms but in other aspects it signals transitivity. Progressive aspect is often associated with decreased transitivity (Hopper & Thompson 1980) and thus appears to derive from middle voice *R-. On the other hand, causative *pa- is clearly responsible for the increased transitivity of the forms in the right hand column.

Table 16. A fragment of the Palawano actor voice paradigm (Zorc 1971)

	Intransitive AV	**Transitive AV**
progressive	məgC_1ə-	məgC_1ə-
perfective	<umin>	nəg-
unbegun/habitual	<um>	məg-
participle	pəg-	pəg-

In many CSP languages south of Tagalog, the "plain" actor voice *<um> paradigm increasingly gives way to a *maR- paradigm, as discussed by Liao (2004:106) and Lobel (2004, 2013:46-47). As noted by Liao (2004:107-121) and Reid and Liao (2004:457), 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.

3.5.6 Reciprocals and reflexives

There are two recurring strategies for forming reciprocals in CSP languages. The first, shown in Tagalog (Sa), 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) and is thus likely reconstructable to PMP. The second, exemplified by Samar-Leyte (Sb), 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 (Sa), as it more often refers specifically to two agents sharing in an activity.

(S)a. nag-patay-an sila
AV-kill-RECP 3p.NOM
'They killed each other.'

b. nag-ka-duːrug hira
AV-CO-sleep 3p.NOM
'They slept together.' (Zorc 1977:144)

Both of these involve a reflex of the *maR- prefix and in some cases the prefix appears to express a reciprocal on its own, as in Tagalog *mag-kitaʔ* AV-see 'to meet', Yakan *mag-sasaʔ* INTR-fight 'fight each other'. There are other reciprocal markers whose etymologies are not so clear. For instance, Tboli marks reciprocals with an *s-* prefix, e.g. *tagak* 'to leave behind' *s-tagak* 'to leave each other'; *tobоŋ* 'to help' *s-tobоŋ* '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* (only *-ay* appears to be reported for the Bisayan languages), as seen in (D). Although both components of this circumfix occur in other derivations, they do not seem to be semantically related.

(D) m<i>g-b<in>ulig-aʔ
<PRF>AV.DUR-<REC$_1$>help-REC$_2$
'They helped each other.' (Post and Gardner 1992:xxiv)

Reflexives are also commonly expressed with a descendant of *maR- and stative reflexives are expressed with a reflex of PMP *maR-pa-ka- AV-CAU-STA- in certain CSP languages among other areas (Zeitoun & Huang 2000 and Blust 2003). Tagalog (E) exemplifies a remnant of this construction although it is not entirely productive as a reflexive.

(E)a. mag-pa-ka-bulag
AV-CAU-STAT-blind
'make oneself blind'

b. mag-pa-ka-matay
AV.TR-CAU-STAT-die
'kill oneself'

c. mag-pa-ka-taːʔo
AV-CAU-STAT-person
'be humane' ('make oneself a person')

3.5.7 Inchoative

The inchoative, termed by Zorc (1977:142) "essive", has barely been investigated from a comparative perspective. In most Central Philippine languages, it is signaled with prefixal morphology but in the Sama languages, the same meaning may be indicated by an independent

word, cognate to Malay *jadi.* Prefixal inchoative markers are shown in Aklanon (X) and Tagalog (Y).

(X) nagiŋ-rayna si neli
AV.PRF.INCH-queen P.NOM Neli
'Nellie became a queen.' (Zorc 1977:142)

(Y) magiŋ bato
AV.INCH rock
'to become a rock'

Note that Tagalog *magiŋ* is not generally considered a prefix in the modern language and is thus written as a separate word in (Y). This is due to a relatively recent degrammaticalization process which has led to the possibility of structures such as (E), where *magiŋ* appears stranded from what would be its complement, *ano*. This is not possible in most languages possessing an inchoative prefix.

(E) ano aŋ gusto mo=ŋ magiŋ?
what NOM want 2s.GEN=LNK AV.INCH
'What do you want to become?'

Other evidence, however, shows that the Tagalog inchoative is less than a full word. Unlike a full verb, it cannot host second position clitics, which must attach to the following lexical stem, as seen in (F).

(F) magiŋ (*ka=ŋ) bato (ka)!
AV.INCH 2S.NOM=LNK stone 2S.NOM
'Become a rock!'

Although less commonly used in modern Tagalog, the inchoative form also has non-actor voice forms, which were still current in the early 20th century. In these forms, we can find the patient and locative voice suffixes following the lexical stem, as seen in (G), not the inchoative morpheme itself.

(G) aŋ maynila aŋ p<in>agin-pariːʔ-an niya
NOM Manila NOM <BEG>INCH-priest-LV 3s.GEN
'It was in Manila where he was ordained a priest.' (Lendoyro 1909:256)

Some Tagalog varieties have carried the degrammaticalization of *magiŋ* to its natural conclusion and now allow the inchoative marker to host second position clitics, which separate it from the lexical stem, as in (D).[40]

(D) magi ka=ŋ tapat!
AV.INCH 2s.NOM=LNK honest
'Be(come) honest!'

[40] I thank Laurie Reid for first bringing this to my attention. Despite my initial disbelief at this construction it is clearly an attested pattern. While Reid believes it is a conservative pattern that predates *magiŋ* as a prefix, I believe it is an innovative degrammaticalization.

Not all CSP languages express the inchoative with a dedicated affix. Maranao, for instance, uses a periphrastic construction, as in (Ha), or a simple reflex of *maR-, as in (Hb), to express change of state.

(H)a. mim-bɤloy a ator
AV.DIST-change LNK stone
'changed into a rock'

b. m<iy>ag-ɤtor
AV<PRF>-rock
'became a rock'

3.5.8 Causative *pa-

The PAn causative prefix *pa- is perhaps the most stable affix in the entire PMP morphological inventory and is found in some form in all the CSP languages. The causative introduces a causer into the argument structure and can co-occur with any voice, mode and aspect. The mapping of roles to arguments in the causative may not be entirely obvious, however. Although there are many complications involving the case frames of the predicate and the definiteness of the arguments involved, Table 17 shows the canonical mapping of roles to arguments in a causative clause, a pattern which is remarkably stable among Philippine and Formosan languages.

Table 17. 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 (V), 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.

(V) nag-pa-su:lat ako naŋ li:ham sa estudya:nte
AV.BEG-CAU-write 1s.NOM GEN letter OBL student
'I had a student write a letter.'

In a patient voice causative clause, as in (B), 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.

(B) p<in>a-sulat-Ø ko naŋ li:ham aŋ estudyante
<BEG>CAU-write-PV 1s.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 (N) 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.

(N) i-p<in>a-suːlat ko sa estudyaːnte aŋ liːham
CV-<BEG>CAU-write 1s.GEN OBL student NOM letter
'I had the student write a letter.'

Tboli, despite its morphological innovations, has preserved a reflex of PMP *pa- as the *h(ə)-* prefix, e.g. *h-bɔŋ* CAU-big 'make big/plenty'; *h-inum* CAU-drink 'cause (someone) to give a drink to (someone/something)' (Forsberg 1992:89-91). The syntax of causative clauses in Tboli is less clear, as the causative prefix seems to not co-occur with the general undergoer prefix *n-*. It is, in fact, unclear if the Tboli causative allows for voice alternations, at all. The examples found in Forberg (1992), two of which are shown in (P) and (O), all express the causer as a genitive argument, the causee as an oblique (when present) and the theme as a nominative, as would be found with the conveyance voice in other CSP languages.

(P) hə-səgyok-u kə kasi yəm bəw
CAU-care.for-1s.GEN OBL Kasi that carabao
'I had Kasi take care of the carabao.' (Forsberg 1992:90)

(O) gɛhɛl h-buluŋ-əm nə yəhən-əm
quickly CAU-medicine-2s.GEN now spouse-2s.GEN
'Quickly have your spouse treated now with medicine.' (Forsberg 1992:91)

Yakan, a Sama language, appears to merge several non-actor voices in the causative morphologically but retains distinct case frames that would be associated with the patient voice and the conveyance voice in other languages. For example, (Wa) appears to select the causee as the nominative/absolutive argument and would thus correspond to a causative patient voice in other CSP languages; (Wb), on the other hand, employs the same verb form but now seems to select the theme as the nominative/absolutive while expressing the causee as an oblique. This would correspond to the conveyance voice causative in other CSP languages.

(W)a. pa-dekdak-ne anak-ne-in
CAU-wash-3s.GEN child-3s.GEN-DEF
'She had her child wash (clothes).' (BB:202)

b. pa-dekdak-ne semmek-in si anak-ne-in
CAU-wash-3s.GEN clothes-DEF OBL child-3s.GEN-DEF
'She had her child wash the clothes.' (BB:203)

Bangingi Sama, which maintains a distinct instrumental voice (with a reflex of PMP *paŋ-), distinguishes the two types of predication in (Wa) and (b) with verbal morphology. The unmarked transitive verb (equivalent to patient voice), promotes the causee to nominative, as in (Qa), while the instrumental voice promotes the theme to nominative, as in (Qb).

(Q)a. pa-inum-na aku bohe?

