

The Central Luzon Group of Languages

Ronald S. Himes

SAN DIEGO STATE UNIVERSITY

The Central Luzon microgroup of Philippine languages is composed of Kapampangan, Sinauna, the three major dialects of Sambal, and the Ayta languages spoken in and around the Zambales Mountains in Zambales, Pampanga, and Bataan provinces. A defining phonological feature of this group is the regular /y/ reflex of Proto–Malayo–Polynesian *R. The languages and dialects in question also share similar pronoun sets and a number of lexical and other innovations. The Northern Mangyan and Bashiic languages also reflect *R as /y/. The former group probably clusters with Central Luzon, but the Bashiic languages lack substantial sharing of innovative items with Central Luzon.

1. THE CENTRAL LUZON LANGUAGES.¹ The Central Luzon (LUZC) group of languages includes Kapampangan (or Pampango), Sinauna (or Sinauna Tagalog), three dialects of Sambal (Bolinao, Tina, and Botolan), and a number of languages spoken by Ayta Negrito populations.²

Kapampangan is spoken primarily in Pampanga and southern Tarlac Provinces. The data used here were collected in the field in Guagua and Angeles, Pampanga, and have been augmented by data from Yap (1977) and McFarland (1977).

1. Most of the information used in this paper was gathered in the field in 1964, 1968, 1995, and 1999. Alangan and Tadyawan data were provided by Mr. James Clifford and Mr. Hugh Porter, respectively. Some of the Central Luzon data are taken from McFarland (1977). The Sinauna data are from Santos (1975), and some of the Ayta information was collected by others and generously provided by Dr. Lawrence A. Reid and Dr. Thomas Headland. Additional comparative data are taken from my language files and from Carro (1956), Headland and Headland (1974), Lambrecht (1978), McFarland (1977), Panganiban (1973), Pennoyer (1986/97), Reid (1971, 1976, 1991), Santos (1975), Scott (1957), Soberano (1976), Steller (1982), Tharp and Natividad (1976), Tryon (1994), Tsuchida (1987), Tweddell (1958), Usup (1980), Vanoverbergh (1933, 1956, 1972), Yamada (1965, 1967), Yamashita (1992), and Yap (1977). I wish to thank Dr. Lawrence A. Reid for his useful comments on an earlier version of this paper. I would also like to thank the anonymous reviewers for their valuable suggestions. Any errors of fact or interpretation are, of course, mine.
2. The abbreviations used in this paper are: ABE, Abellen; ALA, Alangan; AMB, Ambala; BSH, Bashiic; BOL, Bolinao; BOT, Botolan; CBG, Cabangan; IRA, Iraya; IVT, Ivatan; IVTI, Itbayatan Ivatan; KIL, Kakilingan; KPM, Kapampangan; LUZC, Central Luzon; MAN, Mag-antsi; MBK, Magbukun; MIN, Mag-indi; MNGN, Northern Mangyan; PAN, Proto-Austronesian; PHN, Proto-Hesperonesian; PLUZC, Proto-Central Luzon; PPH, Proto-Philippines; PSBL, Proto-Sambalic; PSBLBL, Proto-Bolinao Sambal; PSBLBT, Proto-Botolan Sambal; PSBLSIN, Proto-Sambal-Sinauna; PSBLT, Proto-Sambal Tina; PWMP, Proto-Western Malayo-Polynesian; REM, Remontado; SBLBL, Bolinao Sambal; SBLBT, Botolan Sambal; SBLT, Sambal Tina; SCR, Santa Cruz; SIN, Sinauna; TDY, Tadyawan; VIL, Villar; YAM, Yami.

Sinauna speakers are, or were (Reid 2010), found in remote mountain areas of Tanay municipality in Rizal Province, and in Infanta, Quezon. Our knowledge of Sinauna is limited to the information provided by Santos (1975), and the source for Remontado is a list collected at Paimahuan, Limoutan, General Nakar, Quezon.

Bolinao is spoken in Bolinao and Anda municipalities of western Pangasinan Province. Sambal Tina is spoken in Zambales Province, in the municipalities of Santa Cruz, Candelaria, Masinloc, Palauig, and Iba. The Botolan dialect is spoken in Botolan and Cabangan municipalities.

Wimbish (1986) identifies six distinct Ayta dialects: Abelen, Aberlen, Magqanchi (or Mag-antsi), Magqindi (or Mag-indi), Ambala, and Magbukun (or Magbeken). Mag-antsi speakers, or at least some of them, have been displaced by the 1991 eruption of Mt. Pinatubo. The data used here are from a population originally from Cagmang, Porac, in Pampanga Province who have been moved to Sta. Juliana, Capas in Tarlac.

All of the information available on Mag-indi is from dialects spoken in Florida Blanca, Lumibao, and Maague-ague in Pampanga Province.

The Ambala data used here were collected in Pastolan, Subic Bay Metropolitan Authority, and Gordon Heights, Olongapo City. Two SIL lists of Ambala are from Maliwacat, Cabalan, Olongapo and Batong Kalyo (Pili), San Marcelino municipality, Zambales.

Magbukun data were collected at Biaan, Mariveles, and Canawan, Morong, both in Bataan Province. Additional data on the Canawan dialect and that of Bayan Bayanan, Magbikin, Mariveles, were taken from SIL lists.

There are also Ayta in Botolan and Cabangan municipalities who speak varieties of the Botolan dialect of Sambal. Phonologically they do not differ significantly from the Botolan dialect, although lexically there are some items peculiar to their own communities. Examples of these are the locales of Villar, a *sitio* of Botolan and *sitio* Kakilingan, Santa Fe, Cabangan (Yamashita 1992). The dialect of Abellen, collected from a displaced respondent in Botolan, originally from *sitio* Loob-Bunga in the *barangay* of Poon Bato, Botolan, falls into this category as well.

This paper is an attempt to portray Kapampangan, the Sambalic languages and dialects, and Sinauna as a distinct microgroup of Philippine languages, and to examine the internal structure of this subfamily. First, the phonological changes that have occurred in the Central Luzon sound systems from their parent language, Proto-Malayo-Polynesian, are described. Following the phonological discussion is a treatment of the LUZC pronoun system, focusing on the long nominative forms. Section 4 presents lexical innovations that are attributable to PLUZC and to lower level language/dialect groupings. What little information can be gleaned from lexicostatistical analysis is presented in section 5. Finally, the relationship of the LUZC languages to other Philippine microgroups is discussed in section 6.

2. PHONOLOGICAL CHANGES. This section presents changes from the Proto-Malayo-Polynesian (PMP) sound system to that of Proto-Central Luzon.

2.1 PROTO-MALAYO-POLYNESIAN PHONEMIC INVENTORY. It is generally agreed that Proto-Malayo-Polynesian had the following consonant phonemes:

(1)	*p	*t		*k	*q
	*b	*d	*z	*j	*g
		*s			*h
	*m		*n	*ñ	*ŋ
		*l			
		*r			*R
	*w		*y		

The PMP vowels were *i, *e (schwa), *u, and *a.

2.2 PROTO-PHILIPPINES. Blust (1991), Zorc (1986, 1994), and others (as presented in Wurm and Wilson 1975), including this writer (2001, 2007), have reconstructed lexical items at a level later than PMP and ancestral to reflexes found in two or more subgroups in the Philippines. This level is termed Proto-Philippines (PPH). Others, such as Reid (1982) and Ross (2005) hold that evidence for such a protolanguage is unconvincing. The likely succession of languages proceeds directly from PMP to the descendant languages reconstructible from contemporary languages in the Philippines, Oceania, and elsewhere. For purposes of this paper this is immaterial. Included here are items reconstructed by Blust, Charles, Healey, and Zorc as they appear in Wurm and Wilson (1975), Blust and Trussel (in progress), and Zorc (1986, 1994). Items reconstructed by this author and which do not appear elsewhere and are labeled PPH are listed in the appendix. We can take these as attributable to some level above that of Central Luzon, whether they are traceable to Proto-Malayo-Polynesian or to some descendant stage of that language.

2.3 EARLY CENTRAL LUZON SOUND CHANGES. At some point prior to the dispersal of the Central Luzon (LUZC) languages, certain of the PMP phonemes had undergone change. Among the earliest of sound changes to have affected the reflexes of PMP phonemes in PLUZC are the following:

(2)	Unstressed *e	>	Ø / C __ C
	*q	>	ʔ
	*h	>	Ø ³
	*ñ	>	n
	*Ø	>	ʔ / # __
	*R	>	y
	*j	>	d
	*d	>	l / V [-back] __ V
	*z	>	d

The Central Luzon languages undoubtedly merged PMP *z with d, but this must have occurred after changes affecting the reflexes of *d.

The PLUZC phonemic inventory included those listed in (3). PLUZC reflected the four vowel system of PMP, *i, *i̯, *u, and *a, with the *i̯ representing the PMP schwa.

3. Reid (pers. comm.) suggests that *h became ʔ in initial position, since no Philippine language has VC vs. CVC initial syllable structure. This may well be the case, but *h certainly was lost in all other environments.

- (3) *p *t *k *ʔ
 *b *d *z *j *g
 *s
 *m *n *ŋ
 *l
 *r
 *w *y

2.3.1 *e [- stress] > Ø / C __ C. A sound change rule common throughout the Philippines is schwa syncope (and apocope), which deletes *e when unstressed and not blocked by rules of canonical form:

- | | | | | |
|-----|------------|----------------|--------------|------------------------|
| (4) | | PLUZC | | |
| | ‘lie down’ | PPH *hi-Rezaʔ | *ʔ<um>i-yzaʔ | midaʔ SBLBL |
| | | | | miraʔ SBLT, SBLBT |
| | | | *ka-i-yzaʔ | ke:ra KPM |
| | ‘to fly’ | PPH *l-um-epad | *lumpad | lumpád SBLBL, SBLBT |
| | | | | lumpár SBLT |
| | ‘sated’ | PMP *besúR | *na-bsuy | nabsúy SBLBL |
| | | | | nabhúy SBLT |
| | | | *ma-bsuy | mabsíʔ KPM |
| | ‘rib’ | PMP *tageRaŋ | *tagyaŋ | tagyáŋ SBLBL, SBLT, |
| | | | | SBLBT, KPM, SIN |
| | ‘new’ | PMP *baqeRu | *baʔyu | baʔyú SBLBL, SBLT |
| | | | | ba:yu SBLBT, KPM |
| | | | | báʔyu, ba:yu SIN |

As in Northern Philippines (Cordilleran) languages, at least, schwa syncope also functions synchronically in LUZC languages.

2.3.2 *q > ʔ. In Central Luzon languages, as in all Philippine languages except Tboli and the Kalamianic languages, the PMP phoneme *q is reflected as glottal stop:

- | | | | | |
|-----|-------------|-------------|---------|------------------------|
| (5) | | PLUZC | | |
| | ‘earthworm’ | PMP *qu:lej | *ʔu:lid | ʔu:rír SBLBL |
| | | | | ʔu:lul SBLT |
| | | | | ʔu:wíl SBLBL |
| | | | | u:lad KPM |
| | | | | ʔu:lad SIN |
| | ‘new’ | PMP *baqeRu | *baʔyu | baʔyú SBLBL |
| | | | | baʔyú SBLT |
| | | | | ba:yo SBLBT |
| | | | | ba:yu KPM |
| | | | | báʔyu, ba:yu SIN |
| | ‘blood’ | PMP *daRaŋ | *da:yaʔ | da:yaʔ SBLBL, SBLT, |
| | | | | SBLBT, KPM |
| | ‘nose’ | PPH *qaʔjuŋ | *ʔaʔduŋ | ʔaʔrúŋ SBLBL |
| | | | | ʔaʔlúŋ SBLT |
| | | | | a:ruŋ KPM |
| | | | | ʔaʔduŋ SIN |

2.3.3 *h > Ø. Proto-Philippine *h was lost:

			PLuzC		
(6)	‘to sew’	PMP *tahiŋ	*taiʔ	manayíʔ	SBLBL, SBLT, SBLBT
				tayíʔ	SBLBT, KPM
	‘to rub’	PPH *haprus	*aprus	ʔaprus	SBLBL
				ʔaplús	SBLT
				aplús	KPM
	‘knee’	PMP *tuhud	*tuud	tu:ʔud	SBLBL, SIN
				tu:ʔur	SBLT
				tu:ʔul	SBLBT
				tu:d	KPM
	‘tear’	PMP *luhaŋ	*luaʔ	luwáʔ	SBLBL, SBLT,
				lu:waʔ	SBLBT
				luwáʔ	KPM

2.3.4 *ñ > n. There is very little evidence of reflexes containing *ñ in LUZC languages. But based on one instance we can infer that *ñ merged with *n in the Sambalic languages:

(7)			PLUZC		
	‘carried by current’	PMP *qañud	*ʔanud	ʔa:nur	SBLBL
				ʔa:nul	SBLT, SBLBT

KPM *ʔanyud* seems to indicate that, in this one item at least, the phoneme was split into its component parts prior to this change.

2.3.5 *Ø > ʔ / # __. Words with initial vowels, either through inheritance or loss of *h, acquired an initial glottal stop to fit the canonical form CV(C). Kapampangan (and possibly Remontado) subsequently reversed the effects of this rule.

(8)			PLuzC		
	‘to rub’	PPH *haprus	*aprus	ʔaprus	SBLBL
				ʔaplús	SBLT
				aplús	KPM
	‘pestle’	PPH *halqu	*alʔu	ʔaʔlu	SBLBL, SBLT
				a:lu	KPM
	‘tail’	PMP *ikuR	*ikuy	ʔi:kuy	SBLBL, SBLT, SBLBT, SIN
				i:kiʔ	KPM
	‘vein’	PMP *uRát	*uRát	ʔuyát	SBLBT, SBLT, SBLBT, SIN
				uyát	KPM

2.3.6 *ii > /i/. As in many other Philippine languages the vowel cluster *ii (or *iy) was reduced to /i/. This rule is recursive.

(9)			PLUZC		
	‘urine’	PMP *ʔihiʔ	*ʔiiʔ	i:ʔ	KPM
	‘wind’	PPH *si:riR	*si:riy	si:ri	SBLBL

Other instances of vowel cluster reduction generally involve simplification when a prefix ending in a vowel is appended to a vowel-initial root.

(10)		PLUZC		
‘lie down’	PPH *hi-Reza?	*ʔ<um>i-ya?	mida?	SBLBL
			mira?	SBLT, SBLBT
		*ka-i-ya?	ke:ra	KPM

Cases of diphthong simplification are treated below.

Another feature, one commonly found throughout the Philippines, that is shared by all the LUZC languages except Kapampangan, is the insertion of a glottal stop to break up certain vowel clusters. A sequence of identical vowels (other than *i) is broken up by the glottal stop.

(11)		PLUZC		
‘knee’	PMP *tuhud	*tuud	tuʔud	SBLBL, SIN
			tu:ʔur	SBLT
			tu:ʔul	SBLBT
‘inside’	PPH *luub	*luub	lu:ʔub	SBLT, AMB, MBK
‘banana’	PPH *saha?	*saa?	ha:ʔa	SBLBT, ABE, MAN
			sa:ʔa?	MIN
‘sea’	PMP *tahaw	*taaw	taʔáw	SBLBL

Sequences of unlike vowels are separated into their two components by a glide: /y/ if one of the vowels is a front vowel and /w/ otherwise. These cases are treated below under the rules that are language specific.

2.3.7 *R > /y/. A characteristic of all LUZC languages, as well as the languages of the Bashiic microgroup and the Northern Mangyan languages, is the shift from *R to /y/. There is some environmentally determined variation of this in Sinauna and Kapampangan, and this is discussed in 2.4.1 and 2.5.1, respectively, below.

2.3.8 *d > *z /u __ V. In those items inherited from higher levels that contain *d intervocalically between *u and another vowel, the *d merges with *z.

(12)		PLUZC		
‘to spit’	PMP *ludaq	*luza?	ludá?	SBLBL
			lurá?	SBLT, KPM, SIN
‘cooking pot’	PMP *kúden	*ku:zin	ku:ɖin	SBLBL
			ku:run	SBLT
			ku:rín	SBLBT
			ku:ran	KPM
			ku:dan	SIN
‘to give’	PPH *tudul	*tuzul	tu:ruʔ, tu:zuʔ	MBUK
‘to sleep’	PMP *túduR	*tu:zuR	matu:ruy	SBLT ⁴

Were it not for this early shift from *d to *z, Kapampangan would have ***lula?* and ***ku:lan* for ‘spit’ and ‘cooking pot’, respectively, instead of the actual *lu:ra?* and *ku:ran*.

2.4 SINAUNA. Phonological changes peculiar to Sinauna are treated here.

2.4.1 Sinauna reflexes of *R. In Sinauna, there are very few attested occurrences of word-initial *R. One that is clear is PMP *Rabii ‘night.’ Here we find the reflex *ʔabi*

4. Inexplicably, Botolan and two of the closely related Ayta dialects have *tu:luy* ‘to sleep.’

instead of the expected ***yabi*. Umiray Dumaget also has *ʔabi* for ‘night’, but by the rules of that language we would expect ***gebi* (Himes 2001). This indicates that Umiray Dumaget borrowed the item from Sinauna and that the latter adopted a rule whereby initial *R was lost. The only possible confirmatory evidence of this proposition is the term for ‘back (of a person)’, *ʔilúd* or *ʔilud*. A similar item appears in two dialects of Inagta from Buhi, Camarines Sur, in Bikol. There we find *golód* and *gulód* ‘back.’ These Inagta dialects regularly reflect PMP *R as /g/. Since they are spoken by Negrito peoples, we can hypothesize that there was an ancestral lexeme, Malayo-Polynesian or not, similar to *Rulud ‘back.’