CAU-drink-3s.GEN 1s.NOM water
'He'll give me some water to drink.'

b. pam-[p]a-inum-na ma aku bohe? ilu
CV-CAU-drink-3s.GEN OBL 1s water that
'He'll give that water to me to drink.' (Galton 1999:21)

Note that all Sama languages also allow actor voice causatives, as well, as in Bangingi Sama (A).

(A) ag-pa-inum iya ma aku bohe?
AV-CAU-drink 3s.NOM OBL 1s water
'He'll give me some water to drink.' (Galton 1999:20)

4.0 Word order typology

All the CSP languages are robustly head initial, as can be seen in the basic ordering relations exemplified by Tagalog in (W).

(W)

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:?e
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**
gino?o=ŋ reyes
mister=LNK Reyes
'Mister Reyes'

g. **Complementizer > Clause**
aka:la ni dodoŋ na matalino siya
thought P.GEN Dodong COMP smart 3s.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-hurry=2s.NOM=already
'You should hurry up!'

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

k. **Negation > Verb**
hindi? 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), (b) and (d), allow for alternatives but with different

semantic or pragmatic implications. A third category, which includes (c) and (h), represent tendencies but co-exist with equally unmarked alternative orders. We examine these in the following subsections.

4.1 Word order within the noun phrase

The vast majority of CSP languages are both head and dependent marking, in the sense of Nichols (1986), and possess a set of case marking determiners as discussed earlier (§3.4). While the order of the case markers in relation to the noun phrase is strict, the order of certain modifiers within the noun phrase can be relatively flexible. The canonical order of elements in the Tagalog noun phrase is shown in (W). The elements in square brackets do not co-occur but rather represent two options for expressing possessors.

	CASE	PRE-POSS	NUM	ADJ	ADJ	N	POST-POSS
(W)	aŋ	[kanya=ŋ]	maŋa	ma-ga~ganda=ŋ	pula=ŋ	bulaklak	[niya]
	NOM	3s.OBL=LNK	PL	ADJ-PL~beauty=LNK	red=LNK	flower	3s.GEN

'his/her not beautiful red flowers'

The case marker is in absolute initial position, as is the rule in Philippine languages, and this is followed by the position of the preposed possessor. The more common position for possessors is after the possessum, as shown on the right edge of the sequence although in rare cases, e.g. Hanunoo (Epo 2014), the presposed position appears to have become the norm. The preposed position only hosts pronominals in modern Tagalog although in earlier Tagalog, we find full NP possessors in this position, too, although stylistically marked. When possessors are preposed, they are always in the oblique case and never in the "pure" (typically *n-* initial) genitive case in CSP languages. Furthermore, they are typically connected to the following material in the phrase by the linker, as shown for Central Tagbanwa in (Xa) (Scebold 2003:60), Tagalog in (Xb), and Bikol in (Xc) (although some Central Philippine languages omit the linker here, cf. Wolff 1967:71-72).

(X)a.	kanimi a bavoy	b. inyo=ŋ ba:boy	c. sa=indo=ŋ urig
	2p.OBL LNK pig	2p.OBL=LNK pig	OBL=2p.OBL=LNK pig
	'your (pl.) pig'	'your (pl.) pig'	'your (pl.) pig'

In some languages, preposed oblique possessors have been described as inherently focused, as in Matigsalug Manobo (A). A better description for Tagalog would be that they are *focusable*, as opposed to the unfocusable enclitic pronominals.

(A)a.	ka anak ku	b. ka keddì ne anak	
	NOM child 1s.GEN	NOM 1s.OBL LNK child	
	'my child'	'my child (not his)'	(Wang et al 2006:41)

Following this position we find the plural marker, also ubiquitous in CSP languages.[41] It is only the position of the case marker, at the left edge of the NP, and the plural marker, between the

[41] Zorc (1977:103) claims that the plural marker (or "diversity marker") *maŋa* is found in all the Bisayan languages. CSP Languages outside the Central Philippine group show different markers, e.g. Tboli *kem*, Yakan *me?*. In Sama

case marker and following lexical material, which are in a truly fixed position preceding the head noun. Following the plural marker, the canonical order of elements is adjective followed by noun, but this is variable in most CSP languages, as shown for Tagalog in (X).

(X)a.	ma-bilis	na	paːgoŋ		b.	paːgoŋ na	ma-bilis
	ADJ-speed	LNK	turtle			turtle LNK	ADJ-speed
	'fast turtle'					'fast turtle'	

When modifiers are postposed, as in (Xb), the positioning of the plural marker can occasionally precede the modifier, as in the Mansaka example in (B). In cases such as this, the adjective can also be considered the head of a phrase being modified by the plural. In no CSP language can the plural marker be stranded without any following material and is thus best regarded as a proclitic.

(B)	yaŋ	baboy	na	maŋa maitum
	NOM	pig	LNK	PL black
	'the black pigs' (Svelmoe & Svelmoe 1974: 52)			

In Tboli, where order appears to be more rigid, some adjectives must precede the noun, e.g. *tehe kimu* former property, *dumu lan* other path (Forsberg 1992:39) but most follow the noun, e.g. *lan mahil* path easy, *koyu lembaŋ* tree large. For at least some adjectives, the position with regard to the noun is variable. As discussed by Donohue (2007:359-363), a rigid Noun-Adjective order emerges south of the CSP area 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.[42] A relevant Simunul Island Sama example with canonical Noun Poss Adjective order can be seen in (X), and agrees with the areal trend.

(X)	lansa	ku	heya
	boat	1s.GEN	large
	'my large boat' (Akamine 2005:387)		

The plural marker can often co-occur with a following numeral, yielding an approximative interpretation, as in (W). The plural marker can also follow a modifying numeral, as in (Wb), but here no approximative meaning is obtained and the presence of the plural is completely optional.

(W)a.	aŋ	maŋa	lima=ŋ	bato	b.	aŋ	lima=ŋ	maŋa	bato
	NOM	PL	five=LNK	stone		NOM	five=LNK	PL	stone

languages, plural marking seems optional, unlike in other subgroups of the CSP zone. 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.

[42] Even languages of northern Sulawesi belonging to Blust's Philippine subgroup appear to show Donohue's 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.

‘the approximately five stones’ ‘the five stones’

Demonstratives were left out of the template in (W) above because they are somewhat more difficult to generalize over in the CSP languages. Case is typically marked syncretically on demonstratives, e.g. Tagalog *ito* ‘this (neutral)’, *nito* ‘this (GEN)’, *dito* ‘this/here (OBL)’. In Tagalog, a prenominal demonstrative takes the place of the case marker and is connected to the following material via the linker. It can also occur on the right edge of the noun phrase and here the argument is preceded by the expected case marker. Demonstratives can also sandwich the noun phrase for emphasis, as in (Qc).

(Q)a. ito=ŋ malaki=ŋ aːso
this=LNK big=LNK dog
‘this big dog’

b. aŋ malaki=ŋ aːso=ŋ ito
NOM big=LNK dog=LNK this
‘this big dog’

c. ito=ŋ malaki=ŋ aːso=ŋ ito
this=LNK big=LNK dog=LNK this
‘this big dog’

It is possible that prenominal demonstratives do not co-occur with case markers in languages like Tagalog because the case markers are also derived historically from demonstratives (Reid 2002, Himmelmann 2016:334). Blust (2015) argues that nominative case markers with **a*, which are widespread in the Philippines, replaced an earlier “pure” nominative marker *su, reflexes of which are still found in the Danao languages, Subanen, Pangasinan, and a handful of others. If this is correct, it is possible that languages which preserve a reflex of *su do allow it to co-occur with a following demonstrative marker. This appears true for Pangasinan, where the *s-* initial nominative co-occurs with demonstrative morphology (Benton 1971:51-52, 88-91), but it is unclear for the languages of Mindanao. In Northern Subanen (Daguman 2004:148), demonstratives are described as occurring only on the right edge of the noun phrase, but followed by relative clauses, as in (I).

(I) s<in>aak-an su d-libun kətu nə mig-bələdyaʔ *ice cream*
<RL>ask-LV NOM G-woman that LNK AV.RL-sell ice cream
‘...hey asked that lady who was selling ice-cream.’ (Daguman 2004:159)

The strict postnominal position of demonstratives is a southern feature, as noted by Donohue (2007). In Sama languages, both possessors and demonstratives are restricted to appearing after the head noun. Demonstratives must, in fact, appear *after* any modifying relative clauses in Yakan according to Brainard and Behrens (2002:29), a typologically unusual pattern but one which is common to Malay.

Pallesen (1985:180) observes that the strictly post-nominal order of demonstratives and possessors in Sama languages has entered Tausug, a Central Philippine language, via contact, while its closest relatives in the Eastern Mindanaoan subgroup show the same flexibility found in Tagalog.

Despite some descriptions of classifiers in Central Philippine languages (Lopez 1967, Gonzalez 1973), they are vanishingly rare and appear to have never been obligatory. On the other hand, some languages of the southern CSP zone appear to make more robust use of

classifiers. Daguman (2004:87) describes both sortal (e.g. *buuk* ‘non-flat’, *laad* ‘flat’, *tawan* ‘human’) and mensural (e.g. *dipa* ‘arm span’, *daŋaw* ‘hand span’) classifiers in Northern Subanen. These classifiers follow numeral modifiers and precede adjectives in the pre-head domain, as shown in (B).