Another environment in which *R is lost is in word-final position when preceded by the low vowel /a/:

(13)	‘answer’	PMP *tubaR	SIN tuba (REM)
	‘rattan’	PMP *wakáR ‘root’	SIN waká

Last, *R appears to be replaced in Sinauna by vowel length immediately before a stop:

(14)	‘child’	PPH *quRBun ‘young, small’	SIN ʔu:bun
	‘person’	PPH *qaRta	SIN ʔa:ta

but:

(15)	‘day’	PPH *qaRjaw	SIN ʔaydáw
	‘deer’	PMP *uRsa	SIN ʔi:sa

2.4.2 *ʔ in Sinauna. Sinauna exhibits only one change in its reflexes of the glottal stop. The cluster *Rʔ underwent metathesis:

(16)	‘heavy’	PMP *beReqat	PLUZC *ma-biyʔat	SIN mabáʔyat
------	---------	--------------	------------------	--------------

Elsewhere, Sinauna reflects consonant clusters containing the glottal stop as they were inherited from PLUZC:

(17)	‘new’	PMP *baqeRu	PLUZC *baʔyu	SIN báʔyu
	‘sweet’	PMP *tamqis	*tamʔis	matamʔis
	‘turtle’	PPH *pagquŋ	*pagʔuŋ	pagʔuŋ

2.4.3 Liquids. Sinauna usually shows reflexes of *l and *r as they were inherited:

(18)	‘to play’	PPH *garaw	PLUZC *garaw	SIN garaw
	‘house’	PMP *baláy	*baláy	baláy

However, there are several occurrences of both /l/ and /r/ in items inherited from PMP *l:

(19)	‘tongue’	PMP *dilaq	PLUZC *dilaʔ	SIN di:laʔ ~ diraʔ
	‘palm’	PMP *palaj	*palaj	pa:lad ~ pa:rad

2.4.4 Apicals. PMP *j and *z merged with the reflexes of *d in Sinauna. They may have done this simultaneously or *j may have shifted before *z did, as happened in other LUZC languages. Either way, the result would have been the same. Also, *d is usually reflected as /r/ intervocally.

(20)		PLUZC	SIN
‘worm’	PMP *qúlej	*ʔúlid	ʔu:lad
‘nose’	PPH *qāʔjuŋ	*ʔaʔduŋ	ʔaʔduŋ
‘name’	PMP *ŋajan	*ŋa:dan	ŋa:ran
‘path’	PMP *zálan	*da:lan	da:lan
‘rain’	PMP *quzan	*ʔudán	ʔurán

In Sinauna, if the *z was stem-initial, it is reflected as /d/ even after acquiring a prefix:

(21)		PLUZC	SIN
‘needle’	PMP *zaRum	*ka-da:yum	kada:yum
‘needle’	PPH *zaRem	*ka-da:yim	kadayam
‘far’	PMP *zayuq	*dayuʔ	sadyúʔ

There are some other occurrences of /d/ intervocalically, but all of them are suspected of being loans from Tagalog, Kapampangan, or Umiray Dumaget.

2.4.5 *e > /a/. Sinauna reflects the PLUZC high central vowel as /a/. It is occasionally raised to schwa.

(22)		PLUZC	SIN
‘six’	PMP *a-eném	*ʔaʔnim	ʔaʔnám
‘worm’	PMP *qúlej	*ʔu:lid	ʔúləd, ʔulad
‘hair’	PMP *buhék	*buík	búʔak

2.4.6 Glide insertion. It was noted above that Sinauna inserts a glottal stop to separate a sequence of identical vowels. There is only one unambiguous case of glide insertion to separate another vowel sequence in Sinauna. If the first vowel is stressed (or long) /a/ and the second is the front vowel, the glide /y/ is inserted.

(23)		PLUZC	SIN
‘woman’	PMP *bahi	*ba:i	SIN ba:yi

2.5 KAPAMPANGAN. Phonological changes from PMP and PLUZC to Kapampangan are treated in this section.

2.5.1 *R. *R is reflected as /y/ in all KPM environments except for two very limited ones.

In the data available there is only one instance of *R between *i and *u, in *diRuʔ ‘to bathe.’ Instead of the expected ***dīyu* we find *dilu* ‘to bathe.’ Since there is no apparent donor language, such as Southern Cordilleran, this is unlikely to be a loan. We, then, must construct the rule *R > / / *i __ *u.

Equally limited is a rule to account for the occasional loss of *R in the intervocalic environment *u __ *a. Here, the second vowel must be long (or stressed) and followed by a nonstop or a word boundary. This gives the following:

(24)		Pre-KPM	KPM
‘parent-in-law’	PMP *tuRáj	*ka-tuáj-an	katuwa:ŋan
‘to wash’	PMP *huRás-an	*uás-an	wa:san
‘anger’	PPH *muRá	*muá	muwá

If the low vowel is not stressed (or long), or if it is followed by a stop, *R is reflected as /y/:

(25)			Pre-KPM	KPM
	‘vein’	PMP *uRát	*uyát	uyát
	‘child-in-law’	PMP *tuRáj	*ma-núyaŋ	manu:yaŋ

2.5.2 *j > /d/. In Kapampangan, as in other LUZC languages, PMP *j merges with reflexes of *d, with /d/ becoming /l/ in an intervocalic environment.

(26)			PLUZC	KPM
	‘worm’	PMP *qúlej	*ʔu:lid	u:lad
	‘palm’	PMP *palaj	*palad	pa:lad
	‘charcoal’	PMP *qújiŋ	*ʔu:diŋ	u:liŋ
	‘rice’	PMP *pájey	*pa:diy	pa:le

2.5.3 Glottal stop. In Kapampangan, all instances of *Cʔ were reflected as *ʔC:

(27)			Pre-KPM	KPM
	‘bone’	PPH *butʔul	*buʔtul	bu:tul
	‘headcold’	PPH *sipqun	*siʔpun	si:pun
	‘shame’	PLUZC *di[nŋ]ʔiy	*diʔniy	di:ne

and *ʔ is reflected as vowel length before a consonant

(28)			Pre-KPM	KPM
	‘chin’	PMP *baqbaq ‘mouth’	*baʔbaʔ	ba:baʔ
	‘nose’	PPH *qaʔjuŋ	*ʔaʔduŋ	a:ruŋ
	‘ginger’	PMP *laqia	*laʔya	la:ya
	‘to cough’	PPH *kuʔkuʔ	*kuʔkuʔ	ku:kuʔ

Finally, inherited glottal stop is lost before a vowel:

(29)			Pre-KPM	KPM
	‘charcoal’	PMP *qújiŋ	*ʔu:diŋ	u:liŋ
	‘cogon’	PPH *qil[jdr]ib	*ʔidib	ilib
	‘sour’	PPH *ʔaslem	*maʔaslam	maslám
	‘here’	PMP *-ti	*ka-ʔiti	ke:ti
	‘thigh’	PPH *puʔej	*puʔad	puwad
	‘tree trunk’	PMP *puqun	*puʔun	pun

2.5.4 *r > /l/. PMP *r is reflected as /l/ in Kapampangan:

(30)			Pre-KPM	KPM
	‘outside’	PMP *luwar	*luwar	luwal
	‘bolo’	PMP *paraŋ	*paraŋ	palan
	‘rotten’	PPH *buruk	*buruk	bulúk
	‘to exchange’	PPH *ribay	*ipag-ribay	ipaglibe
	‘to rub’	PPH *haprus	*aprus	aplús

2.5.5 *z > /d/. At some time after intervocalic *d had shifted to /l/, the remaining apical *z was reflected as /d/. Intervocallically, /d/ then is reflected as /r/.

(31)			Pre-KPM	KPM
	‘path’	PMP *zálan	*da:lan	da:lan
	‘needle’	PMP *zaRum	*kadayum	kara:yum
	‘rain’	PMP *quzan	*udán	urán

‘sharp’	PMP *tazém	*ma-tadím	matarám
‘lie down’	PPH *hi-Reza?	*ka-ida?	ke:ra

2.5.6 *e > /a/. As in Sinauna, the high central vowel is lowered to /a/ in Kapampangan:

(32)		Pre-KPM	KPM
‘sibling’	PPH *pated	*ka-patid	kapatad
‘louse egg’	PPH *liqes	*liis	liyas
‘thigh’	PPH *pu?ej	*puid	puwad

2.5.7 Glide insertion. Vowel sequences in Kapampangan are broken up with an inserted glide if they are of a particular type. If the sequence contains a low vowel /a/ and high front vowel, and the cluster is in word-final position (with or without a following consonant), the palatal glide is inserted between the vowels.

(33)		Pre-KPM	KPM
‘bitter’	PMP *paqit	*ma-pait	mapayit
‘woman’	PMP *ba-báhi	*babai	baba:yi
‘louse egg’	PPH *liqes	*liis	liyas
‘alive’	PMP *bihaR	*biay	bye

As a corollary to this rule, sequences of /ai/ that are not in the final syllable of the word are conflated as the vowel /e/.

(34)		Pre-KPM	KPM
‘one’	PMP *hituŋ ‘count’	*ma-ítuŋ	me:tuŋ
‘here’	PMP *-ti	*ka-iti	ke:ti
‘lose’	PMP *wada? ‘exist’	*ma-iwalá?	mewalá?
‘lie down’	PPH *hi-Reza?	*ka-ida?	ke:ra

A sequence of the high back vowel /u/ followed by any other (non-identical) vowel is broken up by a labiovelar glide:

(35)		Pre-KPM	KPM
‘hair’	PMP *buhék	*buak	buwak
‘thigh’	PPH *pu?ej	*puid	puwad
‘tear’	PMP *luhaq	*lua?	luwa?

Sequences of two identical vowels are merged into a single vowel; if there is more than one vowel in the word, the conflated vowel is long or stressed:

(36)		Pre-KPM	KPM
‘knee’	PMP *tuhud	*tuud	tud
‘rich’	PPH *bahandi	*ma-baandi	mabándi
‘yesterday’	PPH *hapun	*na-apun	na:pun
‘tree trunk’	PMP *puqun	*puun	pun

2.5.8 Diphthong reduction. Diphthongs in final position or in closed syllables within a word were simplified. Inherited *-ay is reflected as /-e/, *-aw as /o/, and *-uy and *-iw are both reflected as /-i?/.

(37)		Pre-KPM	KPM
‘termite’	PMP *ʔánay	*a:nay	a:ne
‘house’	PMP *baláy	*baláy	bale

'egg'	PLUZC *ʔeRbun	*aybun	e:bun
'night'	PMP *beRɲi	*bayɲi	be:ɲi
'above'	PMP *babaw	*ba:baw	ba:bo
'housefly'	PMP *lájaw	*la:ɲaw	la:ɲo
'tail'	PMP *ikuR	*i:kuy	i:kiʔ
'lime'	PMP *qápuR	*a:puy	ʔa:piʔ
'to buy'	PMP *sáliw	*sa:liw	sa:liʔ

2.6 THE SAMBALIC BRANCH. The three dialects of Sambal, along with the Ayta languages, experienced some sound changes not shared with Sinauna and Kapampangan.

2.6.1 Final consonants. In Sambal and the Ayta languages, the glottal stop tends to replace a word-final nonobstruent when preceded by a stressed high central vowel *i. *R demonstrates this, in nonreduplicative words, in all the languages of this subgroup except Magbukun.

(38) 'to hear'	PAN *deNéR	riɲiʔ	SBLBL, SBLBT
		luɲúʔ	SBLT
		luɲúy	MBK

But:

(39) 'neck'	PMP *liqeR	li:ʔiy	SBLBL
		lu:ʔuy	SBLT, MBK
		li:ʔiy	SBLBT

Admittedly, there is only one instance of this, and, therefore, it might be more conveniently described as a phonological innovation in Sambal. But the shift to final glottal stop occurs with other word-final phonemes as well, giving the appearance of a more general rule. Glottal stop also replaces word-final *l when preceded by stressed *i in the entire Sambalic branch of the family.

(40) 'many'	PPH *dakel	lakúʔ	SBLT
		lakíʔ	SBLBT
		malakúʔ	MBK
'itch'	PMP *gatél	gatiʔ	SBLBL
		magatúʔ	MBK

Other instances of replacement by glottal stop, at least in some Sambalic languages, affect the nasals *m and *ɲ:

(41) 'deep'	PMP *dálem	ʔara:rim	SBLBL
		mala:luʔ	SBLT, MBK
		mala:liʔ	SBLBT
'arrive'	PMP *daten	ratíʔ	SBLBL
		latúʔ	SBLT
		latiɲ	SBLBT

Loss of a final *s, or its replacement by glottal stop, also occurs sporadically when preceded by a nonfront vowel:

(42) ‘monkey’	PPH *bákes	ba:ki?	SBLBL, SBLBT
		ba:ku?	SBLT
‘to tie’	PPH *tagkes	tagkús	SBLT (SCR), MBK
		tagkí?	ABE, AMB
‘hard’	PMP *teRas	matyá?	SBLBT
‘bad’	PPH *duk[ae]s	duká?	SBLBL, SBLT
‘difficult’	PPH *diRas	madiyá?diyá?	SBLT, SBLBT
‘nose’	PPH *balunus	balu:ŋu	SBLBT
		balu:ŋuh	MAN
		balunus	MIN, MBK
		balu:ŋu?	AMB

The data are sparse because not every etymon has a reflex in each language, and not every language/dialect applies the process to every final consonant. As a result, we are left with the necessity of creating a series of ad hoc rules for each Sambalic language to account for the data at hand. Suffice it for present purposes to note that this process occurs, and that it seems to be limited to this branch of the microgroup. It should also be noted that, since PLUZC did not permit a word to end in the high central vowel, with or without a glottal stop, any innovative terms that do so end should be assumed to have had an earlier form with a final consonant, probably *R, *l, *ŋ, or *s. Thus, the innovative form meaning ‘to sit’ that gives *tumikri?* in the Bolinao dialect and *tumuklí?* in Sambal Tina probably should have a shape such as *tikdíl.

Finally, in Sambal there is a tendency for the unstressed sequence *eR (> *iy) to conflate to /i/. It is unclear if this occurs in some prefixed forms, since *e syncope has the same effect.

(43) ‘to give’	PMP *beRéy	biyán	SBLBL, SBLT, SBLBT
		mamí	SBLBL
		ʔibí	SBLT
		mambí	SBLBT
‘heavy’	PMP *ma-beRat	mabyát	SBLBL, SBLT, SBLBT
		mabiyát	SBLT, AMB
‘rice’	PMP *beRas	byas, biyás	SBLBL
		buyáh	SBLT, SBLBT
		biyáh	AMB
		biyás	MIN
		buyás, buyáh	MBK
‘ankle’	PLUZC *peR-lepu-an	pilpu?án	ABE, MIN ⁵

2.6.2 Glottal stop. The glottal stop has undergone change in most of the Sambalic speech forms, Bolinao being an exception. In Sambal Tina, the only change is loss of the glottal after a consonant.

(44)		PSBLT	SBLT
‘bone’	PPH *butqul	*but?ul	butúl
‘turtle’	PPH *pagquŋ	*pag?uŋ	pagúŋ

5. Cf. Tadyawan *poylupuán* ‘ankle’ < *paR=lepu-an from PWMP *lepuq ‘crooked, of limbs.’

In Botolan and all of the Ayta languages for which data are available, Mag-antsi, Mag-indi, and Ambala, the cluster *ʔl is reflected as *lʔ; the Magbukun data lack confirming evidence.

(45)		PLuzC	PSBLBT	SBLBT, etc.	
	‘peste’	PMP *laqlu	*laʔlu	*lalʔu	la:ʔu

As shown in (57) below, *l is replaced by vowel length when immediately before a voiceless consonant. Were it not for the metathesis of *ʔ and *l in the example in (45), the SBLBT form would be ***laʔi*.

2.6.3 Diphthong reduction. Prior to changes in *l in the Botolan dialect of Sambal and its affiliated Ayta languages, a final diphthong *-iy (and possibly *-ay) underwent consolidation, merging in /i/. This did not occur in Sinauna, which retains final /-ay/, and it does not apply to Kapampangan, where both *-ey and *-ay merge into final /-e/.

(46)	‘rice plant’	PMP *pájey	pa:ri	SBLBL
			pa:li	SBLT, SBLBT
			pa:le	KPM
			pa:ray	SIN
	‘to die’	PMP *matey	matí	SBLBL, SBLT, SBLBT
			mate	KPM
	‘to die’	PMP *patey	patáy	SIN
	‘foot’	PMP *qaqay	ʔáyye, ʔa:yi	SBLBL
			ʔa:yi, ʔáyyi	SBLT
	‘house’	PMP *baláy	balí	SBLBL, SBLT, SBLBT ⁶
			balé	KPM
			baláy	SIN
	‘rattan’	PMP *quáy	ʔuwí	SBLBL, SBLT

Items apparently introduced at a later time are not subject to this rule:

(47)	‘finger’	PPH *garamay	garamáy	SBLBL, SIN
			galamáy	SBLT
			galamáy	SBLBT
			garamáy	SIN
	‘waterfall’	PPH *besáy	busáy	SBLBL
			buháy	SBLT
			biháy	SBLBT
	‘shame’	PLuzC *di[nŋ]ʔiy	diŋʔiy	SBLBL
			kariŋúy	SBLT
			di:ne	KPM

There is an odd reflex in one item in Bolinao, but not in the other Sambal dialects:

(48)	‘eyebrow’	PMP *kiday	ki:ri	SBLBL
			ki:luy	SBLT
			ki:liy	SBLBT
			ki:le	KPM

6. In Abellen, this rule must have operated after the loss of intervocalic *l, giving *ba:iy* ‘house’ instead of the expected ***bali*.

2.6.4 Some reflexes of *l. In the Bolinao and Tina dialects of Sambal, but not in Botolan or the Ayta languages, *l occurring immediately before *j shifts to /w/. This affects only one item in the data, *ʔaljaw ‘day, sun’:

(49)		SBLBL	SBLT	SBLBT
‘day’	PMP *qaljaw	ʔawló	ʔawló	ʔalló

PMP *l was inherited unchanged by the LUZC languages, but very early on it underwent some changes in the Botolan dialect of Sambal. Generally speaking, intervocalic *l is lost in Botolan Sambal between nonfront vowels, but there is quite a bit of variation in this depending on relative vowel height, stress (or length), and syllable coda. To account for the data at hand, a number of somewhat specific rules are needed in explaining the loss of intervocalic *l. Between two occurrences of a nonlow vowel, *l is lost:

(50)	‘deaf’	PMP *telek	SBLBT ti:ik
	‘ten’	PMP *ma-púluq	SBLBT mapúʔ

However, *l is not lost between the high central vowel and /a/:

(51)	‘big’	PPH *seláR	SBLBT hiláy
------	-------	------------	-------------

The lateral consonant is also lost between *u and either of the two other nonfront vowels:

(52)	‘moon’	PMP *bulan	SBLBT bu:wan
	‘blanket’	PMP *qules	SBLBT ʔuwih

Apparently, *l is lost between two occurrences of the low vowel only if the final consonant is a nasal *n:

(53)	‘path’	PMP *zálan	SBLBT da:an
------	--------	------------	-------------

Otherwise the *l is reflected as /l/.