(B) ...k=sala buuk g=əm-bagəl nə d=liun...
NM=one CL:non.flat NM=ADJ-big LNK NM=link
‘one big lion’ (Daguman 2004:158)

The position of negation in the noun phrase, not shown in the schema in (W), largely depends on what is being negated. Negation can precede the entire noun phrase, as in (I), but in this position a predication is being negated rather than a particular element within the noun phrase.

(I) hindiʔ aŋ kanya=ŋ asa:wa
NEG NOM 3s.GEN=LNK spouse
‘(It’s) not his/her spouse’

Within the noun phrase, negation can narrowly negate the prenominal possessor, as in (La), or precede the head noun, as in (Lb). While these constructions are pragmatically odd without a proper context, they are nonetheless grammatical.

(L)a. aŋ hindiʔ kanya=ŋ asa:wa
NOM NEG 3s.GEN=LNK spouse
‘the one who is not his/her spouse’

b. aŋ kanya=ŋ hindiʔ asa:wa
NOM 3s.GEN=LNK NEG spouse
‘his/her “not spouse”’

Negation can even follow the head noun when preceded by the linker, as shown in (U), and thus differs from the proclitic plural marker.

(U) polgas, aŋ aso=ŋ hindiʔ
Polgas NOM dog=LNK NEG
‘Polgas, the dog that is not (a dog)’ (Medina 1995)

The previous example calls our attention to the role of the linker, a thorny problem in the study of the Philippine syntax generally. The linker, which signals all types of modification, is common to the vast majority of Philippine languages but far rarer south of the CSP zone, even among Philippine-type languages and those of Blust’s (2019) Philippine subgroup. The presence of the linker correlates with freer word order within the noun phrase. Witness, for example, the flexible position of the relative clause in relation to the head in Central Philippine languages, as exemplified by Tagalog (D).

(D)a. aŋ bulaklak na b<in>ili-Ø=ko
NOM flower LNK <BEG>buy-PV=1s.GEN
‘the flower I bought’

b. aŋ b<in>ili-Ø=ko=ŋ bulaklak
NOM <BEG>buy-PV=1s.GEN=LNK flower
‘the flower I bought’

As the linker disappears towards the southern range of the CSP area, the order within the noun phrase becomes more rigid. The Bilic and Sama languages again pattern with their southern neighbors in lacking the linker and word order flexibility within the noun phrase (including the position of relative clauses in relation to their head noun).

Due to the word order flexibility and the optionality of the "head" noun in such structures, it is not even clear that the relative clause is subordinate to the head as is generally the case in Indo-European and other language families. Rather what we find is a symmetrical modification structure that largely ignores lexical categories. In the Central Philippine languages, a noun can modify a noun in precisely the same manner that an apparent verb or adjective can modify a noun, as shown in (N).

(N)a.	baːhay	na	bato	b.	ma-bigat	na	bato	c.	b<in>asag-Ø	na	bato
	house	LNK	stone		ADJ-weight	LNK	stone		<BEG>break-PV	LNK	stone
	'stone house'				'heavy stone'				'broken stone'		

As discussed in Kaufman (2018), there are no dedicated relative clause markers in conservative Philippine-type languages that distinguish a structure like (Nc), in which an event denoting predicate modifies a noun, from that in (Nb), where an adjective modifies a noun.[43]

4.2 Word order within the clause

As with all conservative MP languages, the CSP languages are almost without exception predicate initial across lexical category and clause type. Beyond this simple 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. Nearly all CSP languages show the basic order shown in (X) for undergoer voice (i.e. non-actor voice) clauses and actor voice clauses.

(X)	a.	<u>Undergoer voices</u>	b.	<u>Actor voice</u>
		V A_{GEN} P_{NOM}		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 be adjacent to the predicate head. In languages with impoverished case marking, such as those of the Bilic and Sama subgroups, this tendency becomes a rule. Starosta, Pawley and Reid (1982) attributed this tendency to the reanalysis of nominalizations as verbs; just as possessors are tightly bound to the preceding possessum in Austronesian languages, so are genitive agents tightly bound to the preceding non-actor voice verb. In the actor voice, the ordering relations appear to be less fixed although if there is an unmarked order, it is typically said to be the one in which nominative argument follows the patient. At least one Formosan language, Amis, shows a similarly strict word order in non-actor voice clauses, as in (X), with a corresponding word order flexibility in the actor voice, suggesting that the pattern predates PMP. Even the Sama languages, which diverge considerably from their northern neighbors in morphosyntax, still maintain the same basic word order in (X). Yakan (La) shows obligatory agent first order in a transitive

[43] Another approach is to treat the linker as a relative clause marker and all modification, including that between an adjective and a noun, as obligatorily mediated by relativization.

clause, while in the "intransitive" actor voice clause in (Lb), the object is verb-adjacent and followed by the absolutive argument. Both orders can be described as "absolutive last".

(L) kehet dende-hin kenna-hin
cut woman-DEF fish-DEF
'The woman cut up the fish.'

b. ŋ-[k]ehet kenna dende-hin
INTR-cut fish woman-DEF
'The woman cut up fish.'
(Brainard and Behrens 2002:160)

The preverbal domain is typically reserved for pragmatically marked arguments and adjuncts (Naylor 1975, Kroeger 1993, Kaufman 2005 inter alia). All languages discussed here allow for topicalization of the nominative/absolutive argument to a preverbal position (Reid & Liao 2004:447). Typically, the fronted topic is followed by a dedicated topic marker, as in (U), but in languages like Cebuano, there is topic fronting without a topic marker.

(U) aŋ uŋgoy ay <um>akyat sa puːnoʔ
NOM monkey TOP <AV>climb OBL tree
'The monkey climbed the tree.'

Oblique phrases, prepositional phrases and adjuncts can be topicalized in all the languages surveyed here. Genitive arguments and certain types of adjuncts cannot be topicalized so easily. The least extractable phrase is generally the actor voice object, which must occur post-verbally, as shown in (Y). This restriction extends to relativization and cleft-like constructions in addition to topicalization.

(Y)a. aŋ baːtaʔ ay k<um>aʔin (naŋ) maŋga
NOM child TOP <AV.BEG>eat GEN mango
'The child, ate the mango.'

b. *(naŋ) maŋga ay k<um>aʔin aŋ baːtaʔ
GEN mango TOP <AV.BEG>eat NOM child
(For, 'A mango, the child ate.')

There is a sprawling theoretical literature on this pattern, which cannot be reviewed here. The constraint holds in much the same way across CSP languages, including Sama and Bilic, despite their historical restructuring. The grammatical and ungrammatical clefts in Yakan (F) exemplify this.

(F)a. iyan buwaʔ-buwaʔ p<in>oloŋ nakanak-in
that toy TR-break child-DEF
'This is the toy that the child broke.'

b. *iyan nakanak-in p<in>oloŋ(-ne) buwaʔ-buwaʔ
that child-DEF TR-break-3s.ERG toy
(For, 'This is the child that broke the toy.') (Brainard and Behrens 2002:166)

For all the languages discussed here, including the Bilic and Sama languages, the clause initial predicate position typically hosts new information, and the nominative/absolutive argument has typically been previously introduced. Unlike Indo-European languages, in which verbs are unmarked predicators and nouns are unmarked arguments, the predicate and the nominative phrase are generally unselective with regard to lexical category. Whereas the English translation of (Ca) is a monoclausal sentence and that of (Cb) a biclausal sentence (with a copula serving as the main clause predicate), there is little evidence that such a difference exists for most CSP languages. The Tagalog sentences in (C) thus differ only in whether an event denoting phrase fills the predicate position, as in (Ca), or the position of the nominative argument, as in (Cb).

(C)a. [d<um>atiŋ kagabi]$_{Pred}$ [si boboy]$_{Nom}$
<AV.BEG>arrive last.night NOM Boboy
'Boboy arrived last night.'

b. [si boboy]$_{Pred}$ [aŋ d<um>atiŋ kagabi]$_{Nom}$
NOM Boboy NOM <AV.BEG>arrive last.night
'It was Boboy who arrived last night.'

Kaufman (2018) shows how a more English-like structure emerges in Indonesian languages, many of which develop a dedicated relativizer after the loss of case markers. This holds true, too, for the Sama languages, as seen in Yakan (D).

(D)a. tekka si dende dibuhiʔ
arrive PM Dende last.night
'Dende arrived last night.'

b. si dende iye ma-tekka dibuhiʔ-in
PM Dende 3s NMLZ-arrive last.night-DEF
'Dende is the one who arrived last night.' (BB:180)

In (Da), we find that a Yakan clause with a verbal predicate does not look much different from its Central Philippine counterpart. In (Db), however, instead of the expected **Si dende tekka dibuhiʔin*, we find a third singular dummy pronoun *iye* as well as a special form of the verb, with a *ma-* prefix, which Brainard and Behrens (2002) describe as a nominalizing prefix appearing on intransitive/antipassive verbs in relative clauses (see Walton 1997 for similar structures in Pangutaran Sama). There is thus a clear argument for treating (Db) as a more complex sentence than (Da), similar to its English translation. Tboli, which has also lost its case marking, has apparently not developed a structure similar to (Db). The Tboli pair in (E) suggest, at least on the surface, a symmetrical structure with regard to how phrasal categories are mapped to the grammatical roles of predicate and subject.