(54)	‘palm’	PMP *palaj	SBLBT pa:lal
	‘fence’	PMP *qálad	SBLBT ʔa:lal
	‘man’	PMP *la-láki	SBLBT lala:ki

Those lexemes that have *l between the low and high central vowels, lose the *l or reflect it as a glide or a glottal stop:

(55)	‘forehead’	PPH *waleŋ	SBLBT wawín
	‘taro’	PMP *tales	SBLBT ta:ʔi

SBLBT loses *l between /a/ and /u/ unless the latter is stressed and word-final.

(56)	‘palm’	PMP *dalukap	SBLBT dawu:kap ~ dáwkap
	‘widow’	PMP *bálu	SBLBT ba:wu

But cf. PMP *walú ‘eight’ > SBLBT walú.

In Botolan, *l is usually replaced by vowel length when immediately before a voiceless consonant. This happens, as well, in the Ayta languages, but only sporadically.

(57)			SBLBT		
‘left’	PLUZC *ʔulki	ʔu:ki	ʔu:kiʔ	AMB	
			ʔulki	CBG	
‘testicles’	PPH *paltak	pa:tak	pa:tak	VIL	
			paltak	VIL, KIL, ABE	

‘sour’	PMP *qalsim	maʔa:him	maʔa:him	VIL, ABE, AMB, MBK
‘pestle’	PMP *laqlu	la:ʔu	la:ʔu	VIL, ABE, MIN, AMB

When followed by a voiced consonant, *l is reflected as /l/:

(58)			SBLBT
‘chicken louse’	PPH *(k)ulmug		kulmúg
‘day’	PMP *qaljaw		ʔalló
‘other’	PPH *qal(i)wa		ʔalwá

2.6.5 *r. This liquid is reflected as /r/ in Sambal of Bolinao, but as /l/ in the rest of the Sambalic languages.

(59)		SBLBL	SBLBT
‘to play’	PPH *garaw	—	galáw
‘rotten’	PPH *buruk	burúk	bulúk
‘finger’	PPH *garamay	garamáy	galamáy
‘to rub’	PPH *haprus	ʔaprus	—

2.6.6 *d. PMP *d is reflected intervocally as a liquid. In the case of Bolinao Sambal, *d is reflected as /r/ (as it is in Sinauna), and in the remaining Sambal dialects it is reflected as /l/ (as it is in Kapampangan).

In word-initial position, PMP *d is reflected as a liquid:

(60)		SBLBL	SBLT	SBLBT
‘two’	PMP *duwa	ruwá	luwá	luwá
‘water’	PMP *danúm	ranúm	lanúm	lanúm
‘hear’	PAN *deNÉR	riŋʔín	luŋʔún	liŋíʔ
‘arrive’	PMP *daten	ratíʔ	latúʔ	latiŋ
‘wall’	PMP *diŋdiŋ	riŋriŋ	liŋliŋ	liŋliŋ

This shift is blocked if the next consonant in the word is a liquid: *l, *R, or *r:

(61)		SBLBL	SBLT	SBLBT
‘tongue’	PMP *dilaq	di:laʔ	di:laʔ	di:laʔ
‘blood’	PMP *daRaŋ	da:yaʔ	da:yaʔ	da:yaʔ
‘thorn’	PMP *dúRi	duwí	du:wí	du:wí
‘fast’	PPH *daras	darás	—	—

Items entering the Sambalic vocabulary more recently, either through innovative coinage or through diffusion, are not subject to this rule.

In word-final position, *d shows two reflexes in each Sambal dialect. One variant occurs when *d is immediately preceded by the high front vowel /i/, and also when preceded by a stressed nonlow vowel, *é or *ú:

(62)		SBLBL	SBLT	SBLBT
‘narrow’	PPH *kítid	—	ki:til	ki:til
‘rope’	PPH *lúbid	ru:bir	(yu:bil)	(yu:bil)
‘to tie’	PPH *taked	takír	takúl	takíl
‘escort’	PMP *ha(n)ted	ʔatír	ʔatúl	ʔatíl
‘outside’	PPH *likud	—	likúl	likúl

The second variant of word-final *d is /d/ in SBLBL and SBLBT, and /r/ in SBLT. It is found when immediately preceded by the low vowel /a/ or by an unstressed *e or *u:

(63)		SBLBL	SBLT	SBLBT
‘to pay’	PMP *bayad	ba:yad	ba:yar	ba:yad
‘to fly’	PPH *l-um-epad	lumpád	lumpár	lumpád
‘wide’	PPH *la(m)pad	—	mala:par	mala:pad
‘ankle’	PPH *ti(ŋ)ked	—	—	ti:kid
‘nipa’	PPH *pawed	—	pa:wur	pa:wid
‘paddle’	PPH *gaʔud	ga:ʔud	—	—

The only two remaining environments that are permitted by the canonical rules of Sambalic languages are in consonant clusters within a word, *dC and *Cd. If the word consists of a reduplicated monosyllable, the *d follows the language-specific rules for initial *d and, in one attested case at least, for word-final *d:

(64)		SBLBL	SBLT	SBLBT
‘wall’	PMP *diŋdiŋ	riŋriŋ	liŋliŋ	liŋliŋ
‘shore’	PPH *digdig	rigrig	liglig	—
‘warm by fire’	PMP *daŋdaŋ	raŋraŋ	laŋlaŋ	laŋlaŋ
‘to tie’	PMP *bedbed	birbir	—	biłbił

If the cluster is of the type *Cd and it is not a reduplicated monosyllable, the *d is still reflected as a liquid. PSBL *tikdi(l) ‘to sit’ gives SBLBL *tumikriʔ* and SBLT *tumukliʔ* ‘to sit’ and MBK *tuklóʔ*, *tiklíʔ* ‘to live, dwell.’

There are no widespread examples in the data of words containing *dC other than those of the reduplicated monosyllables. The only other type of cluster of the *Cd variety are those wherein the *d is preceded by a nasal. In these cases, *d is reflected as /d/:

(65)		SBLBL	SBLT	SBLBT
‘to know (remember)’	PPH *taNda	tandáʔ	tandáʔ	tandáʔ
‘mother’	PMP *hinduh	—	ʔinduʔ	ʔinduʔ
‘short (person)’	PMP *pandak	—	pandák	pandák

The Ayta languages—Mag-antsi, Mag-indi, Ambala, and Magbukun—follow the same pattern for the reflexes of word-initial *d as the Botolan dialect of Sambal:

(66)		MAN	MIN	AMB	MBK
‘many’	PPH *dakel	lakíʔ	lakíʔ	lakíʔ	lakúʔ
‘water’	PMP *danum	(laním)	(laním)	lanúm	lanúm

as opposed to:

(67)		MAN	MIN	AMB	MBK
‘tongue’	PMP *dilaq	di:laʔ	di:laʔ	di:laʔ	di:laʔ
‘blood’	PMP *daRaq	da:yaʔ	da:yaʔ	ja:yaʔ	ja:yaʔ

In final position, the Ayta languages all have two reflexes of final *d. The data do not provide examples of final *d in all environments. For the data that are available, the rules that account for final *d are *d > /d/ after a low vowel in Mag-antsi, Mag-indi, and Ambala, with /l/ occurring elsewhere. In Magbukun, it appears that *d is reflected as /l/ in final position when immediately preceded by the front vowel /i/, and as /r/ elsewhere.

(68)		MAN	MIN	AMB	MBK	
	‘rope’	PPH *lúbid	(yu:bil)	(yu:bil)	(yu:bil)	lu:bil
	‘to tie’	PPH *taked	takíl	—	—	—
	‘knee’	PMP *tuhud	tu:ʔul	tu:l	tu:l	—
	‘to pay’	PMP *bayad	ba:yad	ba:yad	ba:yad	ba:yar
	‘to fly’	PPH *lumpád	lumpád	lumpád	lumpád	lumpár
	‘wide’	PPH *la(m)pad	—	—	la:pad	la:par
	‘nipa’	PPH *pawed	—	—	—	pa:wir
	‘fence’	PPH *bakud	—	—	—	ba:kur

2.6.7 *j. PMP *j is reflected as a liquid in the Sambal dialects, /r/ in Bolinao, and /l/ in Botolan and Sambal Tina:

(69)		SBLBL	SBLT	SBLBT	
	‘worm’	PMP *qúlej	ʔu:rír	ʔu:lul	ʔu:wíl
	‘palm’	PMP *palaj	pa:rar	pa:lal	pa:lal
	‘name’	PMP *ŋajan	ŋa:ran	ŋa:lan	ŋalán
	‘charcoal’	PMP *qújiŋ	ʔu:riŋ	ʔu:liŋ	ʔu:liŋ
	‘rice’	PMP *pájey	pa:ri	pa:li	pa:li
	‘throat’	PPH *bukjaw	bukráw	buklaw	bukláv

In three of the Ayta languages, PMP *j is reflected as /l/ in all environments for which there is information:

(70)		MAN	MIN	AMB	
	‘worm’	PMP *qúlej	ʔu:wíl	—	ʔu:wíl
	‘name’	PMP *ŋajan	ŋa:lan	—	ŋa:lan
	‘charcoal’	PMP *qújiŋ	ʔu:liŋ	ʔu:liŋ	ʔu:liŋ
	‘rice’	PMP *pájey	pa:liʔ	pa:li	pa:liʔ
	‘throat’	PPH *bukjaw	—	bikláv	—
	‘day’	PMP *qaljaj	ʔallóʔ	ʔallo	ʔallo

In Magbukun, *j is reflected as /l/ in all environments except when word-final and preceded by /a/, in which case it is reflected as /r/:

(71)	‘to dry’	PPH *belaj	MBK bulár
	‘palm’	PMP *palaj	MBK pa:lar
	‘day’	PMP *qaljaj	MBK ʔawló
	‘name’	PMP *ŋajan	MBK ŋa:lan
	‘rice’	PMP *pájey	MBK pa:liʔ
	‘worm’	PMP *qúlej	MBK ʔu:lul

Note that if *j had merged with *d before the application of the rules for word-final *d, the reflex in Magbukun for ‘worm’ would be **ʔu:lur.

2.6.8 *z. At some time after intervocalic *d had shifted to a liquid, the remaining apical *z was reflected as /d/. Within a stem, /d/ < *z is then reflected as /r/ intervocalically in Sambal Tina and Botolan, while it remains /d/ in SBLBL:

(72)		SBLBL	SBLT	SBLBT
‘path’	PMP *zálan	da:lan	da:lan	da:an
‘grass’	PPH *zikut	dikút	dikút	dikút
‘rain’	PMP *quzan	—	—	ʔurán
‘sharp’	PMP *tazém	tadím	tarúm	tarím
‘lie down’	PPH *hi-Rezaʔ	midáʔ	miráʔ	miráʔ

2.6.9 *-aw. In the three major dialects of Sambal, the final diphthong *-aw is simplified to /-o/ if it is not stressed and it is preceded by a voiced obstruent.

(73)		SBLBL	SBLT	SBLBT
‘day’	PMP *qaljaw	ʔáwro	ʔáwlo	ʔálo
‘shallow’	PMP *babaw	ʔaba:bo	maba:bo	maba:bo

If unstressed and preceded by a nasal, the final *-aw is reflected in all three Sambal dialects as /-o/, but Botolan has an alternative pronunciation with final *-iw. The Ayta languages follow the Botolan pattern.

(74)	‘housefly’	PMP *laŋaw	la:ŋo	SBLBL, SBLT, SBLBT
			la:ŋiw	SBLBT, ABE, MAN, MIN, AMB, MBK

If the final vowel is stressed and/or if it is preceded by a voiceless consonant, *-aw is reflected as /-aw/.

(75)		SBLBL	SBLT	SBLBT
‘to steal’	PMP *nakaw	mana:kaw	—	mana:kaw
‘to float’	PMP *le(n)taw	lumtáw	lumtáw	lumtáw
‘soup’	PMP *sa(m)báw	sabáw	sabáw	sabáw

2.6.10 Glide insertion. The Sambalic languages insert a glottal stop to break up sequences of identical vowels, as does Sinauna. Unlike the situation in Sinauna, however, the Sambalic languages and dialects also break up clusters of nonidentical vowels, in this case using glides.

(76)	‘to sew’	PMP *tahíq	manayíʔ	SBLBL, SBLT, SBLBT
	‘woman’	PMP *ba-báhi	baba:yi	SBLBL, SBLT, SBLBT
	‘thorn’	PMP *dúRi	duwí	SBLBL
			du:wi	SBLT, SBLBT
	‘tear (n.)’	PMP *luhaq	luwáʔ	SBLBL, SBLT
			lu:waʔ	SBLT, SBLBT
	‘bad-smelling’	PMP *bahuq	maba:wuʔ	SBLBT, AMB, MBK

2.6.11 Further sound changes. Additional phonological changes are characteristic of the dialects or groups of dialects in this branch of LUZC.

2.6.11.1 Bolinao liquids. In Bolinao Sambal, we find a type of consonant harmony with regard to the liquids. If a lexeme contains both an /r/ (whether as a reflex of *d or *j) and an /l/, the /l/ shifts to /r/.

(77)		PSBLBL	BOL
‘straight’	PPH *tuq/lid	*tuʔlir	matuʔrír
‘worm’	PMP *qúlej	*ʔu:lír	ʔu:rír
‘palm’	PMP *palaj	*pa:lar	pa:rar
‘to push’	PPH *tulud	*tulur	manjiturúr

‘rope’	PPH *lubid	*lu:bid	ru:bir
‘deep’	PMP *dálem	*ʔara:lim	ʔara:rim

2.6.11.2 *s. In Sambal Tina, an ancestral *s is reflected as /h/ in all communities except for the northernmost town of Santa Cruz. This is also the case in all communities where Botolan Sambal is spoken.

(78)			SCR	Other Tina	SBLBT
‘space-under-house’	PPH *siduŋ	si:luŋ	hi:luŋ	hi:luŋ	hi:luŋ
‘navel’	PMP *pusej	pu:sul	pu:hul	pu:hil	pu:hil
‘sated’	PMP *busuR	busúy	buhúy	buhúy	buhúy
‘to squeeze’	PMP *pespes	puspús	puhpúh	pihpúh	pihpúh

As for the Ayta languages, those most closely affiliated with Botolan—Villar, Kakilingan, Abellen, and Mag-antsi—have also undergone this change completely. In Mag-indi, *s is usually reflected as /s/, but occasionally it has become /h/.

(79)	‘to suck’	PMP *sepsep	MIN sipsíp ~ hiphíp
	‘sand’	PMP *badas	MIN balás ~ baláh
	‘spouse’	PMP *qasáwa	MIN ʔasa:wa ~ ʔaha:wa?

Such is also the case in Ambala and Magbukun. This appears to be a sound change in the process of occurring. In a single community, and perhaps even in the speech of a single individual, both sounds are heard.

2.6.11.3 *e. While the Sambal dialects of Bolinao and of Botolan consistently retain the PLUZC high central vowel, Sambal Tina always reflects it as /u/.

(80)			SBLBL	SBLT	SBLBT
‘deaf’	PMP *telek	t̪i:lik	tu:luk	t̪i:ik	t̪i:ik
‘monkey’	PPH *bákes	ba:kiʔ	ba:kuʔ	ba:kiʔ	ba:kiʔ
‘forehead’	PPH *waleŋ	walíŋ	walúŋ	wawíŋ	wawíŋ

The Ayta languages again generally reflect the vowel shift characteristic of Botolan, except for Magbukun. While the latter does permit the occurrence of /i/, the overwhelming number of items with a reflex of *i show the high back vowel instead. This is reflected in the various renderings of the language name itself, Magbeken, Magbukun, Magbukon.

2.6.11.4 Assibilation. A phonetic characteristic that is found in the Ayta communities, and not shared with any of the Sambal dialects, is the assibilation of syllable-initial alveolar stops. The occurrence of [ts] instead of [t] before a front vowel is found in Mag-indi and Mag-antsi.

(81)			MIN	MAN
‘worm’	PMP *bulati	buwa:tiʔ ~ buwatsi	—	—
‘whatchamacallit’	PLUZC (*ʔantiʔ)	ʔantiʔ ~ ʔantsiʔ	ʔantsiʔ	ʔantsiʔ
‘white’	PMP *putiʔ	putiʔ ~ putsi	putsiʔ	putsiʔ
‘penis’	PMP *qu:tin	ʔu:tin	ʔu:tsin	ʔu:tsin

The occurrence appears to be sporadic in Mag-indi, and more complete data would be needed to formulate a general rule to account for it.

Similarly defying a pattern is the alternation of [d] with [j] and [z] in Ambala and Magbukun. Ambala exhibits variation between [d] and [j] (and occasionally [z]) most regularly in word-initial position but occasionally within a word as well:

(82)	‘tongue’	PMP *dɪlaq	AMB ʃi:laʔ ~ di:laʔ
	‘blood’	PMP *daRaʔ	AMB ʃa:yaʔ ~ da:yaʔ
	‘far’	PMP *zayʊq	AMB maʃayúʔ ~ ma:dayúʔ
	‘long’	PPH *kazaŋ	AMB makadánj ~ makajaŋ

Magbukun demonstrates this variation also:

(83)	‘tongue’	PMP *dɪlaq	MBK ʃi:laʔ ~ di:laʔ
	‘blood’	PMP *daRaʔ	MBK ʃa:yaʔ
	‘far’	PMP *zayʊq	MBK maʃayúʔ ~ madayúʔ ~ maza:yuʔ
	‘long’	PPH *kazaŋ	MBK maka:zaŋ ~ maka:jaŋ

Suffice it to note that this variation does occur in some of the Ayta languages, and that within the Central Luzon microgroup this characteristic is limited to them.