(E)a. Ø-blay le do sotu kudaʔ
CV-give 3p.GEN 1s.OBL one horse
'They gave me one horse.'

b. sotu kudaʔ Ø-blay le do
one horse CV-give 3p.GEN 1s.OBL
'It was one horse they gave me.'
(Forsberg 1992:55)

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

(A) sa maynila=na=kami nag-aːˑaːral
OBL Manila=already=1p.ex.NOM AV.BEG-IMPRF~study
'It's in Manila that we study.'

This construction is different from that of (E) and (Fa) in that the notional predicate is not embedded in a nominative phrase. However, the fronted oblique phrase attracts second-position clitics (see §4.4, below) and receives a cleft-like "exhaustive list" interpretation, i.e. 'It's in Manila (and nowhere else) that we study' for (A).[44] This construction is generally uncommon, if attested at all, in languages of the northern Philippines and most likely represents an innovation that took place in some subset of the CSP languages. Note that in several Bisayan languages, focus fronting of an oblique phrase in this manner requires using the dependent paradigm of the verb.

4.3 Types of negation

CSP languages are relatively rich in negators; distinct functional negators exist 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 can be found showing three and four-way distinctions. The negation inventories of six CSP languages are shown in Table 18.

Table 18. Negation in CSP languages

	Tagalog	**Aklanon**	**Subanon**	**Maranao**	**Tboli**	**Yakan**
PERFECTIVE EVENT	hindiʔ	ʔuwaʔ	ʔəndaʔ	diʔ	laʔ	gaʔ
PROSPECTIVE EVENT	hindiʔ	ʔindiʔ	ʔəndiʔ	diʔ	laʔ	gaʔ
PROHIBITIVE	huwag	ʔayaw	ʔəndiʔ	diʔ	béʔ	daʔa
IDENTIFICATION	hindiʔ	bukon	gənnaʔ	kenaʔ	sundu	dumaʔin
EXISTENTIAL	walaʔ	ʔuwaʔ	ʔəndaidun	daraʔ	(laʔ wən)	(gaʔ niyaʔ)

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 & Trussel (ongoing) reconstruct PWMP *beken 'negator of nominals, other, different'. Blust & Trussel (ongoing) also reconstruct

[44] Kroeger (1993) and Kaufman (2005) discuss the relevant oblique fronting construction in Tagalog at length. As far as I can tell, the same facts described there hold for the Central Philippine languages more generally.

PWMP *laqin ‘different’ and this word, too, comes to function as a general negative marker in Sorsogon.[45]

The distinction between perfective and prospective negation is uncommon, occurring mostly in Bisayan languages that employ the negative existential in perfective event-denoting contexts.

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), but most Philippine languages employ distinct unanalyzable roots for the existential and negative existential. As seen in Table 18, Yakan and Tboli, both on the southern periphery of the Philippines, employ an analytic combination, as commonly found further south. There is a degree of fluidity between these functions, as shown by McFarland (1974:254-6). Nonetheless, there are several generalizations that can be made:

i. If a language has distinct negation for perfective events, it will be the same as the negative existential. (Subanen *ʔəndaidun*, above, exceptionally adds the formant *idun* in the negative existential.)
ii. If a language does not have a distinct prohibitive, this function will be carried out by the same form employed in the prospective.
iii. If a language does not have a distinct identification/non-verbal negator, this function will be carried out by the eventive/verbal negator.
iv. If a language does not have a distinct negative existential marker, this function will be carried out by the eventive negation in combination with the (positive) existential.

In a large number of CSP languages, certain negative contexts require the dependent verbal paradigm, as noted by Wolff (1973). Zorc (1977) shows that the negation of perfective verbs in Bisayan generally employs the negative existential with the dependent verbal inflection, as discussed earlier in §3.3.

4.4 Pronominal morphosyntax and second position clitics

Bound pronouns in Philippine languages are almost always second-position clitics, which cannot appear initial in their domain. Free pronouns are positioned more like full noun phrases but may also be banned in the regular, post-predicate argument position. The complementarity between bound pronouns and full noun phrases, typical for all CSP languages, can be seen in Maranao (X). 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 (X). This position is not available for full noun phrases, as shown in (Xb).

(X) a. diʔi[=ako] ma-matiya[*=ako] sa kitab
PROG=1S.NOM AV-read=1S.NOM OBL book
‘I’m reading a book.’

b. diʔi [*so wataʔ] ma-matiya [so wataʔ] sa kitab.

[45] There is a yet unexplained polarity flip that occurred historically with PMP *wada ‘be, exist’, as noted by Dempwolff (1934-38), Blust & Trussel (ongoing) and Wolff (2010). In a number of non-contiguous languages, including ones in the CSP zone, the existential became a negative existential (e.g. Tag. walaʔ NEG.EXT).

PROG NOM child AV-read NOM child OBL book
'The child is reading a book.' (Kaufman 2007:136)

In languages like Tagalog, free pronouns are only used in predicate position (in cleft-like constructions), as independent fragments or as fronted topics. In other Central Philippine languages, such as Cebuano, long forms of the genitive and nominative pronouns show more syntactic freedom than in Tagalog (Wolff 1966). Only long forms can stand independently but both short and long forms can follow the predicate when representing an argument or possessor. In several CSP languages, there is a constraint against using the short forms of both the genitive and nominative pronouns in a single clitic cluster, which can be resolved in several ways. Where we would expect such a combination in Agutaynen, the argument lower on the person hierarchy is expressed as an oblique pronoun, as in (Xa). When one clitic can attach to a higher host, as in (Xb), then both arguments can be expressed as clitics. This is seen in (47), as well, where the second person pronoun is expressed as a free oblique pronouns in (47a) but as a genitive clitic in (47b).

(X)a. I-tabid=ami nandia
IRR:PV-accompany=1P.NOM 3S.OBL
'S/he will include us'

b. Indi=ami i-tabid=na
NEG=1P.NOM IRR:PV-accompany=3S.GEN
'S/he will not include us.' (Quakenbush 2005)

(47) a. Indi=o nio i-tabid
NEG=1S.NOM 2S.OBL IRR:PV-accompany
'Don't include me!'

b. Indi=o i-tabid=mo
NEG=1S.NOM IRR:PV-accompany=2S.GEN
'Don't include me!' (Ruch & Quakenbush 2006:9)

In Obo Manobo (Brainard & Vander Molen 2005), when both arguments of a transitive clause are first or second person, there is optionality as to which pronoun cliticizes to the preceding host. The other pronoun is expressed in its long, non-clitic form, as seen in (Y). Using the short forms of both pronouns (in either order) is strictly ungrammatical.

(Y)a. Od suntuk-on=du siyak
IRR hit-PV=2S.GEN 1S.NOM
'You hit me.'

b. Od suntuk-on=a nikkow
IRR hit-PV=1S.NOM 2S.GEN
'You hit me.'

In Hanunoo, there seems to be variability for at least some pronoun combinations without a visible difference in clisis, as we saw in Obo Manobo. In (V), from an orally narrated story, we find both orders of the first and second singular pronouns within the same sentence.

(V) ʔi-sukad **niku** **kawu** sa kaldiru, gatuŋ-an **kawu** **niku**
CV-put 1S.GEN 2S.NOM OBL pot boil-LV 2S.NOM 1S.GEN
'I'll put you inside the cooking pot and boil you!' (Epo 2014:157)

There are constraints from at least three domains that determine the ordering of pronominals within the clitic cluster (Schachter 1973, Liao 2004, Billings and Kaufman 2004, Kaufman 2010). In the domain of prosody, short forms precede longer forms; in the domain of case, genitive forms precede nominative forms; and in the domain of person, first precedes second which precedes third. Different constraints are active in different languages, but if a particular domain is active, it will always follow the above stated scales. In Tagalog, prosodic factors dominate so that monosyllabic clitics always precede disyllabic ones. If the syllable count is equal than case "breaks the tie" by having genitive pronouns precede nominative ones. In the languages of Mindanao, person constraints often play the dominant role. Maranao orders clitics strictly according to the person hierarchy with case again breaking the tie, all else being equal (Kaufman 2007).

Clitic doubling, which is found occasionally in the northern Philippines and is most developed in Kapampangan, is rare in the CSP zone. Tboli, however, does show clitic doubling with certain preverbal elements, as seen in (I), where the second position clitic *le* doubles the nominative argument *kem dumu*.

(I) deŋ=le ma koyu kem dumu
already=3p.NOM AV.fetch wood PL companion
'The others already fetched some wood.' (Forsberg 1992:63)

4.5 Multiple predicate constructions

Below we review finite complement clauses and several types of non-finite complements, including control clauses and a rarer construction that requires the actor voice. Raising, another type of biclausal construction in which an argument from a subordinate clause appears in the matrix clause, has been claimed to exist in at least some CSP languages (Kroeger 1993, Wegmüller 1998 inter alia) but for reasons of space raising will not be reviewed here.

4.5.1 Finite complement clauses

All CSP languages allow for finite clause complements, as in (X), where the embedded clause has all the hallmarks of a main clause predicate and can serve as such.

(X) s<in>abi-Ø ko sa iyo na gaː~gaw-in niya bukas.
<BEG>say-PV 1s.GEN OBL 2s LNK IMPRF~do-PV 3s.GEN tomorrow
'I told you that s/he will do (it) tomorrow.'

Although the relation between a complement clause and a matrix clause is not considered one of modification, complement clauses are generally introduced by the linker, as in (X). Aside from this marker, there is no overt sign of subordination. While it is relatively trivial to subvert the more common Adjective-Noun order in Central Philippine languages (§4.1), reversing the order of clauses in a sentence like (X) is far more marked.

Arguments interpreted as part of the subordinate clause can also be questioned in the cleft like structure shown in (Y).

(Y) ano aŋ s<in>abi-Ø mo=ŋ gaː~gaw-in niya?

what NOM <BEG>say-PV 2S.GEN=LNK IMPRF~do-PV 3S.GEN
'What did you say he would do?'

In such cases, it seems that both the subordinate and matrix clause verbs must employ the voice that would select the interrogative phrase as the nominative argument. In the context of (Y), this would be the patient voice for both predicates.[46]

4.5.2 Interrogative complements

Interrogative complements, also referred to as "embedded questions", are used in subordinate clauses as complements to matrix predicates of cognition as well as subjunctive type complements. In the majority of CSP languages, these complements look exactly like questions except that the interrogative phrase is introduced by a conditional marker, as seen in Central Tagbanwa (V), Matigsalug Manobo (B), Tagalog (Z) and Hanunoo (D). Note that Scebold glosses Tagbanwa *iŋ* as HYPOTHETICAL but, as in Tagalog, one of its primary functions is to introduce conditional clauses.

(V) pog-tuʔma iŋ kali ka nag-giʔit.
INC.AV-ask HYP where 2S.NOM PRF.AV-depart
'He is asking where you came from.' (Scebold 2003:73)

(B) Su mig-inse sikandan ke hendei key eg-pa-bulus
so AV.PRF-ask 3P.NOM if where 1P.EX.NOM AV.PROG-CAU-continue
'So they asked where we were going.' (Wang et al 2006:112)

(Z) nag-tanoŋ siya kuŋ saʔan ka p<um>unta
AV.BEG-ask 3S.NOM if where 2S.NOM <AV.BEG>go
'He is asking where you went.'

(D) sabi-hun nimu sa kaŋku nu hayga
tell-PV 2S.GEN OBL 1S.GEN COND why
'Tell me why (it's) that way.' (Epo 2014:22)

4.5.3 Nonfinite complement clauses

Clausal complementation with verbs of wanting, trying, and certain non-verbal predicates are typically non-finite, meaning that they appear in a neutral form that does not indicate aspect, as shown in Tagalog (Y), Cebuano (Z) and Agutaynen (T).

(Y) naʔis ko=ŋ mag-ʔaːral
want 1s.GEN=LNK AV-study
'I want to study.'

[46] It has been argued by Rackowski and Richards (2005) that the matrix verb is actually selecting the subordinate clause as a whole in examples like (Y), thus making it transparent for extraction. More work is required to tease apart these analyses.

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

(T) mambeŋ aŋ mag-pa-layog ta boradol
fun LNK AV-CAU-fly OBL kite
'It's fun to fly a kite.' (Quakenbush et al 2010:13)

Note that voice marking is still present in most non-finite subordinate clauses. Other morphosyntactic categories discussed above, including the abilitative, causative, reflexive, etc. can also appear in such contexts. In a small minority of CSP languages, aspect in the subordinate clause appears to agree with the matrix predicate in what are typically non-finite contexts for other languages. Agutaynen appears to show such agreement, as shown in (C).

(C)a. nam-[p]ag-t<ar>abaŋ-an tanira=ŋ naŋ-ayeg
AV.PRFPL-TR-<PL>help-LV 3P.NOM=LNK AV.PRF.DIST-harvest
'They helped one another to harvest.'

b. mam-[p]ag-t<ar>abaŋ-an tanira=ŋ maŋ-ayeg
PRF.PL-TR-<PL>help-LV 3P.NOM=LNK AVDIST-harvest
'They will help one another to harvest.' (Quakenbush et al 2010:20)

Kroeger (1993) discusses restructuring/clause union in Tagalog, which allows second position clitics from the lower clause to appear in the higher clause, in addition to other material. Clitic positioning in a regular biclausal structure is shown in (Va) and in its restructured counterpart in (Vb).[47]

(V)a. hindiʔ kaːya ni Pedro=ŋ bigy-an=siya naŋ peːra
NEG able GEN Pedro=LNK give-LV=3S.NOM GEN money
'Pedro cannot give her money.'

b. hindiʔ=siya kaːya ni Pedro=ŋ bigy-an naŋ peːra
NEG=3S.NOM able GEN Pedro=LNK give-LV GEN money
'Pedro cannot give her money.' (Kroeger 1993)

A less common type of complementation pattern attested in Central Philippine languages involves treating the subordinate predicate as a case marked complement, as shown in Tagalog (V) and Cebuano (S).

(V) b<in>ilis-an ko aŋ pag-kaʔin
<BEG>fast-LV 1S.GEN NOM GER-eat
'I speeded up my eating.'

(S) nag-siːge ug sunod sa iya=ŋ bukog

[47] See Kroeger (1993) for a discussion of raising and other types of biclausal constructions, which cannot be covered here.

AV.BEG-continue OBJ follow OBL 3S.GEN=LNK bone
'He continues following his bones.' (riddle)

4.5.4 Control patterns

Control, referred to as "equi NP deletion" in the early generative literature, refers to coreference between an argument in a matrix clause and a missing argument in a (typically non-finite) subordinate clause. In Tagalog, for instance, an agent in a subordinate clause that co-refers with a matrix argument must be null, as shown in (E).

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

Unlike in English, this has nothing to do with a non-finite predicate being unable to license a subject in the subordinate clause. When the subordinate agent does not co-refer with a matrix clause argument, it must be overt, as shown in (F), but the subordinate verb remains in the infinitive form.

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

Not just any argument in the subordinate clause can be controlled by a matrix clause argument. As shown by (G), a matrix argument cannot control the missing undergoer of a volitional transitive clause. The sentence is grammatical, but not under the interpretation where the missing nominative argument of 'call' corefers with the matrix subject.

(G) gusto ko=ŋ tawa:g-an mo
want 1S.GEN=LNK call-LV 2S.GEN
'I want you to call (someone).' Not: 'I want you to call me.'

But as has been noted (Kroeger 1993, Schachter 1976, 1994), the volitionality of the subordinate predicate determines which argument can be controlled, as seen in the minimal pair in (V).

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

b. gusto ko=ŋ ma-tawa:g-an
want 1S.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 the vast majority of CSP languages. The Sama languages, however, appear to differ in privileging the absolutive argument without regard to thematic role or volitionality. Trick (2006) exemplifies this for Southern Sama with the data in (F) and (D). Not only does a matrix argument control the undergoer of a volitional clause in (F), it appears that an ergative argument cannot be controlled by an argument in a higher clause, as is generally possible in Central Philippine languages and exemplified with Tagalog in (E), above.

(F) kabilahian-ku ni-liŋan-an leh si ben
want-1S.ERG PV-call-APPL ERG PM Ben
'I want Ben to call [me].'

(D) kabilahian si ben ni-liŋan-an aku
want PM Ben PV-call-APPL 1S.ABS
OK: 'Ben wants that I will be called'
Not: 'Ben wants to call me.'

Trick (2006) thus analyzes the Sama control pattern as following an ergative pattern in that only the absolutive argument can be controlled, whereas the Central Philippine pattern has been analyzed as being sensitive to a number of factors including thematic role and grammatical function (Schachter 1976, 1994).

4.5.5 The actor voice restriction

A more unusual phenomenon which has not received any notice in the literature is found in the Danao languages. For fully biclausal sentences, Maranao and Maguindanao show structures similar to Tagalog and other Central Philippine languages, as seen in Maguindanao (D), where the embedded verb is an infinitive in the locative voice.

(D) Kalinian=neŋka tawag-an=ko seka?
want=2s.GEN call-LV=1s.GEN 2s.NOM
'Do you want me to call you?'

However, in the structures analyzed as clause reduction above, the Danao languages require that the subordinate verb appear in the actor voice, as seen in (Z). Here, the undergoer of 'call' appears to obtain case from the matrix verb and is positioned in the matrix clause.

(Z) t<in>ekaw-an ko seka tawag!
<PRF>try-LV 1s.GEN 2s.NOM <AV>call
'I tried to call you!'

The Maranao examples in (S) show that the use of the expected voice in this construction, as in (Sb), is degraded. Only the structure in (Sb) would be grammatical in the Central Philippine languages thus far described.