3. PRONOUNS. All of the LUZC languages have pronoun systems that include eight semantic slots. These slots are common to many, but not all, languages in the Philippines, since a number do not have a distinctive term for the 1+2SG (or dual) slot different from the 1+2PL form (Liao 2008). The pronoun set dealt with here is one of the prominent forms common to Philippine languages, the long nominative form. Proto-Central Luzon inherited a set of pronoun stems from higher levels (PMP or whatever) that includes the following:

(84)	1SG	*aku	1PL	*i-kami
	2SG	*ika	2PL	*i-kamu
	1+2SG	*ita, *i-kita	1+2PL	*i-tamu
	3SG	*ia, *siya	3PL	*ida, *sida

3.1 KAPAMPANGAN LONG NOMINATIVE PRONOUNS. With these as the inherited PMP forms, all the long nominative pronouns begin with the sequence *i- except for the 1SG *aku. Kapampangan has the set in (85):

(85)	PLUZC	KPM	PLUZC	KPM
	1SG	*ʔaku	1PL	*ʔi-kami
	2SG	*ʔika	2PL	*ʔi-kamu
	1+2SG	*ʔita, *ʔi-kita	1+2PL	*ʔi-tamu
	3SG	*ʔia, *siya	3PL	*ʔida, *sida
				ikami
				ikayu
				ita:mu
				ila

The unexpected occurrence of /a/ rather than /i/ in the 1+2SG may have occurred under the influence of the 2SG item, or it may have occurred by analogy with the other three-syllable forms, all of which have /a/ as the second vowel. Alternatively, it may be a combination of the 2SG pronoun *ʔika + the 1+2SG possessive suffix *-ta. Reid (1979) attributes the Kapampangan (and Tagalog) 2PL item to diffusion from a Cordilleran language.

3.2 THE REMAINING LUZC LANGUAGES. Aside from KPM, all of the other LUZC languages adopted the case-marking particle *si- to the bases inherited from

their immediate ancestor. This feature is shared with the Northern Philippine (Cordilleran), Manobo, and other groups. In LUZC it provides the following set:

(86)	1SG	*si-ʔaku		1PL	*si-ʔi-kami
	2SG	*si-ʔika		2PL	*si-ʔi-kamu
	1+2SG	*si-ʔita, *si-ʔikita		1+2PL	*si-ʔitamumu
	3SG	*si-ʔiya, siya		3PL	*si-ʔida, *sida

For the sake of convenience, this level of development will be referred to as Proto-Sambal-Sinauna (PSBLSIN). Characteristic of this set of languages is the shortening of at least some of the reconstructed pronoun lexemes.

3.2.1 Sinauna long nominative pronouns. Sinauna presents us with the following set:

(87)		PSBLSIN	SIN		PSBLSIN	SIN
	1SG	*si-ʔaku	saku		1PL	*si-ʔi-kami siʔkami
	2SG	*si-ʔika	siʔka, siʔika		2PL	*si-ʔi-kamu siʔkamu
	1+2SG	*si-ʔita	si-ʔita-daw, siʔdarʔa		1+2PL	*siʔitamumu siʔtamumu
	3SG	*siya	siya		3PL	*si-ʔida, *sida siʔra, sira

The 3SG form and one form of the 3PL item were inherited unchanged. The 1SG item has undergone simplification by losing the second and third segments of the word. All of the plural forms, one variant of the 2SG form, and one of the 1+2SG items have lost the first vowel of the stem. With regard to the 1+2SG items, *-daw* may be a variant of some ancestral form, such as PAN *Dewha ‘two’. The construction of a 1+2SG item from the word for ‘two’, usually together with *-ta*, occurs elsewhere, as in Botolan *luwa:ta*, and also in Northern Bontok *dowa:ta* and Aborlan Tagbanwa *duwata*.

3.2.2 Sambalic long nominative pronouns. The dialects of Sambal have also shortened some of the ancestral pronoun forms. SBLBL has the following set:

(88)		PSBLSIN	SBLBL		PSBLSIN	BOL
	1SG	*si-ʔaku	siʔku		1PL	*si-ʔi-kami siʔkami
	2SG	*si-ʔika	siʔka		2PL	*si-ʔi-kamu siʔkamu
	1+2SG	*si-ʔita	siʔta		1+2PL	*siʔitamumu siʔtamumu
	3SG	*siya	siya		3PL	*si-ʔida, *sida sira, sara

Here we see that the first vowel of the stem has been lost in all lexemes other than the 3SG and 3PL items, which were inherited unchanged. The seemingly aberrant 3PL item *sara* may have been formed by analogy with the other plural items that have /a/ as the first vowel of the stem.

Sambal Tina and Botolan show only minor differences from Bolinao. One feature is the loss of prenasal glottal stop. While this is a regular phonological process in Botolan, it occurs in Tina only in the pronouns.

(89)		PSBLSIN	SBLT	SBLBT		PSBLSIN	SBLT	SBLBT
	1SG	*si-ʔaku	hiku	hiku		1PL	*si-ʔi-kami	hikami hikayi
	2SG	*si-ʔika	hika	hika		2PL	*si-ʔi-kamu	hikamu hikawu
	1+2SG	*si-ʔita	hita	hita		1+2PL	*siʔitamumu	hitamumu hita:mu
	3SG	*siya	hiya	hiya		3PL	*si-ʔida, *sida	hila hila

The one noticeable way in which the Botolan pronouns differ from those in Tina is the loss of the intervocalic nasal in the 1PL and 2PL items (but not the 1+2PL form), with the addition of a compensatory glide.

3.2.3 Ayta long nominative pronouns. Reliable pronoun data are available for Mag-antsi, Magbukun, and Ambala. The pronouns in these three languages show some notable similarities to, and some differences from, the pronouns in Botolan Sambal. In addition to the regular phonological processes of change from *s and *d to /h/ and /l/, respectively, they also show loss of the preconsonantal glottal stop. They demonstrate, as well, the loss of the intervocalic nasal in the 2PL form. Mag-antsi and Magbukun, but not Ambala, also lose the nasal in the 1PL item. All three languages employ the longer of the two forms for the 1+2SG slot, *siʔikita rather than *siʔita. Further, these languages have innovated a longer form for the 1+2PL slot: *siʔikitamu.

(90)	PSBLSIN	MAN	MBK	AMB
1SG	*si-ʔaku	hiku	haku	haku
2SG	*si-ʔika	hika	hika	haka
1+2SG	*si-ʔita, *si-ʔikita	hita, hikita	hikita	hakita
3SG	*siya	hiya	hiya	hiya
1PL	*si-ʔi-kami	hikayi, hikay	hikayi	hakami
2PL	*si-ʔi-kamu	hikawu, hikaw	hikawu	hakawu
1+2PL	*siʔitam	hikitamu	hikatamu	hakitamu
3PL	*si-ʔida, *sida	hila	hila	hila

The pronouns in Mag-antsi are perhaps the most regular of the group. Those in Magbukun show only two minor differences, both involving vowels. Both Magbukun and Ambala have *haku* rather than the expected ***hiku* for the 1SG slot, as occurs in Sinauna. Magbukun *hikatamu*, rather than ***hikitamu*, shows the same sort of change that Kapampangan demonstrates in regularizing the second vowel of the plural pronouns. Ambala takes this process even further by shifting all of the first vowels of the nonthird person pronouns to /a/.

3.3 SUMMARY OF LUZC PRONOUN SYSTEMS. The inheritance of the long nominative pronoun bases with the prefix *ʔi- is not limited to the Central Luzon group of languages. It is, nevertheless, a feature that they all share. The incorporation of this prefix into a new stem is also found outside of this group. The addition of another case-marking prefix *si- is a characteristic of Sinauna, Sambal, and the Ayta languages, as it is in other Philippine microgroups. Sinauna differs from the Sambalic languages in some minor features of the pronoun system, such as the retention of the ʔi-base unchanged in several of the long nominative forms. Mag-antsi, Magbukun, and Ambala differ from the three dialects of Sambal in their use of the longer stem *siʔikita for the 1+2SG semantic slot and the innovation of a longer form for the 1+2PL slot, *siʔikitamu. On the basis of these data, it appears that Sinauna is somewhat closer to the Sambalic branch of the LUZC group than it is to Kapampangan.

4. LEXICAL RECONSTRUCTION. The LuzC languages under consideration here demonstrate some sharing of lexical innovations. Naturally, each of them has a

large number of innovative items not shared with any other languages; by and large, these “uniques” are not treated here.

A lexical item with reflexes in at least two of the three established branches of the microgroup—Kapampangan, Sambal, and Sinauna—is attributed to the level of Proto-Central Luzon. Lexically, Botolan and the Ayta languages share a large vocabulary not shared with the Bolinao and Tina dialects nor, of course, with Kapampangan and Sinauna. This situation is as we would expect if the Ayta populations acquired their version of Austronesian from contact with the people of Botolan, with some communities experiencing more sustained interaction than others (Reid 1987, 1994, 2010).

Naturally, we cannot assume that any of the lexical diffusion that led Botolan and Ayta to share vocabulary to the exclusion of other languages and dialects was unidirectional. Further, lexical items that are shared among Ayta communities could be reflexes of some precontact language, or they could be products of later innovation and subsequent diffusion from one Ayta community to another.

With all of this in mind, lexical data are presented here in categories (with varying degrees of confidence) reflecting level of reconstruction and community sharing. A lexeme that appears to be inherited, as opposed to borrowed, in at least two of the Sambal dialects is considered to be from Proto-Sambalic. Within Sambal, a Proto-Bolinao lexeme is represented in both Bolinao and Anda municipalities. An item found in at least two of the municipalities of Sta. Cruz, Candelaria, Masinloc, Palauig, and Iba is assigned to Proto-Tina.

If a lexical item is found in Botolan and Cabangan, or either of these and one similar Ayta community—Villar, Kakilingan, or Abellen—it is considered to be inherited from Proto-Botolan. A separate list presents lexical items shared by Botolan (and/or Cabangan) and other Ayta languages. Lexemes that are shared by two or more Ayta languages are presented separately, keeping in mind that they may be inherited from a common ancestor or they may be shared because of diffusion. Finally, lists of lexemes assignable to the specific Ayta speech communities are presented.

4.1 PLUZC. Those items that may be ascribed to Proto-Central Luzon with some degree of confidence include those that follow.

4.1.1 Phonological/morphological innovations

- (91) *ʔiRbun ‘egg’; cf. PAN *Rebun ‘egg’
 KPM e:bun; SBLT ʔúybun
 *ʔikit ‘to see’; cf. PPH *ʔakit ‘to see’
 KPM ikit ‘saw, was seen’; SBLBT, MBK maʔikit ‘to see’

4.1.2 Semantic innovations

- (92) *ʔu:ɲut ‘coconut’; cf. PPH *huɲut ‘coconut shell dish’
 KPM u:ɲut; SBLBL, SBLT, SBLBT, MIN ʔu:ɲut ‘(ripe) coconut’
 *gawaʔ ‘swidden; to prepare a field’; cf. PPH *gawaʔ] ‘to make’
 KPM gawa ‘cut grass to cultivate a field’; SBLBL gagawá ‘field’; MIN ginawaʔ ‘swidden’
 *kuskus ‘to wipe’; cf. PAN *kuskus ‘to scrape’
 KPM kuskus; SIN kuskusan; MIN kuskusan ‘to wipe’

- *wakáR ‘rattan’; cf. PMP *wakaR ‘root’; cf. also Bontok, Kankanaey
wakal, waka ‘vine’
 SIN waká; MIN, AMB, MBK wakáy ‘rattan’

4.1.3 Lexical innovations

- (93) *ʔili KPM ili, ayli; SBLBT kaʔi:li; MAN maŋkaʔi:liʔ; MIN
 maŋkaʔiliʔ; AMB maŋkáyliʔ; MBK magkakayli ‘to laugh’
 *gana KPM keganagana, eganáganá; SBLBT, MAN kaganawán ‘all’
 *gu:[djr]ut KPM gu:lut; SBLBL gu:rut ‘back (of a person)’
 *kimul KPM akmul ‘to swallow’, KPM akmu:lan ‘throat’; MBK
 ʔiku:mul ‘to swallow’
 *(l)usuŋ MIN ʔoson; SIN lusuŋ, lusuŋ ʔid tu.ʔid ‘tree’
 *pitís KPM yaptas; VIL pithin; MAN ʔipítih ‘to tie (bundle)’
 *s[au]la:paw KPM sula:po; SIN sarápaw ‘to fly’
 *saluʔ KPM sa:lu ‘chest; VIL, MAN, AMB halóʔ; MIN salóʔ ‘throat’
 *samát KPM samat; SBLBL samát; SBLT, SBLBT; MAN, AMB
 hamát; MIN, MBK samat ‘betel leaf’. This item has
 diffused into Tagalog.
 *ta[djr]an KPM talnan; tatalanan; SBLBT talanin; MIN talanin, talnin
 ‘to hold’
 *takáʔ KPM pitakan ‘anus’; SBLT, SBLBT, MAN, MBK takáʔ
 ‘excrement’
 *tulaʔ KPM tulaʔ ‘happiness’; MIN magtulaʔ ‘to play’

4.2 SINAUNA DEVELOPMENTS. There are a number of items that occur in Sinauna (and Remontado) that appear to be limited to this branch of the Central Luzon microgroup. As with the LUZC list above, this one is divided into phonological/morphological, semantic, and lexical innovations. A possible pre-Sinauna form is given. Following is a sample of such items.

4.2.1 Phonological/morphological innovations

- (94) *bakis bákis ‘old (person)’; cf. PPH *bakes ‘old (woman)’
 *darʔa darʔá ‘two’; cf. PPH *da-duwa ‘two’
 *laʔwa láʔwa ‘not’; cf. PMP *wada ‘exist, not exist’

4.2.2 Semantic innovations

- (95) *ʔu:tik ʔu:tak ‘head’; cf. PMP *hutek ‘brain’
 *mamaʔ mamaʔ ‘to eat’; cf. PAN *mamaq ‘to chew betel’
 *saluŋ salúnj ‘to live, dwell’; cf. Tagalog *salóng* ‘low-roofed
 hut’, Pangasinan salúnj ‘going to another place’
 *sukub su:kub ‘body’; cf. KPM *sukub* ‘breast of animal’
 *tibanluʔ tibanlúʔ ‘torch’; cf. Isnag *sibálo* and Casiguran
 Dumagat *tibalo* ‘to cook food in a bamboo tube’

4.2.3 Lexical innovations

- (96) *ʔap[ai]d ʔapád ‘near’
 *ʔigan ʔi:gan ‘stone’ But cf. Northern Kankanaey ʔi:gan
 ‘small sharp pieces of stone or gravel’

*ʔi[lr]am	makaʔírám	‘cold’
*ʔiriŋ	ʔumi:riŋ	‘to wait’
*ʔumid	ʔúmid	‘below’
*bulid	ʔibulíd	‘to push’
*da:nut	dánut	‘to smell’
*dilap	diláp	‘lightning’
*gataw	gunnátaw	‘to swim’
*landap	lándap	‘to hear’
*la:pad	lumápad	‘to sit’
*raŋbun	raŋbun	‘many’
*raʔtun-an	raʔtúnan	‘bamboo water container’
*rayra	ráyra	‘mother’
*sabud	sabúd	‘mountain, forest’
*sulʔaŋ	miʔsulʔáŋ	‘to see’
*taʔan	taʔán	‘neck’
*tabiʔu	tabiʔu	‘deep’
*tariʔ	miʔtáriʔ	‘to fear’
*tuʔus	tu:ʔus	‘thigh’

4.3 KAPAMPANGAN DEVELOPMENTS. Kapampangan has a number of items that show some sort of innovation not shared by other languages, either in LUZC or elsewhere. Presented below is a set of items found in the basic vocabulary, and it is not intended to be exhaustive. The left-hand column contains a possible pre-Kapampangan form for ease of comparison with other languages and sources.

4.3.1 Phonological/morphological innovations

(97) *ayun	ayún	‘earthquake’; cf. PPH *dayun ‘earthquake’ and PMP *hayun ‘swing, shake’
*imiR	imiʔ	‘urine’; cf. PMP *miqmiq ‘to urinate’
*dagul	dagúl, maragul	‘big’; cf. PPH *dakel, dakul ‘big’
*ku:lut	ku:lut	‘back’; cf. PLUZC *gu:[dj]ut ‘back’
*ŋu:ŋut	ŋu:ŋut	‘coconut’; cf. PLUZC *ʔuŋut ‘coconut’
*paslu	paslu	‘bow, arrow’; cf. PPH *palsu(k) ‘bow-and-arrow’
*pusit	pusít	‘to squeeze’; cf. PMP *su(m)pit ‘tongs’
*siyas	masyas	‘hard’; cf. PAN *teRas ‘hard’
*takday	takde	‘arm’; cf. PPH *taklay ‘arm’
*tariziʔ	tali:riʔ	‘finger, toe’; cf. PAN *dariziʔ ‘finger’
*tudtud	tudtud	‘to sleep’; cf. PAN *tuDuR ‘to sleep’
*yumaʔ	yu:maʔ	‘old (object)’; cf. PMP *lumaq ‘weak, tired; worn out’

4.3.2 Semantic innovations

(98) *a:yup	a:yup	‘bird’; cf. PPH *hayup ‘animal’
*banuwa	banwa	‘year’; cf. PMP *banua ‘inhabited land’
*damdam	damdam	‘to hear’; cf. Tagalog <i>damdam</i> ‘feeling, sensing’
*ma-hituŋ	me:tuŋ	‘one’; from *ma- + PAN *hituŋ ‘count’
*pagaw	pago	‘shoulder’; cf. PPH *p[ae]:gaw ‘chest’
*suksuk	suksuk	‘thorn’; cf. PAN *suksuk ‘to stick in’

4.3.3 Lexical innovations

(99) *abit	abit	‘to pull’
*a:yap	ma:yap	‘good’
*i:gu	i:gu	‘to pull’
*ubiŋ-an	ubiŋan	‘snake’
*ugsay	ugse	‘to throw away’
*ba-bagwa?	babagwa?	‘spider’
*b/al/ugbug	balugbug	‘ear’
*ba:tal	ba:tal	‘neck’
*da:pu	da:pu	‘crocodile’
*du:tun	du:tun	‘tree’
*gandus	gandús	‘taro’
*gulyut	gulyut	‘to pull’
*ki:nis	kumi:nis	‘to bite’
*kitig	kitig	‘ant’
*lagyu	lagyú	‘name’. This term has diffused into MIN and AMB.
*lamlam	malamlam	‘slow’
*lati	malati	‘small’
*palaypay	pale:pe	‘path’
*pa:liy ~ *pa:liw	mapa:li?	‘hot’
*pisaŋ	apsaŋ	‘chicken louse’. This item has diffused into MIN and AMB.
*sagu	sagu	‘horn (of animal)’
*tikdaw	tikdo	‘to stand’
*tusug	tusúg	‘to push’

4.4 SAMBALIC DEVELOPMENTS

4.4.1 PSBL. As was mentioned above, some innovative developments appear to be widespread within the Sambalic branch of LUZC and are, therefore, ascribed to the level of Proto-Sambalic.