(S)a. kəbəyaan=ko seka m-ugup b ?kəbəyaan=ko seka ugup-an
want=1s.GEN 2s.NOM AV-help want=1s.GEN 2s.NOM help-LV
'I want to help you.' 'I want to help you.'

This corresponds to the so called "actor voice constraint" discussed by Aldridge (2004) for Seediq and Chang (2010) for both Kavalan and Seediq. Attestations of this construction in

Philippine languages are vanishingly rare although an example from Ivatan cited in passing by Reid and Liao (2004) and shown in (55) appears to evince the same pattern.[48]

(55) kakey=da a s<om>idoŋ sira
want=GEN.3P LIG help NOM.3P
'They wanted to help them.' (Larson 1986:8)

Further south, Kroeger (2008, 2014) has described the same phenomenon in Kimaragang, a Dusunic language of Sabah. As the actor voice constraint appears in multiple branches of Formosan languages and is found in Kimaragang as well as the Danao languages, it most likely represents the retention of a pattern that predates PMP but which was lost in the vast majority of MP languages. It is also possible that the pattern may still turn up in some form in Central Philippine languages. For instance, in Cebuano (N), we find what appears to be restructuring (with dependent voice morphology due to the prohibitive) and the use of the actor voice in the subordinate clause where we would expect a patient voice form. The nominative case on the subordinate undergoer argument is clearly assigned by the matrix predicate, *halwu:ta*. This is a rich area for further exploration.

(N) ayaw halwu:t-a pag-lu:tuʔ aŋ tirati:ra
PROH hard-PV.DEP AV.DEP-cook NOM taffy
'Don't cook the taffy so hard.' (Wolff 1972:294)

4.5.6 Adjunct clauses

Nominalizations are used to form temporal adjuncts in a wide range of Austronesian languages and this is particularly true among the CSP languages (Kaufman 2011). We can compare the finite Tagalog clause in (Za) with the temporal adjunct in (Zb), which employs a gerundive. The argument which would normally be assigned nominative case, as in (Za), here takes genitive case, and is typically followed by a finite clause with a nominative case argument.

(Z)a. <um>alis si maria
<AV.BEG>leave NOM Maria
'Maria left.'

b. pag-ʔalis ni maria...
GER-leave GEN Maria
'When Maria leaves/left...'

A similar structure is seen in (W) for Sarangani Manobo, which also shows the use of oblique case on objects of such gerundive nominalizations.

(W) peg-dineg te amay din kenyan
GER-hear GEN father 3S.GEN that.OBL
'When his father heard that.' (DuBois 1976:94)

5.0 Conclusion

[48] This example from Larson's (1986) text was not accepted by several speakers of Batanic varieties and its status thus remains to be investigated. Thanks to Kristine Gallego for confirming this with her collaborators in the Batanes.

This chapter has attempted to give a broad overview of the phonology, morphology and syntax of the CSP languages while focusing on several phenomena of interest that are particularly characteristic of the region. I have also attempted to highlight areas in further need of documentation. In the phonology, gradient phonotactic generalizations have largely gone unexplored beyond Tagalog and the study of word prosody and intonation is also a rich area for further research. The CSP languages have played a large role in our understanding of PMP morphosyntax but we still have an incomplete understanding of how the dependent paradigm was deployed as well as various types of subordination. The actor voice constraint has been presented here for the first time as a Philippine phenomenon, in addition to its presence in Sabah and Formosan languages.

Finally, a note on the general typology of the region. Himmelmann (2005) defines Philippine-type languages as having symmetric voice in addition to the following three characteristics:

(a) at least two formally and semantically different undergoer voices
(b) at least one non-local phrase marking clitic for nominal expressions
(c) pronominal second position clitics

These features, all of which are understood to be retentions from PMP, begin to erode in the Bilic and Sama languages, thus opening a typological rift between Bilic and Sama on the one hand and all other Philippine languages on the other hand. Yet it is clear that Bilic and Sama represent independent developments. If it is not a coincidence that they share certain simplifications in the voice and case system, it may be due to historical contact with non-Austronesian languages rather than a purely autochthonous development. In any case, these developments are seen to be extremely rare in the rest of the Philippines despite the large number of languages. Kalamianic, on the other hand, despite representing a higher level branch of the putative Philippine subgroup in comparison to the Greater Central Philippine languages, does not differ greatly from a typological perspective when compared with its neighbors. This could be due to centuries of convergence, as these languages show clear hallmarks of long term contact with Central Philippine languages.

Despite progress, there is still much work to be done in the description of CSP languages outside the Central Philippine group. While contact relations between Sama and Bisayan have been analyzed in great detail by Pallesen (1985), no such effort has been made for other areas within CSP. The Bilic languages, in addition to the Palawan and Mindoro areas, are in special need of further work with an eye towards contact relations. Blust (1992), looking at the Tiruray lexicon, has already shown that the emerging picture is complex and multilayered.

Bibliography

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

Akamine, Jun. 2002. The Sinama derived transitive construction. In Wouk and Ross: 355- 366.

Akamine, Jun. 2005. Sama (Bajau). In Karl Alexander Adelaar & Nikolaus P. Himmelmann (eds.), *The Austronesian languages of Asia and Madagascar*, 377-396. London: Routledge.

Aldridge, Edith. 2004. Ergativity and Word Order in Austronesian Languages. Ph.D. dissertation, Cornell University.
Benton, R.A., 1971. Phonotactics of Pangasinan. Department of Linguistics, University of Hawaii.
Billings, Loren and Daniel Kaufman. 2014. Towards a typology of Austronesian pronominal clisis. *ZAS Papers in Linguistics*. pp. 15-29.
Blake, Frank. 1917. Reduplication in Tagalog. *The American Journal of Philology*. 38:4, p. 425-431.
Bloomfield, L. 1917. 2 parts. *Tagalog texts with grammatical analysis*. University of Illinois Studies in Language and Literature. Urbana, Illinois: University of Illinois.
Blust, Robert. 1971. A Tagalog consonant cluster conspiracy. PJL 2: 85-91.
Blust, Robert. 1991. The Greater Central Philippines hypothesis. OL 30: 73-129.
Blust, Robert. 1992. On speech strata in Tiruray. In M.D. Ross, ed., Papers in Austronesian Linguistics, No. 1: 1-52. Canberra: Pacific Linguistics.
Blust, Robert. 2003. Three notes on Early Austronesian morphology. OL 42: 438-478.
Blust, Robert. 2004. Austronesian nasal substitution: a survey. *Oceanic Linguistics* (43):73--148.
Blust, Robert. 2005. The linguistic macrohistory of the Philippines: some speculations. In Liao and Rubino: 31-68.
Blust, Robert. 2005b. A note on the history of genitive marking in Austronesian languages. *Oceanic Linguistics* 44:215-22.
Blust, Robert. 2007. The linguistic position of Sama-Bajaw. SPLC 15: 73-114.
Blust, Robert. 2015. The case-markers of Proto-Austronesian. *Oceanic Linguistics*. 54(2):436-491.
Brainard, S. and D. Behrens. 2002. A grammar of Yakan. Manila: Linguistic Society of the
Burton, Scott L., (ed.). 2018. Tagakaulo-English Dictionary. SIL Philippines.
Casanova, Arthur P. and Ligaya Tiamson Rubin. 2001. *Retorikang pangkolehiya*. Manila: Rex Bookstores.
Cena, R.M. 1971. Case opposition in Tagalog. *Ohio State Working Papers in Linguistics* 10 133-148.
Chang, Henry Yungli. 2010. "On the syntax of Formosan adverbial verb constructions." *Austronesian and theoretical linguistics*, pp. 183–212. Ed. by Raphael Mercado, Eric Potsdam, and Lisa deMena Travis. Amsterdam: John Benjamins.
Charles, M. 1974. Problems in the reconstruction of Protophilippine phonology and the subgrouping of the Philippine languages. OL 13: 457-509.
Collins, James N. 2018. Definiteness determined by syntax: A case study in Tagalog. *Natural Language and Linguistic Theory*.
Conklin, Harold. 1956. Tagalog speech disguise. *Language* 32: 136-139.
Daguman, Josephina Sanicas. 2004. *A grammar of Northern Subanen*. Ph.D. diss., 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, V. P. 1988. Ergative analysis for Philippine languages: An analysis. In R. McGinn (Ed.), Studies in Austronesian Linguistics, 323{345. Athens, OH: Ohio University Center for Southeast Asia Studies.
De Guzman, Videa P. 1991. Inflectional morphology in the lexicon: Evidence from Tagalog. *Oceanic Linguistics* 30:33-48.

Donohue, Mark. 2004. Typology and linguistic areas. *Oceanic Linguistics* 43: 221–239.
Donohue, Mark. 2007. Word order in Austronesian from north to south and west to east. *Linguistic Typology* 11:349–391.
Dubois, C.D. 1976. Sarangani Manobo: an introductory guide. Manila: LSP.
Epo, Yrrah Jane S. 2014. Discourse analysis of suyot: a Hanunuo-Mangyan folk narrative. MA thesis. Payap University.
Fincke, Steven C. 2002. Sto. Domingo Bikol verbal portmanteau morphemes: Transitivity in conversation. PhD diss. UC Santa Barbara.
Foley, W.A. 1976. Comparative syntax in Austronesian. Unpublished Ph.D. Dissertation. Berkeley, California: Department of Linguistics, University of California.
Foley, William. 2012. A comparative look at nominalisations in Austronesian. Keynote address presented at 12-ICAL, Denpasar, Bali, July 2, 2012. Ms., 50pp.
Gerdts, D. B. 1988. Antipassives and causatives in Ilokano: Evidence for an ergative analysis. In R. McGinn (Ed.), Studies in Austronesian Linguistics, 295{321. Athens, OH: Ohio University Center for Southeast Asia Studies.
Vivian M. Forsberg. 1992. A Pedagogical Grammar of Tboli. *Studies in Philippine Lingiustics* 9. 1-110.
Geraghty, P.A., L. Carrington, and S.A. Wurm, eds. 1986. 2 vols. FOCAL I, II: Papers from the Fourth International Conference on Austronesian Linguistics. Canberra: Pacific Linguistics.
Gonzalez, A.B. ed., 1973. Parangal kay Cecilio Lopez: Essays in honor of Cecilio Lopez on his seventy-fifth birthday. LSP Special Monograph Issue No. 4. Quezon City: Linguistic Society of the Philippines.
Halle, M. 2001. Infixation versus onset metathesis in Tagalog, Chamorro, and Toba Batak. In M. Kenstowicz, ed., Ken Hale: a life in language: 153-168. Cambridge, MA: MIT Press.
Harmon, Carol. 1977. Kagayanen and the Manobo subfamily of Philippine lan- guages. Ph.D. dissertation, University of Hawaii.
Himes, Ronald S. 2006. The Kalamian Microgroup of Philippine Languages. Paper Presented at the Tenth International Conference on Austronesian Linguistics 17-20 January 2006 Palawan, Philippines. 23pp.
Himmelmann, N.P. 1991. The Philippine challenge to Universal Grammar. Arbeitspapier Nr. 15 (Neue Folge). Köln: Institut für Sprachwissenschaft, Universität zu Köln.
Himmelmann, N.P. and J.U. Wolff. 1999. Toratán (Ratahan). Languages of the World/Materials 130. Munich: Lincom Europa.
Himmelmann, Nikolaus P. 2005. The Austronesian languages of Asia and Madagascar: Typological characteristics. In Karl Alexander Adelaar & Nikolaus P. Himmelmann (eds.), *The Austronesian languages of Asia and Madagascar*, 110-181. London: Routledge.
Himmelmann, Nikolaus P. 2006. How to miss a paradigm or two: Multifunctional ma- in Tagalog. In: Felix Ameka, Alan Dench and Nicholas Evans (eds) *Catching language: The standing challenge of grammar writing*, 487–526. Berlin: Mouton de Gruyter.
Himmelmann, Nikolaus P. 2008. Lexical categories and voice in Tagalog. In Peter Austin & Simon Musgrave (eds.), Voice and Grammatical Relations in Austronesian Languages, 247-293. Stanford: CSLI.

Himmelmann, Nikolaus P. 2016. Notes on 'noun phrase structure' in Tagalog. In Jens Fleischhauer, Anja Latrouite & Rainer Osswald (eds.), Explorations of the Syntax-Semantics Interface, 319-341. Düsseldorf: dup.
Hopper, Paul and Sandra Thompson. 1980. Transitivity in grammar and discourse. Language 56, 2:251-299.
Johnson, R. A. 1996. A bibliography of Philippine linguistics. Manila: LSP.
Kaufman, Daniel. 2011. 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*, pp. 721-754. Amsterdam: John Benjamins.
Kerr, Harland B. 1965. The case-marking and classifying function of Cotabato Manobo voice affixes. *Oceanic Linguistics* 4:15-47.
Kroeger, Paul R. 1993. Phrase structure and grammatical relations in Tagalog. CSLI Publications, Stanford CA.
Kroeger, Paul R. 1998. "Clitics and clause structure in Tagalog." In Bautista (ed.), Pagtanaw: Essays on language in honor of Teodoro A. Llamzon, pp. 53–72. Linguistic Society of the Philippines, Manila.
Kroeger, Paul. 2008. "Result complex predicates in Kimaragang Dusun." In Phil Quick & Joanna Smith (editors), 10-ICAL Borneo, Sulawesi, and Philippines papers. Studies in Philippine Languages and Cultures vol. 19. preliminary draft available at: https://sil-philippines languages.org/ical/papers/Kroeger-Predicates%20in%20Kimaragang%20Dusun.pdf
Kroeger, Paul. 2014. An affectedness constraint in Kimaragang restructuring. In Miriam Butt & Tracy Holloway King (eds.), Proceedings of the LFG14 Conference, pp. 283–303. Stanford, CA: CSLI Publications. http://web.stanford.edu/group/cslipublications/cslipublications/LFG/19/papers/lfg14kroeger.pd
Larson, Virginia. 1986. Ivatan texts. Manila: Summer Institute of Linguistics.
Leben, William. 1973. Suprasegmental Phonology. PhD dissertation, MIT. Distributed by Indiana University Linguistics Club.
Lendoyro, Constantino. 1909. *The Tagalog Language*. Manila: Juan Fajardo.
Liao, H.C. 2004. Transitivity and ergativity in Formosan and Philippine languages. Ph.D. dissertation. Honolulu: Department of Linguistics, University of Hawaii.
Liao, H.C. 2008. A typology of first person dual pronouns and their reconstructibility in Philippine languages. OL 47: 1-29.
Liao, Hsiu-chuan. 2004. Transitivity and ergativity in Formosan and Philippine languages. Ph.D. dissertation, University of Hawai'i at Mānoa.
Lobel, J.W. 2004. Old Bikol -um- vs. mag- and the loss of a morphological paradigm. *Oceanic Linguistics* 43: 469-497.
Lobel, J.W. 2013. Philippine and North Bornean languages: issues in description, subgrouping, and reconstruction. Unpublished Ph.D. dissertation. Honolulu: Department of Linguistics, University of Hawai'i.
Lobel, Jason William and Hall, William C. 2010. 'Southern Subanen Aspiration.' *Oceanic Linguistics* 49(2):319--338.

Lobel, Jason William and Riwarung, Labi Hadji Sarip. 2009. 'Maranao revisited: An overlooked consonant contrast and its implications for lexicography and grammar.' *Oceanic Linguistics* 48(2):403--438.
Lobel, Jason William and Riwarung, Labi Hadji Sarip. 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. 2004. 'Old Bikol -um- and mag- and the loss of a morphological paradigm.' *Oceanic Linguistics* 43(2):469-497.
Lynch, John, Malcolm Ross & Terry Crowley. 2002. *The Oceanic languages*. London/New York: Routledge.
McFarland, C.D. 1976. A provisional classification of Tagalog verbs. SLCAA Monograph Series, No. 8. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
McFarland, Curtis. 1978. 'Definite objects and subject selection in Philippine languages.' In Edrial-Luzares, Casilda & Austin Hale (eds.) Studies in Philippine Linguistics, Vol. 2. Manila: Linguistic Society of the Philippines, 139-182.
McKaughan, Howard P. 1958. *The inflection and syntax of Maranao verbs*. Manila: Institute of National Language.
McKaughan, Howard P. 1958. *The inflection and syntax of Maranao verbs*. Publication of the Institute of National Language. Manila, Bureau of Printing.
Medina, Pol Jr. 1995. *Polgas: Ang Asong Hindi*. Manila: Pugad Baboy Inc.
Miller, Jeanne, and Helen Miller. 1976. *Mamanwa grammar*. Language Data, Asian-Pacific Series, Number 8. Huntington Beach, California: Summer Institute of Linguistics.
Miller, Jeanne. 1973. Semantic structure of Mamanwa verbs. *Linguistics* 110:74- 81.
Naylor, P.B. 1986. On the pragmatics of focus. In Geraghty, Carrington, and Wurm, 1: 43-57.
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 Geraghty, Paul, Lois Carrington & S. A. Wurm (eds.) FOCAL 1: Papers from the Fourth International Conference on Austronesian Linguistics. 43-57.
Nolasco, Ricardo & Maureen Saclot. 2005. 'M- and S-transitivity in Philippine type languages.' Proceedings of the RRG Conference 2005. Taiwan: Taipei.
Nolasco, Ricardo Ma. 2003. Ang pagkaergatibo at pagkatransitibo ng mga wikang Pilipino: Isang Pagsusuri sa Sistemang Bose. Ph.D Dissertation. Unibersidad ng Pilipinas Diliman Quezon City.
Nolasco, Ricardo. 2005. 'What ergativity in Philippine languages really means.' Proceedings of the Taiwan-Japan Joint Workshop on Austronesian Languages, 215-238.
Pallesen, Kemp A. 1985. *Culture contact and linguistic convergence*. Linguistic Society of the Philippines Special Monograph Issue 24. Manila: Linguistic Society of the Philippines.
Pittman, Richard. 1966. Tagalog -um- and mag-, an interim report. Linguistic Circle of Canberra Publications A-8. Canberra: The Australian National University.
Porter, D. 1977. A Tboli grammar. Philippine Journal of Linguistics Special Monograph No. 7. Manila: LSP.
Quakenbush, J. Stephen and Edward Ruch. 2006. Pronoun ordering in Kalamianic. *Tenth International Conference on Austronesian Linguistics*, 17-20 January 2006, Palawan,