4.4.1.1 Phonological/morphological innovations

- (100) *ʔaliŋsa:ŋaw ‘to stink’; cf. PPH *qaliŋsaw ‘to stink’
SBLT paŋalinha:ŋaw, VIL ʔaliŋha:ŋaw, ABE ʔalinha:ŋaw
‘to stink’
- *ʔa:nag ‘termite’; cf. PAN *ʔánay ‘termite’; cf. also Northern
Kankanaey *ʔanig* ‘termite’
SBLBL, SBLT, SBLBT, MAN, MIN, AMB, MBK ʔa:nag
‘termite’
- *ʔi(N)san ‘later’; cf. PPH *miNsan ‘[+ time]’
SBLBL maʔinsan, SBLT (SCR) paysán, páʔisan,
SBLBT paʔiŋhan, papaʔiŋhán ‘later’
- *ʔumi? ‘urine (n.)’; cf. PAN *iSiq via *ʔ-um-ihiq ‘to urinate’ >
*ʔ-um-ihí? > *ʔ-um-ii? > ʔumi?
SBLBL, SBLT ʔumi? ‘urine’

*bibiR	'mouth'; cf. PPH *bibiR 'mouth'; cf. also PAN *bíbíR 'lip' SBLBL, SBLBT, MAN, MÍN, AMB bibíy, SBLT, MBK bubúy 'mouth'
*buʔid	'heel'; cf. PPH *buq[eu]l 'heel' SBLBL bu:ʔid, SBLT buʔur 'heel'
*lawaw	'saliva'; cf. PPH *la:way 'saliva' SBLBL, SBLT lawáw 'saliva'
*tuʔzik	'housepost'; cf. PMP *tezek 'erect, upright' SBLBL tuʔdik, SBLT (SCR) tuʔduk, SBLT tuʔrúk 'housepost'
*yu:bid	'rope'; cf. PPH *lubid 'cord' SBLT, SBLBT, MAN, MÍN yu:bil 'rope'

4.4.1.2 Semantic innovations

(101) *ʔaba:gat	'rain'; cf. PAN *habagat 'south, west wind' SBLT, AMB, MBK ʔaba:gat 'rain'
*ʔapay	'mat'; cf. PPH *hapaR 'to spread leaves' SBLBL, SBLT ʔapáy '(sleeping) mat'; SBLBL maʔapáy 'to weave (mat)'
*babáʔ	'west'; cf. PMP *babaq 'low, beneath' SBLT libabáʔ, SBLBT, AMB babáʔ 'west'
*ba:kir	'mountain'; cf. PPH *ba:kir 'forest' SBLT, SBLBT, MAN ba:kil 'mountain'
*bari:taʔ	'to say'; cf. PAN *bari(Ct)ʔa 'news' SBLBL ʔibari:taʔ, SBLT, SBLBT ʔibali:taʔ 'to say'
*buʔlun	'to swallow'; cf. PWMP *buqul 'stopped in its course, as something that sticks in the throat' SBLT ʔibuʔlun, MAN, MÍN, AMB ʔibu:ʔun 'to swallow'
*laguʔ	'young'; cf. PPH *laguʔ 'vigorous (of vegetation)' SBLBL, SBLT, SBLBT malagúʔ 'young (especially coconut)'
*sabut	'feather, hair'; cf. PAN *sabut 'husk', PPH *sabut 'pubic hair' SBLBL, SBLT (SCR) sabut, SBLT, SBLBT, MAN, MÍN habút 'feather, hair'
*sikin	'old man'; cf. PPH *siken 'strong' SBLBL masikín, SBLT masakún 'old man'
*ya:but	'cogon (<i>Imperata cylindrica</i>)'; cf. PAN *RabuC 'to pull up by the roots' SBLT, SBLBT ya:but 'cogon'

4.4.1.3 Lexical innovations

(102) *ʔakuláw	SBLBL, SBLT ʔakuláw 'old woman'
*ʔambay	SBLT, SBLBT ʔambáy 'shore'; SBLT ʔambáy 'sea'
*ʔantíd	SBLBL ʔantíd, SBLT maʔantur 'short (object)'
*ʔilis	SBLBL kaʔlís, SBLT (SCR) ʔulís, SBLT ʔumlíh 'to laugh'
*ʔimit	SBLBL maʔmít, SBLT maʔmút 'to lose (something)'
*ʔibut	SBLBL, SBLT ʔibút 'rat'

*ʔi:kap	SBLBL, SBLT ʔi:kap ‘to play’; cf. Tagalog <i>hikap</i> ‘gallivanting around’
*ʔin	SBLBL si:ʔin; SBLBT habaʔin, baʔin, haʔin ‘that (near addressee)’; SBLT ʔinin, SBLBT ʔinʔin ‘whatchamacallit’
*ʔi-sin	SBLBL ʔisín, SBLT ʔisún, SBLBT bahín, bayhín ‘there (near addressee)’
*ʔu[dʒr]andis	SBLBL ʔurandis, SBLT (SCR), ʔulándis, SBLT ʔulándi ‘ant’
*ʔuzit	SBLBL múʔdit, SBLBT, MÍN maʔurít, MAN naʔurít, MÍN naʔudít, AMB maʔudít, AMB mawujít, MBK mawzit ‘red’; MBK maʔuzít ‘yellow’
*ʔVmút	SBLBL, SBLT maʔmút, SBLT, SBLBT, MAN maʔamút, SBLBT, MÍN, AMB, MBK maʔumút ‘hot’
*bu:y[au]t	SBLT (SCR) mabu:yat, SBLT, SBLBT mabu:yut ‘long (time)’
*dalu:nut	SBLBL mandalu:nut, SBLBT dau:nut ‘smooth’
*diblim	SBLBL madiblim, SBLT marublúm, madublúm ‘dark’
*dilap	SBLBL, SBLT diláp ‘tomorrow’
*gaʔgít	SBLBL gaʔgít, SBLT gaʔgut ‘to hold’
*gasák	SBLT, SBLBT, AMB, MBK gahák, MÍN, AMB, MBK gasák ‘swidden’; cf. Tagalog <i>gasák</i> ‘pruned, cut at the top (said of grasses, plants, etc.)’
*gawgaw	SBLBL maŋawgáw, SBLT ʔigawgáw ‘to dry’
*gi:gaŋ	SBLBL, SBLT, SBLBT, MAN, MÍN gi:gaŋ ‘spider’. This item has diffused into Pangasinan.
*kabunlalakaw	SBLBL kabunlala:kaw, SBLT kabulala:kaw, SBLT, AMB, MBK bulala:kaw, ‘rainbow’
*ki[dʒr]íp	SBLBL kiríp, SBLT kulúp ‘chest’
*kisaw	SBLBL, MÍN maksáw, SBLT, SBLBT, MAN makháw ‘strong’
*kudpal	SBLBL makúbpal, SBLT (SCR) makudpal, SBLT makurpál, MÍN nakudpá? ‘thick’
*lamul	SBLBL lamúʔ, SBLT, SBLBT kalamú, SBLT kalamuʔán, SBLBT lamú, MAN kalamúʔ, MBK lamúl, kalamulan ‘companion’
*lasput	SBLT, SBLBT malaspút ‘to lie (untruth)’
*la:taʔ	SBLBL mila:taʔ, SBLT mala:taʔ ‘thirst’
*li:ŋiw	SBLT malu:ŋuw, SBLBT mali:ŋiw, SBLBT ʔaŋkali:ŋiw ‘lonely’
*lubay	SBLBL, SBLT, SBLBT, KIL, MAN, AMB, MBK lu:bay ‘G-string’; cf. PPH *labay ‘cord’
*maʔin	SBLBL, SBLT, SBLBT maʔin; AMB, MBK mayín ‘there is/are’
*pa:ka	SBLT pa:ka, SBLBT, MAN, AMB paka:ʔin ‘to split’
*pastaj	SBLT, SBLBT pastáj, ‘question’
*paysip	SBLBL paysip, SBLT payhupún, SBLBT payhipán ‘to blow’
*piyis	SBLT pumyúh, SBLBT napiyáh ‘to turn’
*puni:tiʔ	SBLT, SBLBT puni:tiʔ ‘to hit’. This item has diffused into Pangasinan.

*pu:ʔut	SBLT, AMB, MBK pu:ʔut ‘anger’; cf. Tagalog <i>poót</i> ‘rancor, hatred’
*rayʔip	SBLBL marayʔip, SBLT malayúp, SBLBT, AMB, MBK malayʔip ‘cold’
*ri:riʔ	SBLBL mari:riʔ; SBLT malu:luʔ, SBLBT, MAN mali:liʔ ‘sad, lonely’
*ri:ga	SBLBL mari:ga, SBLT, SBLBT mali:ga ‘happy’
*rigsa	SBLBL marigsa, SBLT kaligsawán; ‘healthy’; SBLT, SBLBT malighá ‘happy’
*saʔgili	SBLT ʔihagi:li, mihaʔgi:li, SBLBT ha:gili ‘to exchange’
*sakalaku	SBLBL sakalaku, SBLT hakalakú ‘other’
*sarba(n)	SBLBL sarba; SBLT halbán ‘all’
*sayɲit	SBLBL sayɲit, SBLT háyɲut, SBLBT, MAN hayɲit, MÍN sayɲit ‘sweat’
*siʔi[dʒr]	SBLBL si:ʔir, SBLT, SBLBT, MAN hi:ʔil ‘floor’
*s[iu]lyaw	SBLBL masilyáw, SBLT, SBLBT mahulyáw, MÍN nasilyaw ‘yellow’; MÍN holyawin ‘green’; cf. Proto-Bashiic *sulaw ‘yellow’
*taʔgan	SBLBL, SBLT manaʔgán, SBLT taʔgán ‘to wait’
*talugtug	SBLBL talugtúg ‘mountain’; MAN taúgtug ‘hill’; cf. Northern Kankanaey (Bauko) <i>patugtug</i> ‘hill’
*talunasan	SBLT (SCR), taluna:san, SBLT talunahán, SBLBT, MAN tawuna:han ‘eel’
*tapúʔ	SBLT, SBLBT, MAN, AMB, MBK tapúʔ ‘semen’
*taw	SBLBL, SBLBT ʔitáw ‘there (far), SBLBL si:taw, SBLT ʔudtaw, yatáw ‘that (far)’
*titiʔ	SBLT, SBLBT titiʔ ‘vagina’; cf. Tagalog <i>titi</i> ‘penis’
*ya:ɲat	SBLT, SBLBT, MÍN, AMB maya:ɲat ‘dirty’

4.4.2 PSBLBL. Presented here are those items found in both Bolinao and Anda municipalities and reconstructed as Proto-Bolinao Sambal.

4.4.2.1 Phonological/morphological innovations

- (103) *ʔalina ʔaliná ‘shadow’; cf. PPH *qalinaw ‘shadowy outline’
 *takliʔ takliʔ ‘excrement’. Blend of PAN *(t)aki and PPH *taklaʔ ‘excrement’

4.4.2.2 Semantic innovations

- (104) *ʔitiɲ naʔtiɲ ‘long (time)’; cf. Proto–Southern-Central Cordilleran *ʔitiɲ ‘big, old (person), to grow’
 *bijɲ bi:riɲ ‘rooster’; cf. Isnag *bixing*, Northern Kankanaey *biding* ‘k.o. bird’
 *da:pig ra:pig ‘rain’; cf. Isnag *dápag* ‘typhoon’
 *ra:but ra:but ‘grass, cogon’; cf. PAN *rabu(Ct) ‘pull out s.t. rooted in soft matter’ and PSBL *Ra:but ‘cogon’ < PAN *RabuC ‘to pull up’

4.4.2.3 Lexical innovations

- (105) *ʔad + *ti ʔadtí ‘where?’
 *bigat bi:gat ‘scar’

*bikuy	bikúy ‘seed’
*butla?	mamutlá?, butlá? ‘to split’
*da?ikli(nŋ)	dikliŋ, da?iklin ‘small’
*kasa	kasá ‘none’
*(k)irup	panrupán ‘west’
*libit	nalbit ‘wet’
*litu?in	litu:ʔin, litu:win ‘(slaked) lime’
*palna?	mapalna? ‘slow’
*pikŋa?	pikŋaʔin, mamikŋá? ‘to split’
*sa:gir	sa:gir ‘anger’
*sama?	kasamá?, samá? ‘dirty’; cf. Tagalog <i>masamá</i> ‘bad’
*sa?paŋ	sina?páŋ ‘cooked rice’
*sa:ya	sa:ya ‘one’
*tamu	tamu, ʔatamú? ‘only’; cf. PPH *mu ‘only’
*tu:pa?	tu:pa? ‘mud’
*yu:pa?	yu:pa? ‘below’; ʔayu:pa? ‘short (person), low’

4.4.3 PSBLT. Listed here are those items that occur in three or more Sambal Tina municipalities or at least in two that are not contiguous.

4.4.3.1 Phonological/morphological innovation.

(106) *dutak duták ‘mud’; cf. PWMP *lutak ‘mud’

4.4.3.2 Semantic innovation

(107) *duka? duká? ‘don’t’; cf. PPH *duk[ae]s ‘bad’

4.4.3.3 Lexical innovations

(108) *ʔay + *ti	ʔaytí ‘where?’
*bilbi	bilbí, bilbí? ‘to know (someone)’. This could be from *bilbiR.
*buwat	buwát ‘field’; cf. Itawis <i>hiwat</i> ‘to pick fruit from a tree’
*da?u:tu?	da?u:tu, dara?u:tu?, da?da?u:tu, da?u:tu? ‘few’
*din	yaʔrin, yarín ‘that (near addressee)’
*duna[ʔt]	maruná?, marunát, maduná? ‘soft’; cf. PAN *ruñay ‘become soft’
*kal[iu]g	makalúg ‘small’; makalúg ‘short’
*katiŋ	makatíŋ, makatí? ‘small’; cf. Casiguran Dumagat <i>kátí</i> ‘low tide; shallow’
*labas	labas labáh, nabáh ‘good’; cf. Northern Kankanaey <i>labásna</i> ‘favorable moment; opportunity’
*luput	lu:put, lupút ‘wind’
*s[iu]bs[iu]b	subsub, hubhúb ‘anger’
*si:ba?	si:ba?, hi:ba? ‘to steal’
*si?ban	masí?ban, mahí?ban, mahi?bán ‘big’
*sumin	humín ‘none’
*talabit	matalabít ‘fast’
*talakba?	talakbá? ‘frog’
*tíbya?	matíbya? ‘red’; cf. Proto–Nuclear–Central Cordilleran *tiba ‘yellow, green’
*tu?pak	tu?pák ‘wet’; cf. Bontok <i>tópak</i> ‘to flow, of blood’

4.4.4 PSBLBT. Listed below are those items found, minimally, in Botolan (BOT) and Cabangán (CBG). Included are items that are also shared with Villar (VIL), Kakilingan (KIL), and/or Abellen (ABE), dialects spoken by Ayta peoples whose dialects are very close to that of Botolan.

4.4.4.1 Phonological/morphological innovations

- (109) *bisu BOT mamabhó, ABE ?ipabhó, VIL pabhu:win ‘to boil’; cf. PMP *sebu ‘sudden meeting of water and heat’
 *kudpaw BOT, ABE makudpáw, SBLBT makugpáw, CBG makidpaw, VIL makubpáw ‘thick’; cf. PSBL *kudpal ‘thick’
 *na:win BOT, ABE na:win, VIL na:won, KIL na:in ‘our (exclusive)’; cf. PPH *namen our (exclusive)’
 *nibnib BOT, CBG, VIL, ABE nibnib ‘chest’; cf. PPH *debdeb ‘chest’

4.4.4.2 Semantic innovations

- (110) *bayumbu:ka BOT, VIL bayumbu:ka, CBG ba:yimbu:kah ‘morning’. Possibly from *baRu ‘newly’ + *buka ‘open’
 *ka-patay-?an BOT, VIL, ABE kapati:?an ‘sand’; cf. PAN *pa(CtT)aR ‘level’
 *matu?a CBG ka:matun ta:wán, ABE kamatu?an ta?uwan ‘aunt’; cf. PAN *ma(n)-tuqás ‘parent-in-law (WF); uncle (MB)’
 *si:lim BOT, CBG, VIL, KIL mahi:lim ‘afternoon’; cf. PPH *si[r]em ‘dark’

4.4.4.3 Lexical innovations

- (111) *?a:ka[djl] BOT, CBG ?a:kál ‘to hunt’
 *?añán BOT, CBG ?añán ‘thirst’
 *?a:ñaw BOT, VIL ?a:ñaw ‘to cry, weep’
 *?ugik BOT KIL ?ugik ‘frog’
 *?ugnat BOT, VIL ?ugnát ‘to pull’
 *?uligtán BOT ?uligtan, BOT, CBG, VIL, ABE ?uligtanán, KIL ?ulitán ‘to hold’
 *ba?in BOT ba?in, ABE habayín ‘that (near addressee)’
 *bu:[djr]añ BOT, CBG, VIL bu:lanj ‘ankle’
 *dimik BOT dimik, KIL nadimik, ABE marimik ‘dirty’
 *kata[k?] BOT nikaták, CBG mikaták, VIL nikatá?, ABE mikatá? ‘to lose (s.t.)’
 *limbu?iy BOT malimbi:?iy, CBG malimbu:?uy, VIL malimbóy, ABE malimbu:?iy ‘round’
 *pasánj BOT, CBG, KIL pahánj, CBG, VIL pamahánj ‘anger’
 *sa:nib BOT, KIL ha:nib ‘morning’

4.4.5 Shared Botolan and Ayta. Included here are items found in Botolan and shared more widely with the Ayta languages, that is, beyond Villar, Kakilingan, and Abellen. These speech communities are Mag-antsi (MAN), Mag-indi (MIN), Ambala (AMB), and Magbukun (MBK).