Philippines. Linguistic Society of the Philippines and SIL International. 19 p. http://www.sil.org/asia/philippines/ical/papers.html
Quakenbush, J. Stephen, Hendrickson, Gail R. and Josenita L. Edep. 2010. A brief overview of Agutaynen grammar. Manila: SIL Publishing. www.sil.org/asia/philippines/plb_download.html
Quakenbush, J. Stephen. 2005. Some Agutaynen grammatical details: personal pronouns nominal markers, and tense/aspect/mode. In Danilo T. Dayag and J. Stephen Quakenbush (eds.) *Language and Education in the Philippines and beyond: A festschrift in honor of Maria Lourdes S. Bautista*, 437-477. Manila: De La Salle University Press.
Quakenbush, Stephen J. 1991. "Agutaynen glottal stop". *Work Papers of the Summer Institute of Linguistics*, University of North Dakota Session: 35 (4).
Rackowski, Andrea and Norvin Richards. 2005. Phase Edge and Extraction: A Tagalog Case Study. *Linguistic Inquiry* 36:4, p.565-599.
Ramos, Teresita V. 1971. *Tagalog structures*. PALI Language Texts: Philippines. Honolulu: University of Hawaii Press.
Reid, Lawrence and Hsiu-chuan Liao. 2004. A brief syntactic typology of Philippine languages. Language and Linguistics 5(2):433-490.
Reid, Lawrence. 1978. Problems in the reconstruction of Proto-Philippine construction markers. In Wurm and Carrington: 33-66.
Reid, Lawrence. 1982. The demise of Proto-Philippines. In Halim, Carrington, and Wurm 2: 201-216. Canberra: Pacific Linguistics.
Reid, Lawrence. 1992. 'On the development of the aspect system in some Philippine languages.' *Oceanic Linguistics* 31(1):65-91.
Reid, Lawrence. 2002. Determiners, nouns, or what? Problems in the analysis of some commonly occurring forms in Philippine languages. OL 41: 295-309.
Reid, Lawrence. 2007. Another look at the marking of plural personal noun constructions in Austronesian languages. *Oceanic Linguistics* 46(1): 232-252.
Reid, Lawrence. 2009. The reconstruction of a dual pronoun to Proto Malayo Polynesian. In Evans: 461-477.
Revel-Macdonald, Nicole. 1979. Le Palawan (Philippines). Langues et Civilisa- tions de l'Asie du Sud-Est et du Monde Insulindien Monograph # 4. Paris: Centre National de la Recherche Scientifique.
Ross, Malcolm. 1995. Reconstructing Proto Austronesian verbal morphology: evidence from Taiwan. In Li et al: 727-791.
Ross, Malcolm. 2002. The history and transitivity of western Austronesian voice and voice marking. In Wouk and Ross: 17-62.
Ross, Malcolm. 2005. Reconstructing Proto Austronesian Verb Classes. *Language and Linguistics* 16(3) 279–345.
Ross, Malcolm. 2006. Reconstructing the case-marking and personal pronoun systems of Proto Austronesian. In Chang, Huang and Ho: 521-563.
Ross, Malcolm. 2009. Proto Austronesian verbal morphology: a reappraisal. In Adelaar and Pawley: 295-326.
Ryan, Kevin. 2010. Variable affix order: grammar and learning. *Language* 86.4: 758-791.
Salas Reyes, V, N.L. Prado and R.D.P. Zorc. 1969. *A study of the Aklanon dialect*. Kalibo, Aklan (Philippines).

Scebold, Robert A. 2003. *Central Tagbanwa: a Philippine language on the brink of extinction: sociolinguistics, grammar, and lexicon*. Linguistic Society of the Philippines Special Monograph Issue, Number 48. Manila: Linguistic Society of the Philippines.
Schachter, Paul and Fe Otanes. 1972. Tagalog Reference Grammar. Berkeley: University of California Press.
Schachter, Paul, and Fe T. Otanes. 1972. *Tagalog Reference Grammar*. Berkeley, Los Angeles, and London: University of California Press.
Schachter, Paul. 1973. Constraints on clitic order in Tagalog. Parangal kay Cecilio Lopez. ed. by Andrew Gonzalez, pp. 214-31. Manila: Linguistic Society of the Philippines.
Schachter, Paul. 1976. The subject in Philippine languages: actor, topic, actor-topic, or none of the above. In Li, ed. 491-518.
Schachter, Paul. 1994. The subject in Tagalog: still none of the above. UCLA Papers in Linguistics, 15, 1-61.
Shand, Jean. 1964. Categories and markers of tense, focus, and mode in Ilianen Manobo. OceanicLinguistics3:58-68.
Shibatani, M. 1988. Voice in Philippine languages. In M. Shibatani, ed., Passive and voice: 85-142. Amsterdam/Philadelphia: John Benjamins Publishing Company.
Starosta, Stanley, Andrew K. Pawley, and Lawrence A. Reid. 1982. The evolution of focus in Austronesian. In *Papers from the Third International Conference on Austronesian Linguistics, vol. 2, Tracking the travellers*, ed. by Amran Halim, Lois Carrington, and S. A. Wurm, pp. 145-170. Pacific Linguistics C-75. Canberra:The Australian National University.
Stevens, Alan M. 1969. Case grammar in Philippine languages. Paper read at the Fifth Meeting of the Linguistic Society of America, San Francisco.
Svelmoe, Gordon, and Thelma M. Svelmoe. 1974. *Notes on Mansaka grammar*. Language Data, Asian-Pacific Series, Number 6. Huntington Beach: Summer Institute of Linguistics.
Trick, Douglas. 2006. Ergative control of syntactic processes in Sama Southern. 10th International Conference on Austronesian Linguistics, 17-20 January, Palawan.
Venago, Martin F. 1929. Ang mga paring Pilipino sa kasaysayan ng Inang bayan. Maynila: [s. n.].
Walton, Charles. 1986. *Sama verbal semantics: Classification, derivation and inflection*. Linguistic Society of the Philippines, Special Monograph Issue.
Walton, Charles. 1997. The uses of pseudo-cleft sentences in Sama narrative discourse. *Journal of Translation and Textlinguistics* 9. 15-34.
Wang, Peter, Robert Hunt, Jeff McGriff and Richard E. Elkins. 2006. *The Grammar of Matigsalug Manobo*. The Summer Institute of Linguistics and the University of Texas at Arlington.
Wegmüller, Ursula. 1998. Sentence structure and ergativity in Tagalog. (Arbeitspapiere / Universität Bern, Institut für Sprachewissenschaft, 36.) Bern: Univ. Bern, Inst. f. Sprachwiss.
Wolfenden, Elmer. 1961. *A Re-statement of Tagalog Grammar*. Manila: Summer Institute of Linguistics and the Institute of National Language.
Wolff , John U. 1974. Proto-Austronesian *r and *d. *Oceanic Linguistics*. 13 (1/2): 77-121.
Wolff, John U. 1966. *Beginning Cebuano I*. New Haven: Yale University Press.
Wolff, John U. 1967. *Beginning Cebuano II*. New Haven: Yale University Press.

Wolff, John U. 1973. Verbal inflection in Proto-Austronesian. In *Parangal Kay Cecilio Lopez*, ed. by Andrew B. Gonzalez, pp. 71-91. Philippine Journal of Linguistics Special Monograph Issue No. 4.
Wu, Jiang. 2017. Plural Words in Austronesian Languages: Typology and History. MA thesis, Leiden University.
Zorc, R. David. 1977. *The Bisayan dialects of the Philippines: Subgrouping and reconstruction*. Pacific Linguistics C-44. Canberra. The Australian National University.
Zorc, R.D. 1972. Current and Proto-Tagalic stress. PJL 3.1: 43-57.
Zorc, R.D. 1978. Proto-Philippine word accent: innovation or Proto-Hesperonesian retention? In Wurm and Carrington: 67-119.
Zorc, R.D. 1979. On the development of contrastive word accent: Pangasinan, a case in point. In N.D. Liem, ed., South-East Asian Linguistic Studies, vol. 3: 241-258. Canberra: Pacific Linguistics.
Zorc, R.D. 1983. Proto Austronesian accent revisited. PJL 14.1: 1-24.
Zorc, R.D. 1986. The genetic relationships of Philippine languages. In Geraghty, Carrington, and Wurm 2: 147-173.
Zorc, R.D. 1996. The reconstruction and status of Austornesian glottal stop–chimera or chameleon? In *Reconstruction, classification, description: festschrift in honor of Isidore Dyen,* ed. by Bernd Nothofer, 41-72. Hamburg: Aberla Verlag.
Zuraw, Kie. 2000. Patterned Exceptions in Phonology. Ph.D. dissertation, UCLA.
Zuraw, Kie. 2007. The Role of Phonetic Knowledge in Phonological Patterning: Corpus and Survey Evidence from Tagalog Infixation. Language 83.2: 277-316.