4.4.5.1 Phonological/morphological items

- (112) *ʔain BOT, VIL, KIL, ABE, MAN ʔa:yin, BOT, MÍN, ʔain CBG, MÍN ʔayín, MÍN, AMB ʔayin ‘none’; cf. PAN *am(e)ʔin ‘all, finished’
- *ʔa:yaʔ BOT, CBG, VIL, MÍN ʔa:yaʔ, KIL ʔa:ya ‘ant’; cf. PPH *la(R)a(q) ‘ant’
- *ʔ[ii]ŋgan BOT, CBG ʔiŋgan, BOT, KIL ʔiŋgán, VIL ʔiŋganán, MAN ʔiŋga:nan, MÍN ʔiŋga:nan, AMB maŋiŋgan ‘to wait’; cf. PPH *tiŋ(g)aq ‘to wait’
- *diglim BOT, VIL, ABE, MÍN mariglim, CBG madiglim, KIL diglim ‘dark’; cf. PSBL *diblim ‘dark’
- *diŋʔiy BOT, VIL, ABE diŋʔiy, CBG ʔaŋkariŋʔiy, KIL, AMB madiŋʔiy, KIL, MAN, MÍN mariŋʔiy, AMB jŋiŋʔiy, MBK najujuŋʔoy, nazuzúŋʔoy ‘shame’; cf. PLUZC *di[ŋŋ]ʔey ‘shame’
- *ditá BOT padtah, ditah, dumtá, BOT, KIL, dinumtá, CBG ʔimpacta, VIL ʔidutá, ʔipagtá, ditá, KIL dumtáh, ABE ʔipadtá, MÍN domtaʔ, dinumtaʔ ‘to stick to’; cf. PAN *ditaq ‘sticky substance, adhesive’
- *laním BOT, VIL, KIL, ABE, MAN, MÍN lanim ‘water’; cf. PMP *danúm ‘water’
- *lisay BOT, CBG, VIL, KIL, ABE malháy ‘big’; cf. PPH *seláR ‘big’
- *lilim BOT, CBG, VIL, KIL, MAN, MÍN li:im, VIL, ABE, MÍN li:ʔim ‘cloud’; cf. PPH *lulem ‘raincloud’
- *zi:ŋan BOT, CBG, VIL, ABE, MÍN di:ŋan ‘handspan’; cf. PMP *zaŋan ‘handspan’

4.4.5.2 Semantic items

- (113) *ʔidí BOT ʔampaʔiri, CBG, ABE paʔiri, CBG makiʔiri, KIL ʔampaʔidí, MAN paʔiriʔ ‘to live, dwell’; cf. PPH *idi ‘here’
- *balas BOT, VIL, ABE, AMB baláh ‘river’; cf. PMP *badas ‘sand’
- *sa:ʔa BOT, CBG, VIL, KIL, ABE ha:ʔa, MAN ha:ʔaʔ, MÍN, AMB saʔaʔ, MÍN sa:ʔa ‘banana’; cf. PPH *sahaq ‘banana trunk sheath’
- *tiik BOT, KIL, MÍN ti:ik, MAN, AMB ti:k, MÍN, AMB tiik, tuʔuk, MBK tiʔik ‘ear’; cf. PAN *[tT]eLek ‘deaf’

4.4.5.3 Lexical items

- (114) *ʔabu:ʔug BOT, MBK ʔabu:ʔug ‘fence’
- *ʔalindag BOT, VIL ʔalindág, MAN maŋalindag ‘to play’
- *ʔasiʔ BOT, CBG, KIL, MAN ʔa:hiʔ, CBG ʔa:hi, VIL, ABE, AMB ʔahíʔ, MÍN ʔa:si ‘not’
- *ʔay + *di BOT, CBG, VIL, KIL, ABE ʔayí, MAN ʔay ʔiri ‘where’?
- *ʔiŋat BOT, CBG, VIL, ABE ʔi:ŋat, KIL magʔi:ŋat, MBK ʔiŋatín ‘to call’
- *ʔipík BOT, CBG, VIL, ABE, MÍN, MAN, AMB maʔipík, AMB maʔipik ‘short (object)’
- *ʔukúy BOT, CBG, VIL, ABE, MAN, MÍN, AMB ʔukúy ‘egg’; cf. Bontok *okol* ‘to be developing, of mushrooms’
- *ʔulki BOT, KIL, MAN ʔu:ki, CBG ʔúlki, VIL, ABE ʔu:kiʔ ‘left (side)’
- *ba:lay BOT, CBG, VIL, ABE ba:lay ‘to know (someone)’
- *bi:taʔ BOT bi:taʔ, MAN, MÍN mi:taʔ, ‘to walk’; cf. Casiguran Dumagat *ita* ‘go ahead (on path) (imperative)’

*buŋkuk	BOT, CBG, VIL, KIL, ABE, MÍN buŋkúk ‘to lie’
*dagáw	BOT, CBG, VIL, ABE dagáw, MÍN magdaragaw, MBK magdagaw ‘to play’
*du:day	BOT, CBG, VIL, KIL, ABE, MAN, MÍN, MBK du:ray, MÍN du:day, AMB du:ɟay, AMB, MBK ju:ɟay ‘urine’
*g[íi]tan	BOT, CBG, VIL gitan, BOT magtán, CBG maŋtan, KIL, ABE, AMB gitán ‘to bring, carry’
*g[íi]gat	BOT, ABE giga:tin, MAN maŋigat, AMB gi:gat, BOT, VIL gi:gat, MÍN giga:tin ‘to scratch’
*gugút	BOT, MAN, MÍN, AMB gugót ‘tooth’; cf. Ilokano <i>gugú</i> ‘gum’, Northern Kankanaey <i>gugút</i> ‘nasal septum (of hogs)’
*kana:yun	BOT, CBG, VIL, KIL, ABE, MÍN, AMB, MBK kana:yun ‘other’
*[In]jawi:ni	BOT, CBG, VIL, KIL, ABE, AMB, MBK nawi:ni, MÍN, AMB, MBK nawi:ni? ‘body’; BOT, KIL, MAN lawi:ni ‘body’
*pasi:ŋa?	BOT, CBG, VIL, KIL, MÍN pahi:ŋa?, MÍN pasiŋa? ‘frog’
*pitá?	BOT, CBG, VIL, ABE mapitá?, MAN napitá? ‘soft’; cf. Bontok <i>pita</i> ‘to mash cooked cassava’
*sambak	BOT, KIL, ABE, AMB mahambák ‘morning’
*sa + pa:ʔig	BOT, CBG, VIL, KIL, MÍN, AMB hapa:ʔig ‘now, today’
*títiw	BOT lubut títíw, KIL, AMB títíw, MAN ti:tiw, MÍN títíw, AMB títíw ‘vagina’
*tuŋʔuy	BOT, MÍN tuŋʔuy ‘summit’
*(y)amu?	BOT, CBG, VIL, ABE, MAN ʔaʔamu?, ʔa:mu? ‘few’; BOT, CBG, KIL, ABE, MAN maya:mu?, VIL maya:mu ‘small’

4.4.6 Shared Ayta. This is a list of items shared by two or more of the Ayta languages and not appearing in any of the three major dialects of Sambal.

4.4.6.1 Phonological/morphological innovations

(115) *balʔu	‘widow(er)’; cf. PWMP *bálu ‘widow(er)’ MÍN, MBK, balʔó? MÍN, AMB balʔó, ‘widow, widower’
*dalipapá(n?)	‘palm, sole’; cf. PPH *d/al/apa and PMP *dapan ‘sole’ KIL dalipapán, MAN, MÍN, AMB dalipapá?, MÍN dalipapa ‘palm of hand, sole of foot’
*di:wi	‘thorn’; cf. PMP *duRi ‘thorn’ MAN, MÍN, AMB, MBK di:wi?, AMB, MBK ji:wi? ‘thorn’
*lu:yu?	‘to bathe’; cf. PAN *DiRuq ‘to bathe’ MÍN, AMB magpalu:yu?, AMB palu:yu? ‘to bathe’
*palaŋka?	‘frog’ An apparent blend of PPH *paNkaq and PPH *p/al/aká?, both meaning ‘frog’ AMB, MBK palaŋká? ‘frog’
takya?	‘excrement’; cf. PPH *taklaq ‘excrement’ MÍN, AMB takyá? ‘excrement’

4.4.6.2 Semantic innovation

(116) *palakpak	‘wing’; cf. PPH *palakpak ‘clap hands’; cf. also PAN *pakpak ‘flap wings, clap hands’ MÍN, AMB, MBK palakpák ‘wing’
-----------------	---

4.4.6.3 Lexical innovations

- (117) *ʔamis MAN ʔamihín, MÍNʔamisin, ʔaminsin ‘now, today’
 *ʔantiʔ MAN ʔantsí, MÍN, AMB ʔantiʔ, MÍN ʔantsiʔ ‘whatchamacallit’
 *ʔa:ɲid MÍN, AMB ma:ɲid ‘good’
 *ʔayan MAN naʔa:yan, MÍN naʔayán ‘sweet’
 *ʔitín MÍN, AMB marji:tiŋ, MBK ʔuturjun, marju:tun ‘to wait’;
 cf. SBLBL ʔitiŋ ‘long (time)’
 *ʔilab KIL, MÍN ʔi:lab ‘pain’
 *ʔiyag KIL, MAN, MÍN manyág ‘to make, do’
 *ʔu:ɲiy VIL, MAN ʔu:ɲiy ‘headcold’
 *ʔuyán KIL nauyán, MÍN, AMB naʔuyán ‘black’
 *baʔig MÍN bi:gin, AMB nagbaig, baigin ‘to call’
 *bantaw MÍN, AMB bantáw ‘shoulder’
 *banjiban KIL banʔi:ban, MÍN banjiban, Mlw banjiban ‘feather’;
 cf. Tagalog *bangibang* ‘head plumage, used by tribesmen
 as part of head dress and also on army helmets’
 *b[ɲu]liw MÍN mabuliw, maboliw, MÍN, AMB mabiliw ‘to sleep’
 *biyag MÍN biyág ‘floor’; AMB biyág ‘house’
 *bulit AMB bulit, AMB, MBK mamulit ‘to spit’
 *buŋʔuy MÍN buŋúy, MÍN, AMB buŋʔúy ‘mountain’; cf. Ilokano
buŋ-ur ‘protrusion’
 *daʔip AMB daʔip, daip, AMB, MBK da:wup ‘wind’
 *da:nun MAN, MÍN mara:nun ‘morning’; cf. Ilokano *da:nun*
 ‘to arrive at’
 *danliŋ ABE da:ɲiy, MÍN, AMB, MBK danliŋ, MÍN danláy, MBK
 danlúy ‘eel’
 *datuʔ AMB da:tuʔ, MBK ha:ratu, datúʔ ‘moming’. This could be
 *datiʔ from PAN *dateŋ ‘to arrive’
 *didiŋ MÍN, MBK diriŋ, AMB jǐjiŋ, didiŋ, ‘eyelash’
 *duliŋ MÍN naduliŋ, AMB maduliŋ ‘smooth’
 *du:sin KIL, MAN du:hiŋ ‘charcoal’; cf. Tagalog *dusing* ‘dirt on face’
 *gabaw KIL, MÍN gabaw ‘to lie’
 *gi[nŋ]im VIL, KIL gi:ɲim, KIL, MÍN gi:nim, Amb maginim,
 MBK ginum ‘raincloud’; MÍN, MBK ginem ‘cloud’
 *kaʔawat MÍN kaʔa:wat, AMB kaʔawatan ‘afternoon’
 *kama:na(t) VIL kama:na, MAN kama:nat ‘ghost’; cf. below *ma:na ‘old’
 *k[ɲu]bil KIL nakibil, MAN makubil ‘cold’
 *ku:nat MÍN, AMB, MBK maku:nat ‘hard’
 *labyún MÍN, AMB nalabyún ‘green’
 *la:buk MBK la:buk ‘feather’; MÍN, AMB, MBK la:buk ‘hair’
 *la:tug ABE nila:tug, MAN, MÍN latu:gin ‘to stab’
 *lusuʔ MÍN lu:suʔ, AMB luhuʔ ‘boil (n.)’
 *ma:naʔ MÍN ma:naʔ, manna ‘old’; MAN mananta:wuʔ ‘ancestors’
 *misák⁷ AMB, MBK mamhák, AMB mahmak ‘big’
 *papay AMB, MBK papáy ‘cheek’
 *pidaʔ MÍN piraʔ, AMB pi:daʔ ‘spider’
 *pun KIL, AMB pun ‘yet’

7. Or, perhaps better, *misák ~ *simák.

* <i>(sa)ʔantu</i>	KIL ʔatú, MÍN ʔantoy, saantuʔ, AMB ʔantú, ʔantuya, ha:ʔantuʔ ‘where?’
* <i>sabiʔ</i>	MÍN sabíʔ, AMB habíʔ, MBK habúʔ ‘soup’
* <i>salsal</i>	KIL halhál, MAN manalhál, AMB salsalín ‘to weave (mat)’
* <i>sapak</i>	MÍN sapak, AMB ha:pak ‘floor’
* <i>sulu(k)</i>	KIL nanuluk, KIL naghuluk, MAN nahulúʔ ‘anger’; cf. Ilokano <i>sólok</i> ‘to be proud, arrogant’ and Itawis <i>túluk</i> ‘to push (s.o. around, to defeat)’
* <i>tagiy</i>	KIL, MÍN, AMB tagíy, MÍN taʔgíy ‘fruit’
* <i>tagul</i>	MÍN, AMB tagúl ‘knee’
* <i>tambuk</i>	MÍN, AMB ʔitambuk ‘to throw away’; AMB ʔitambúk ‘to throw’
* <i>tubág</i>	MÍN, AMB tubág ‘anger’
* <i>tu[lw]a:pu(k)</i>	KIL tuwa:puk, MAN tuwa:puʔ ‘dust’
* <i>yaŋgíŋ</i>	MÍN, AMB nayaŋgíŋ, AMB mayaŋgíŋ ‘dirty’
* <i>yawini</i>	AMB, MBK yawini ‘body’; cf. shared Botolan and Ayta *[ln]awi:ni ‘body’
* <i>ya:ʔuʔ</i>	VIL, KIL, AMB, MBK ya:wuʔ, MAN ʔiya:wuʔ ‘arrow’

4.4.7 Mag-antsi. Data from the available Mag-antsi-speaking communities show a number of innovative items.

There are three items that exhibit phonological change:

- (118) **ʔaya* ʔaya ‘not (verb, past)’; cf. PPH **ayaw* ‘negative’
 **lipaw* lumpáw ‘to float’; cf. PWMP **le(n)táw* ‘to float’
 **suyúʔ* huyóʔ ‘torch’; cf. PAN **suluq* ‘torch’

Another three items show a semantic shift from a recognizable antecedent:

- (119) **ʔu:liŋ* ʔu:liŋ ‘field, swidden’; cf. PAN **qujiŋ* ‘charcoal’
 **la:ʔu* la:ʔuʔ ‘mortar’; cf. PWMP **laqlu* ‘pestle’
 **su:buʔ* hu:boʔ ‘to suck’; cf. PPH **su:buq* ‘to put s.t. in mouth’

The remaining items appear to be lexical innovations.

- (120) **ʔadáw* ʔaráw ‘flood’
 **ʔaŋál* ʔaŋkaʔaŋál ‘thirst’
 **bagwát* bagwát ‘comb’
 **[bp]a:kiʔ* mama:kiʔ ‘to call’; cf. Ifugaw *báki* ‘invocations and prayers addressed to supernatural beings’
 **bitis* bithin ‘to exchange’
 **bu:diŋ* mabu:riŋ ‘raincloud’
 **daŋsik* darañhik ‘pestle’
 **di:yaŋ* nadi:yaŋ ‘dew’
 **kanan* kanán ‘pond, lake’
 **kiyáʔ* kiyáʔ ‘bolo’
 **lamit* ʔila:mít ‘to stick to’
 **la:pig* la:pig ‘ashes’
 **ŋi:suʔ* naŋi:hoʔ ‘sour’
 **pa:nat* mapa:nat ‘to lose’
 **pintúnŋ* pintónŋ ‘belly, intestines’; cf. SIN *pítúʔ* ‘belly’

- (127) *tikdi(l) tiklí?, tuklú? ‘to live, dwell’; cf. PSBL *tikdi(l) ‘to sit’
 *zikut di:kut, jíkút ‘forest’; cf. PPH *zikut ‘grass’

The remaining innovative items are lexical developments.

- (128) *ʔalun ʔalún, ha ʔalún ‘there (far from speaker and addressee)’
 *(ʔam)baŋan mamba:ŋan, ‘big, fat’
 *ʔandiŋ maʔandiŋ, maʔandi? ‘small, few’
 *ʔayma ʔáyma ‘finger, hand’
 *ʔik[íi]n ʔikún, ʔikín ‘not’
 *ʔitl[iu]k maʔitlúk, maʔikluk ‘short (object)’
 *da:pi? manda:pi?, ʔida:pi? ‘to throw, to throw away’
 *kapaw nakapáw ‘many’, makapaw ‘all’
 *k[íu]ltimin kultí:min ‘debt’
 *kumit maŋu:mit, maŋumit ‘to steal’
 *lamʔig ʔilamʔig, ʔilalamʔig ‘to swallow’
 *sa + ʔatu haʔatu, ha:tu? ‘now, today’
 *saʔay haʔáy, hayi ‘where?’
 *sagmak mahagmák ‘many’
 *sali(?) mahalima:sali?, mahalí? ‘near’
 *takbaw matakbaw ‘fast’; manakbáw ‘to run’; cf. Tagalog *takbó* ‘to run’
 *tamudaw tamuráw ‘vagina’
 *tiktik natuktúk, tuktúk, tuktukún, maniktik ‘to burn’
 *tulbuy matulbúy ‘straight’
 *tuŋʔu tuŋʔú ‘forehead’; cf. Tagalog *tungó* ‘with the head bowed’

5. LEXICOSTATISTICS. As a tool for calculating the relative closeness of relationships between or among related languages, lexicostatistics can provide some clues to the internal structure of a microgroup. Glottochronology, which attempts to convert lexicostatistic calculations into an absolute chronology of language divergence, is largely out of favor, and for a good number of reasons. The value of lexicostatistics, however, is highly limited without an assurance that the data under comparison are truly comparable. That is to say, lexical items for a concrete object or a notion in one language must be truly cognate with such items in another language. So it is absolutely necessary for the comparisons to be made only on such lexical input as has been verified by sound application of reconstruction techniques. Without the latter, we would inevitably find excessively high relationships between languages that have borrowed from each other (no matter how distantly related they may be genetically), and we would miss the match between items that, while truly cognate, appear not to be because of significant differences in the output of their language-specific rules of phonological change.

In the Central Luzon microgroup, there has been an enormous amount of diffusion among the dialects and languages themselves, as well as between these languages and Tagalog. Since loans may well be phonologically naturalized, the diffusion situation is made even more complex. This problem has occurred with other attempts to establish the internal structure of the LUZC microgroup (Stone 2008).

With regard to the Ayta languages, the issue of the original language acquisition even further limits the utility of lexicostatistics. It is apparent that the Ayta first acquired their

versions of Austronesian from Sambal speakers, most likely those speaking the Botolan dialect. This is quite similar to the situation that obtained in the Cagayan Valley, in northeastern Luzon, in the Bikol region, and elsewhere in the Philippines. Given this, and given the various degrees of continued interaction between the Ayta and Sambal speakers, and among Ayta communities themselves, the lexicostatistical conclusions are not very informative. Whether calculated by the “syllolexicostatistical” method used by Stone (2008) or by other methods based on reconstructed material (such as I have used), we come to the not very startling conclusion that the Sambal dialects are more closely related to each other than any of them is to Kapampangan. The results of these two investigations are not widely divergent from each other, except perhaps in calculating the relationship of Kapampangan to Bolinao and Sambal Tina (Sinauna was not included in Stone’s work).

(129) Dyadic Comparison	Stone	Himes
SBLBL : SBLT	.75	.75
SBLBL : SBLBT	.66	.68
SBLBL : KPM	.46	.39
SBLBL : SIN	—	.34
SBLT : SBLBT	.70	.76
SBLT : KPM	.45	.40
SBLT : SIN	—	.34
SBLBT : SIN	—	.38
KPM : SIN	—	.44

Sinauna shows low percentages of shared cognates with the Ayta languages, as it does with the dialects of Sambal, ranging from a high of .35 with Mag-antsi to a low of .31 with Ambala. On the basis of these figures, we would come to the conclusion that Sinauna is reasonably closer to Kapampangan than it is to Sambalic. With regard to the pronouns, however, Sinauna appears to be somewhat closer to the western branch of the family. It is noted above that Sinauna shares only one distinctive rule of phonological change with Kapampangan, the lowering of Proto-Philippine schwa, and none with the Sambalic languages. Further, with regard to the innovative lexicon, Sinauna shares one item with Kapampangan, two with Sambalic, and one with both of them. This is very poor evidence for the establishment of a linkage within the microgroup.

The most acceptable conclusion, then, is to posit the internal structure of the Central Luzon microgroup as consisting of three equidistant branches: Sambalic, Kapampangan, and Sinauna.

6. CENTRAL LUZON AND OTHER MICROGROUPS. Zorc (1974) first proposed that the Central Luzon languages may cluster with three Mangyan languages—Iraya (IRA), Alangan (ALA), and Tadyawan (TDY)—of northern Mindoro in a single group. Their sharing of a /y/ reflex of *R, together with at least a few exclusively shared lexemes, formed the basis of this hypothesis. Blust (1991) apparently accepts this grouping. Zorc also pointed out that the Bashiic languages also reflect *R as /y/. This might indicate that Yami, Ivatan, and so on, subgroup with the Central Luzon and Northern Mangyan languages in contrast with other Philippine languages.

6.1 CENTRAL LUZON AND NORTHERN MANGYAN. There are no generally held criteria for the establishment of a microgroup as distinct from others. Yet there is a consensus that certain languages do indeed form a group distinct from other languages and groups. Manobo, Palawan, Minahasan, and other language microgroups are well established. Minimally, of course, such groups of languages must share some phonological regularities and some other features, such as innovative lexicon, with each other distinct from other languages (Blust 2008).

In the present state of our knowledge, it is clear that the Central Luzon languages form a group distinct from all others, and that this group includes Kapampangan, the three dialects of Sambal (Bolinao, Sambal Tina, and Botolan), the languages and dialects spoken by various Ayta groups in the Zambales Mountains, and Sinauna (and Remontado) of Rizal Province. Nor can there be any doubt of the close relationship among the three Northern Mangyan languages of Mindoro.

What we need is evidence that the Central Luzon languages and the Northern Mangyan languages do share enough to be so linked. And once established, we will have confirmed the grouping first suggested by Zorc. Blust (1991) provides convincing evidence that this Central Luzon group once extended farther south, into Palawan, as is demonstrated by the occurrence of /y/ in Palawan languages where /g/ is expected as the reflex of *R. The same is true for the Kalamianic languages of northern Palawan, where /y/ occasionally occurs when /l/ is expected (Himes 2007).

6.1.1 Phonological Evidence. It appears that the Northern Mangyan languages share a number of phonological shifts from PMP to PLUZC discussed in 2.3 above. But then again, many Philippine microgroups also share these rules. PPH *j and *z merge with the reflexes of *d, but it remains to be determined if this occurred in a specific order. Some MNGN items with word-initial *d reflect this phoneme as a liquid, under the same or similar conditions as the process requires in the Sambalic languages, such as in initial position when the next consonant is not a liquid.

(130)		IRA	ALA	TDY
	'big'	PPH *dakul	—	—
	'deaf'	PAN *dɔŋɔn	—	—
	'to hear'	PAN *deNéR	kariŋiy	kariŋiy
	'lonely'	PPH *demdem	magrimrim	magrimrim
	'wind'	PPH *dejes	—	ʔariris
	'other'	PAN *duma	ru:ma	ru:ma
				lakul
				luŋún, kariŋin
				riŋyan
				pandimdim

Those items, such as the following, that seem to contravene this rule, may well be borrowed from Tagalog or Visayan languages.

(131)		IRA	ALA	TDY
	'wall'	PMP *diŋdiŋ	—	—
	'breast'	PWMP *dúdu?	du:du	—
	'leaf'	PMP *dahun	da:ʔun	da:ʔun
	'thick'	PPH *dameŋ	—	—
				madamil

Most convincing of the evidence for a link between the LUZC and MNGN languages, however, remains the regular reflex of *R as /y/.

(132)		IRA	ALA	TDY
‘new’	PMP *baqeRu	bayú	ba:yu	bayu
‘heavy’	PMP *beReqat	mabyat	mabiyát	ʔabyat
‘to bite’	PAN *kaRát	kayát	kayát	kayát
‘mouth’	PPH *bibíR	bibí	bibí	bibíʔ
‘neck’	PPH *leqeR	li:ʔiy	liʔiy	—

Further study of the phonology of the Northern Mangyan languages is necessary for an understanding of the exact nature of the ways the two language groups correspond.

6.1.2 Shared innovations. Evidence for a linkage between one group of languages and another is sought in other areas besides that of regular sound changes. Innovative change can also provide some confirmation of a relationship. Those items that are shared by Central Luzon and Northern Mangyan, then, could be attributed to a level somewhere in time between Proto-Malayo-Polynesian and the current microgroup levels. Some of these innovations are lexical, some phonological/morphological, and some semantic.

6.1.2.1 Phonological/morphological evidence. There are at least three such innovations:

(133) *ʔakit	‘to see’; cf. PAN *kítaʔ ‘to see’ SBLBL, SBLT maʔkít, SBLBT, MAN ma:kit, AMB naʔakit, KPM ʔa:kit; IRA ʔa:kit, ALA paŋkiton ‘to see’
*gaRaw	‘to scratch’; cf. PAN *kaRaw ‘scratch’ MBK gayowin, maŋga:yo; IRA, ALA ga:yaw ‘to scratch’
*liyaki	‘man, male’; cf. PHN *la-láki ‘man, male’, most probably via *l<in>aki ‘man male’ MIN, AMB liya:kiʔ, AMB ya:kiʔ; ALA kalyaki:yan ‘man, male’

6.1.2.2 Semantic evidence. Shared semantic innovations also provide some evidence of a common history:

(134) *ʔutan	‘snake’; cf. PMP *qútan ‘forest’ SBLBT, MIN, MAN, AMB, MBK ʔu:tan; IRA ʔután ‘snake’
*beles	‘answer’; cf. PPH *be(jd)es ‘talk’ KPM ʔablas; ALA mi:lis, bulusan ‘answer’
*labu:yu	‘rooster’; cf. PPH *labu:yuʔ ‘jungle fowl’ SBLT labu:yu, MAN labu:yuʔ; ALA, TDY labu:yo ‘rooster’
*liqes	‘louse’; cf. PPH *liqes ‘louse egg, nit’ MIN lis, MBK le:yos; IRA liis, ALA luyos, TDY lis ‘louse’
*peR-lepu-an	‘ankle’; cf. ISNEG <i>lappó</i> ‘thigh’ ABE, MIN pilpuʔán; TDY poylupuán ‘ankle’; cf. PWMP *lepuq ‘crooked, of limbs’
*sa(m)puk	‘to kick’; cf. PWMP *sa(m)puk ‘collide, bump into’ SBLBT sapúk; ALA sampukin ‘to kick’; cf. Tagalog <i>sapók</i> ‘uppercut (in boxing)’
*tamʔis	‘sugarcane’; cf. PAN *tamqis ‘sweet’ MBK matamʔis; TDY katamis ‘sugarcane’

6.1.2.3 Lexical innovations. These include:

- (135) *ʔazuk ‘bad’
 KPM marúk ‘bad’; TDY karok ‘bad-smelling’; cf. Cebuano and Hiligaynon *ʔálok* ‘witchcraft’
- *bagbag ‘forest’
 SBLBT, KIL, MÍN bagbág, binagbág; ALA bagbág ‘forest’; cf. Casiguran Dumagat *bagbag* ‘a newly cleared rice paddy (that has recently been cleared from forest land)’; cf. also Proto–Nuclear–Central Cordilleran *pagpag ‘forest’
- *d[ae]gsaʔ ‘to push’
 KIL ʔiragsaʔ; IRA digsiʔ, digsiʔin; ALA digsin ‘to push’
- *dimla ‘cold’
 KPM marimla; IRA, ALA madimlaʔ; TDY kadimla ‘cold’
- *liʔmu ‘to fear’
 SBLBL, SBLT maliʔmú; SBLBT, ABE mali.mu; MAN, MÍN, AMB, MBK mali.muʔ; IRA, TDY limu; ALA, TDY kalimu ‘to fear’
- *mamaw ‘ghost’; cf. PPH *limaw ‘ghost’
 MBK ma.moʔ; TDY maw, mamaw ‘ghost’
- *(r)apak ‘ginger’
 KIL ʔapak; TDY rapak ‘ginger’, But cf. Agusan Manobo *ʔapaʔ* ‘ginger’
- *tupaR ‘saliva, to spit’
 MÍN mantupay; MÍN, AMB manupáy; IRA tumupay; ALA tu:pay; TDY tupáy ‘to spit’, TDY tupáy ‘saliva’; cf. Northern Kankanaey *topá* ‘an expression of hatred, spite, usually pronounced after spitting’; cf. also Central Bontok *tobba* ‘spit’. This item has diffused into Buhid as *tufáy*.

While the evidence for a Central Luzon–Northern Mangyan link is not overwhelming, it is probably sufficient to justify a closer relationship between these two micro-groups than that enjoyed by either of them with other Philippine groups.

6.2 BASHIIC AND THE CENTRAL LUZON AND NORTHERN MANGYAN LANGUAGES. Zorc (1974) noted that the Bashiic languages spoken in the islands of the northern Philippines and southern Taiwan also show a /y/ reflex of *R. But he noted, as well, that the lexicon of Bashiic languages is substantially different from those of the more southerly languages.

6.2.1 Phonological developments. The common rules that shift PMP *q to the glotal stop, loss of unstressed *e, and the merging of *j and *z with the reflexes of *d are seen in the Bashiic languages, as they are in many of the languages of the Philippines. PPH *h is lost in Yami (YAM) and Ivatan (IVT), but it is preserved in Itbayatan Ivatan (IVTI).

The Bashiic languages do indeed have /y/ as a reflex of *R.

- | | | | | |
|---------|------------|---------|---------|---------|
| (136) | | YAM | IVTI | IVT |
| ‘root’ | PMP *Ramút | yamút | yamút | yamút |
| ‘bathe’ | PHN *diRus | mariyús | mariyús | mariyús |

‘new’	PMP *baqeRu	vayú?	vaʔyú?	vaʔyú?
‘sated’	PMP *ma-besúR	mabsúy	mabsúy	mabsúy

Another phonological process is shared exclusively by the Bashiic languages and Kapampangan. This is a rule whereby, in KPM, a word-initial sequence of consonant-schwa-consonant-vowel metathesizes the initial consonant and schwa; the schwa later shifts to /a/. The rule is:

$$(137) *C_1eC_2 > eC_1C_2 > aC_1C_2 / \# _ _ V$$

Examples are:

(138)			PLUZC	KPM
‘answer’	PPH *be(jd)es		*bilis	ablas
‘husked rice’	PAN *beRas		*biyas	abyas
‘sugarcane’	PMP *tebuh		*tibu	atbu
‘to swallow’	—		*kimul	akmul
‘to split’	PWMP *sepak ‘break’		*sipak	aspak
‘to bundle’	—		*i-pitis	yaptas

The process in Bashiic is nearly identical, except that the lowering of the schwa is not a regular feature in this microgroup. The lowering of schwa, then, must be seen as a part of the original rule:

$$(139) *C_1eC_2 > aC_1C_2 / \# _ _ V$$

Examples:

(140)	‘three’	PAN *telú	ʔatlú?	YAM, IVTI
	‘answer’	PPH *tebaR	ʔatbay	YAM, IVTI, IVT
	‘hear’	PAN *deNÉR	adjijy	IVTI, IVT
	‘pond’	PPH *lebeŋ	ʔahbiŋ	YAM
			ʔaybiŋ	IVTI
			ʔahbiŋ	IVT

6.2.2 Lexical evidence. If we look at the lexical evidence, we come to the conclusion that the Bashiic languages are distinct enough from the languages of Central Luzon for them to be classified as a separate microgroup. Let us look specifically at the four groups of languages of the northern Philippines.

In the data available, Bashiic shows 235 items from the higher (that is, extra-Philippine) levels—Proto-Austronesian, Proto-Malayo-Polynesian, and so on. Of these, Bashiic shares 161 (68.5%) with Central Luzon, 105 (44.7%) with Northern Mangyan, and 196 (83.4%) with Northern Philippine (Cordilleran). If we look at those items that are innovative within the Philippines, Bashiic shares with Central Luzon, to the exclusion of other microgroups, 19 items, and it shares three items exclusively with Northern Mangyan. There is one further item that is shared only by Bashiic, Northern Mangyan, and Central Luzon. In contrast, Bashiic and Cordilleran share 81 items to the exclusion of all other microgroups. Clearly, the Bashiic and Cordilleran languages appear to be somewhat more closely related lexically, yet Bashiic shares with Cordilleran no phonological processes that it does not also share with other Philippine microgroups. On the basis of these figures, it is most logical to conclude that BSH is not appreciably closer to LUZC and/or MNGN than it is to Cordilleran.

That said, any lexical items that are exclusively shared by BSH and LUZC and/or MNGN should be added to the lexicon attributable to the level of PPH or some other source at a higher level than the individual microgroups. These include the following:

- (141) PPH *qah 'negative'
 IVTI ʔah; MIN, AMB, MBK ʔa 'not'
- PPH *qaptus 'to rub'
 IVTI maŋaptus 'masseur'; IVT ʔaptusan; SBLBT ʔaptuh;
 MAN ʔaptuhín 'to rub'
- PPH *qasel 'foot, calf of leg'
 IVTI ʔasóy; SBLBT, VIL, ABE ʔa:hił; MAN, AMB, MBK
 ʔa:hił 'calf of leg'; MIN ʔasil 'foot'
- PPH *qubid 'rope'; cf. PPH *lubid 'cord'
 YAM ʔovíd; SBLT ʔubil 'rope'
- PPH *qudiq 'left (side)'
 YAM ʔoriʔ; MIN ʔudiʔ, ʔu:diʔ, ʔuriʔ 'left'
- PPH *quya 'this'
 YAM ʔoyáʔ; SBLT ʔu:ya 'this'
- PPH *buklas 'morning'
 IVT mavikhas, mavokhás; SBLBL, SBLT (SCR) buklás;
 SBLT buklá, bukláh 'morning'
- PPH *hilak 'white'
 YAM meylák; IVTI mahilák; IVT maydák, mayidák; IRA
 ʔinlakan 'white'
- PPH *mu 'only'
 IVT; KPM mu 'only'; SBLBL tamu 'only'
- PPH *muRa 'anger'
 YAM mioyáʔ; KPM muwáʔ; REM minamuyan 'anger,
 angry'
- PPH *tapiq 'floor, covering' ? Cf. Tagalog *tapi* 'apron'
 IVT tapiʔ, ta:piʔ; SBLT ta:piʔ; SBLBT tapiʔ 'floor'
- PPH *t/ar/ini[kŋ] 'straight'
 YAM, IVTI matarinóŋ; IVT mataliniŋ; MIN matinik
 'straight'
- PPH *ŋepen 'tooth'; cf. PAN *ŋipen 'tooth'
 YAM ŋəpən; MBK ŋu:pun 'tooth'

7. CONCLUSIONS. It is clear from the treatment above that the cluster of languages of Central Luzon are related to each other, in varying degrees of proximity, because of a common ancestry. The three major branches of the LUZC microgroup evolved from a common ancestor through the usual processes of language change, ultimately differing among themselves as a result of separation and development in relative isolation after their initial split. The PLUZC-speaking peoples settled the western coast of Luzon south of the Ilocos region, and from Manila Bay they spread into the Central Plain. The coastal dwellers came into contact with a previously settled Negrito population, and the latter acquired their Sambalic speech forms from the newcomers. Those who settled on the coast east of the Zambales Mountains were most probably forced northward by the immigrating Taga-

log speakers. At some point this Central Luzon-speaking population encountered the group of Negritos ancestral to the modern Sinauna and Remontado populations, as well as the now extinct Tayabas Ayta, if indeed these latter were speakers of a Central Luzon language. These Negrito populations acquired the form of LUZC spoken by this new population, and subsequently remained in their traditional lands (Reid 2010) or filtered east of the Central Plain into the mountainous areas of Rizal and Quezon provinces.

The migration that brought the LuzC speakers to the coast of Zambales and the Central Plain apparently included a group or groups that settled in northern Mindoro, and evidently the island of Palawan. These would be the ancestors of the modern Northern Mangyan and the speakers of languages in Palawan that were subsequently absorbed by populations speaking Kalamianic and Palawanic languages.

APPENDIX. PPH RECONSTRUCTIONS

Lexical items reconstructed in this paper that are attributable to the level of Proto-Philippines, or higher, and that have not appeared in other publications include these.

*bahandi	KPM mabandi; Cebuano bahandi, Samar-Leyte bahándi? ‘rich’
*ba:kir	Iloko bákir ‘forest, woods’; Northern Kankanaey bákil ‘to prepare lumber’
*baluŋus	SBLBT balo:ŋo, MIN baloŋos, MBK balu:ŋu?; Tasaday Manobo baluŋús, Blit Manobo bliŋús ‘nose’; Tagalog baluŋus ‘snout of fish’
*bibiR	IVT viví?, IVTI vivih; IRA, ALA bibí, TDY bibí?; Gaddang bi:fig; Northeastern Luzon bibig; Tagalog bibíg; Hanunóo, Buhid bibíg, Aborlan Tagbanwa bibíg; Mongondow bibig ‘mouth’; cf. PAN *bibíR ‘lip’
*buq[eu]	SBLT bu?ól, SBLBT bu?í, KIL bu?ih, MIN bo?í?, AMB bu:wí?; Southern Alta bu?ol; Bikol bu?ól, Kuyonen bu:l; Romblonun bo?óy; Karamianan bu:l ‘heel’; cf. PPH (Charles) *buqel ‘ankle’
*d/al/apa	SBLBL dapa; ALA darapa(?), TDY dalapa; Kuyonen raparapa?, Hiligaynon dapádapá, Cebuano lapálapá, Samar-Leyte rapádapá; Mamanwa rapa?rapa?, Hanunóo raprapá, Buhid yapyáfa; Karamianan dáparápa? ‘sole’; Kallahan da:pá? ‘palm of hand’
*daras	SBLBL madaras; Iloko nadaras; Bontok, Northern Kankanaey, Kalingadálás, Kagayanan dalas ‘fast’
*dayun	SBLBL rayon, SBLT, SBLBT, KIL, VIL, ABE, MIN, MAN layún, AMB dayún; Ilongot diyundiyun ‘earthquake’; cf. PMP *hayun ‘swing, shake’
*dejes	ALA ?ari:ris, ?arurus; Bikol du:rus; Batak, Aborlan Tagbanwa, Molbog diris, Southern Palawano dARAS ‘wind’; cf. PMP *deRes ‘heavy, of rain’
*demdem	Atta, Itawis, Gaddang maraddam, IRY magrimrim, TDY pandimdim; Aborlan Tagbanwa dimdim ‘lonely’; cf. Tagalog <i>panimdim</i> ‘misgivings’
*digdig	SBLBL rigrig, SBLT liglig; Central Cagayan Agta zigzig; Northeastern Luzon digdig; Southern Alta padigdig ‘shore’
*diRas	SBLT, SBLBT madiyá?diyá?, Isneg dixat, Malaweg nadi:gas, Yogad madi-gat, Gaddang middyat ‘difficult’
*duk[ae]s	SBLBL, SBLBT duká?, SIN marúkas; Northern Alta madú?as; Northeastern Luzon madukas, dukis ‘bad’; IRA ru:kas ‘rotten’
*garamáy	SBLBL, SIN garamáy, SBLBT, MAN, MIN galamáy, Ilongot giyamáy; Northern Alta gala:máy ‘finger’; Tagalog galamáy ‘all the fingers of the hand’
*kay[eu]p	MAN kayi:pan, MIN kayu:pan, AMB ka:yip; Magindanaw kayup ‘to blow’

- *ka:zaŋ SBLBT, ABE maka:raʔ, KIL maka:raʔ, MAN makaráʔ, MIN makaraŋ, AMB maka:daŋ, makaʔaŋ, maka:zaŋ ‘long’; YAM, IVT, IVTI makaraŋ ‘tall, high’; Iloko kadaŋkádaŋ ‘stilts’; cf. Proto-Southern Cordilleran *kadaŋ ‘handspan’
- *kuqkuq SBLBL, SBLT koʔkóʔ, SBLBT, ABE, MAN, MIN, KPM ku:koʔ; Umiray Dumagat ku:koʔ ‘to cough’; cf. PMP *ukuk ‘to cough’
- *(k)ulmug SBLBL, SBLT ʔulmúg, SBLBT, ABE, MAN kulmúg; Iloko, Northeastern Luzon, Pangasinan ulmúg; Ilongot kólmug ‘chicken louse’
- *labu:yuq SBLT labu:yo, MAN labu:yoʔ; ALA, TDY labu:yo ‘rooster’; cf. Tagalog *labuyò* and KPM *labu:yuʔ* ‘jungle fowl’
- *laguq Tagalog lagò ‘luxuriant growth’; Casiguran Dumagat lagu ‘healthy (of crops)’
- *laway SBLBT, VIL, ABE, MAN la:way; Tagalog, Bisayan, Aborlan Tagbanwa laway; Agutaynon lawáy ‘saliva’
- *lebeŋ YAM, IVT ʔahbiŋ, IVTI ʔaybiŋ; Kankanaey libiʔiŋ, libiŋ, Ifugaw lobóŋ, I-wak dibíŋ, Kallahan lebéŋ, Ilongot ʔibiŋ ‘lake, pond’; cf. PWMP *le(m)beŋ ‘valley’
- *leqeR SBLT lu:ʔuy, SBLBT, KIL, VIL, ABE, MAN li:ʔiy, AMB, MBK li:ʔiy; IRA, ALA liʔiy; Kagayanen liʔig ‘neck’; cf. PAN *liqeR ‘neck’
- *iimaw SBLBT mallili:mo, ABE malili:mo; Brookes Point Palawano moŋlilimow ‘ghost’
- *liques SBLBL li:ʔis, SBLT li:ʔuh, SBLBT li:ʔih, KPM lias; Itneg, Pangasinan liyís; Aborlan Tagbanwa liʔis; Agutaynon, Kalamian Tagbanwa likít ‘louse egg, nit’
- *luksu Tagalog luksó; Iloko loksó; Casiguran Dumagat luksu ‘to jump’
- *lulem SBLBL lu:lím, SBLT lu:lum, KPM lu:lám; Hanunóo rúrom, Buhid rurom ‘cloud’; Iloko lúlem ‘cloudy, overcast’
- *miNsan SBLBL maʔinsan, SBLT paysán, páʔisan, SBLBT paʔiŋhan, papaʔiŋhán ‘later (same day)’; Tagalog minsan ‘once’; Northern Kankanaey miŋsan ‘next time’
- *mu KPM, IVT mu ‘if’
- *muRa KPM muwá ‘anger’; SIN minamuyan; YAM mioyáʔ ‘angry’
- *p[ae]:gaw SBLT, pa:gaw, KIL po:gaw; Bontok piɡʔiw, piɡpiɡʔiw, Kankanaey pa:giw, Ifugaw pa:go, Isinai pahaw, Pangasinan paɡiw, Ibaloy pa:gow, Karaw pa:xow, Northern Alta paɡúʔ ‘chest’
- *palsu(k) ABE palhóʔ; TDY palaso ‘arrow’; Kalinga palsúʔ, palsúk ‘bow-and-arrow’; Casiguran Dumagat palsok ‘a kind of arrow’
- *paltak SBLBL, SBLT, VIL, KIL, ABE, MAN; Pangasinan, Ibaloy, Karaw paltak; SBLBT pa:tak ‘testicles’
- *pated SBLBT, MAN, MIN, ABE, MBK patíl, KPM kapatad; Tagalog kapatíd; Sindangan Subanun patid ‘sibling’; Samar-Leyte patúd ‘cousin’
- *qalinjasaw SBLT ʔalinjaháw, ʔampaliŋa:haw; Tagalog alingasaw; Iloko aliŋásaw; Pangasinan aliŋasew ‘to stink’; Bontok aliŋásew ‘to give off steam’
- *qal(i)wa SBLBT, KIL, ABE, MIN, MAN, AMB ʔalwá, KPM aliwa; Pangasinan aliwá; Ibaloy ʔaligwán, ʔalivwá; Kallahan ʔaliwwá, ʔaliwwán ‘not (of nominals), other’
- *qaRjaw SIN ʔaydaw; Tasaday, Blit, and Kalamansig Cotabato Manobo ʔagdaw ‘day’
- *qaslem KPM maslam; IRA maʔaslim; Kuyonen ʔaslim; Aklanon, Hiligaynon, Tausug, Magindanaw maʔaslum; Buhid maslóm; Agutaynon makaʔlím, Kalamian Tagbanwa makaklím ‘sour’
- *qi[jdr]jib KPM ilib; Agutaynon and Kalamian Tagbanwa kirib ‘cogon’
- *qikluR Ibaloy (Atok) ʔekdúl; Central Cagayan Agta iklug; Marinduque Tagalog, Tausug ʔiklúg; Magindanaw ʔiklog ‘egg’; cf. PMP *qiteluR ‘egg’
- *quRbun KIL, AMB, MAN ʔoybon; IRA ʔibín ‘small’; SIN ʔubun; IRA ʔibun ‘young child’; Iloko urbún ‘young horse or cow’; Casiguran Dumagat ógbun ‘baby birds’; cf. Proto-Oceanic *ubu ‘young coconut’

*ribay	KPM ?ipaglibe; Bikol ríbay ‘to exchange’
*sabut	SBLBL, SBLT (SCR) sabut, SBLT, SBLBT, MAN, MÍN habút ‘feather, hair’; Northern Alta, Southern Alta, Northeastern Luzon, Arta, Umiray Dumaget sabút; Aklanon, Kagayanen sabút ‘pubic hair’
*sahaq	SBLBT, ABE ha:aʔ, MAN ha:ʔaʔ, MÍN, AMB saʔaʔ, MÍN sa:ʔa ‘banana’; Tagalog sahaʔ ‘petioles or sheathing of banana trunks’
*siduŋ	SBLBL si:ruŋ, SBLT, SBLBT hi:luŋ, MAN, MÍN, AMB si:luŋ, KPM si:luŋ, SIN siruŋan; Northern Alta siduŋ; Isneg, Adasen si:duŋ; Northeastern Luzon siduŋ; Tagalog, Bisayan si:long, Bikol si:ruŋ; Inati sidóŋ; Umiray Dumaget se:duŋ; Batak, Aborlan Tagbanwa siruŋ; Agutaynon, Kalamian Tagbanwa siruŋ ‘space-under-house’
*sikan	KPM nasikan; Isneg, Malaweg, Ibanag nasikan; Atta, Itawis nasikan; Central Cagayan Agta, Northeast Luzon masikan ‘strong’; Isneg, Itawis nasikán; Central Cagayan Agta masikan ‘fast’
*si:riR	SBLBL si:ri ‘wind’; Hanunóo sirig ‘storm’
*su:buq	MÍN hu:boʔ ‘to suck’; Iloko so:bo ‘to feed, to convey food to the mouth of’; Casiguran Dumagat subu ‘to put something into the mouth’
*tagkes	SBLT (SCR), MBK tagkús, MBK tagkíh; Cotabato Manobo, Binukid tagkis ‘to tie’
*taked	SBLBL takir, SBLT takúl, SBLBT, ABE, MAN takíl; Bontok, Northern Kankanaey, Southern Kankanaey, Itneg takíd; Bolinao taʔad, Ifugaw taʔód, Kalinga taʔud, takud ‘to tie, bind’
*taklaq	KPM; SIN; IRI, ALA taklaʔ; Bikol (Northern Catanduanes) takráʔ ‘excrement’
*taklay	Central and Southern Cordilleran taklay ‘arm’; Mamanwa pataklayan ‘forearm’
*tebaR	SBLBT, KIL, MÍN tibáy; YAM, IVT, IVTI ?atbáy; Proto-Cordilleran (Reid) *tebaR ‘to answer’; cf. PPH (Zorc) *tubaR ‘to answer’
*waléŋ	SBLBL walíŋ, SBLT walúŋ, SBLBT wawíŋ ‘forehead’; Surigaonon wayóŋ, Batak, Kagayanen walíŋ, Binukid wiliŋ ‘face’
*zaRem	SBLBL kada:yim, SBLBT, KIL, AMB kara:yim, SIN kadayam, IVT dayim, IVTI rayim; Kiniray-a, Kagayanen, Dibabawon Manobo dáŋim ‘needle’; cf. PAN *ZáRum ‘needle’
*zikut	SBLBL, SBLT, SBLBT dikót, KPM di:kut; TDY rikut; Libon (Albay) Bikol rikót; Hanunóo rikut; Aborlan Tagbanwa dikót; Gorontalo-Mongondow hiʔuto, hikuto, rikut ‘grass’; cf. PMP *zukut and PPH (Charles) *dikut ‘grass’

REFERENCES

- Blust, Robert. 1991. The Greater Central Philippines hypothesis. *Oceanic Linguistics* 30:73–129.
- . 2008. Is there a Bima-Sumba Subgroup? *Oceanic Linguistics* 47:45–113.
- . 2009. *The Austronesian languages*. Canberra: Pacific Linguistics.
- Blust, Robert, and Stephen Trussel. In progress. Austronesian comparative dictionary. Online at <http://www.trussel2.com/acd>. Accessed June 22, 2012.
- Carro, Andres. 1956. *Iloko-English dictionary*. Translated, augmented, and revised by Morice Vanoverbergh. Manila: No publisher given.
- Headland, Thomas N., and Janet D. Headland. 1974. *A Dumagat (Casiguran)-English dictionary*. Canberra: Pacific Linguistics.
- Himes, Ronald S. 2001. The relationship of Umiray Dumagat to other Philippine languages. *Oceanic Linguistics* 41:1–20.

- . 2007. The Kalamian microgroup of Philippine languages. *Studies in Philippine Languages and Cultures* 15:54–72.
- Lambrecht, Frans Hubert. 1978. *Ifugaw–English Dictionary*. Baguio City: The Catholic Vicar Apostolic of the Mountain Province.
- Liao, Hsiu-chuan. 2008. A typology of first person dual pronouns and their reconstructibility in Philippine languages. *Oceanic Linguistics* 47:1–29.
- McFarland, Curtis D. 1977. *Northern Philippine linguistic geography*. Monograph Series No. 9. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
- Panganiban, Jose V. 1973. *Diksiyunaryo-Tesaurus Pilipino–Ingles*. Quezon City: Manglapaz Publishing Co.
- Penoyer, F. Douglas. 1986/87. Inati: The hidden Negrito language of Panay, Philippines. *Philippine Journal of Linguistics* 18/19:1–36.
- Reid, Lawrence A. 1976. *Bontok–English dictionary*. Canberra: Pacific Linguistics.
- . 1979. Towards a reconstruction of the pronominal systems of Proto-Cordilleran, Philippines. In *South-East Asian Linguistic Studies*, vol. 3, 259–75. Canberra: Pacific Linguistics.
- . 1982. The demise of Proto-Philippines. In *Papers from the Third International Conference on Austronesian Linguistics*, vol. 2, ed. by Amran Halim, Lois Carrington, and S. A. Wurm, 201–16. Canberra: Pacific Linguistics.
- . 1987. The early switch hypothesis: Linguistic evidence for contact between Negritos and Austronesians. *Man and Culture in Oceania* 3 (Special Issue): 41–59.
- . 1991. The Alta languages of the Philippines. In *VICAL 2: Papers from the Fifth International Conference on Austronesian Languages*, ed. by Ray Harlow, 265–97. Auckland: Linguistic Society of New Zealand.
- . 1994. Possible non-Austronesian lexical elements in Philippine Negrito languages. *Oceanic Linguistics* 33:37–72.
- . 2010. Historical linguistics and Philippine hunter-gatherers. In *Piakandatu ami Dr. Howard P. McKaughan*, ed. by Loren Billings and Nelleke Goudswaard, 234–60. Manila: Linguistic Society of the Philippines and SIL Philippines.
- Reid, Lawrence A., ed. 1971. *Philippine minor languages: Word lists and phonologies*. Oceanic Linguistics Special Publication No. 8. Honolulu: University of Hawai'i Press.
- Ross, Malcolm. 2005. The Batanic languages in relation to the early history of the Malayo-Polynesian subgroup of Austronesian. *Journal of Austronesian Studies* 1(2):1–23.
- Santos, Pilar C. 1975. Sinauna Tagalog: A genetic study examining its relationship with other Philippine languages. Unpublished MA thesis, Ateneo de Manila University.
- Scott, William H. 1957. *A vocabulary of the Sagada Igorot dialect*. Philippine Studies Program, Transcript No. 6. Chicago: Department of Anthropology, University of Chicago.
- Soberano, Rosa P. 1976. The dialects of Marinduque Tagalog. PhD diss., Ateneo-Philippine Normal College Consortium.
- Stone, Roger. 2008. The Sambalic languages of Central Luzon. *Studies in Philippine languages and cultures* 19:158–83.
- Tharp, James A., and Mateo C. Natividad. 1976. *Itawis–English wordlist, with English–Itawis finderlist*. HRAFlex Books OA29-001. New Haven: Human Relations Area Files.
- Tryon, Darrell T., ed. 1994. *Comparative Austronesian dictionary: An introduction to Austronesian studies*. Berlin and New York: Mouton de Gruyter.
- Tsuchida, Shigeru, Yukihiro Yamada, and Tsunekazu Moriguchi. 1987. *Lists of selected words of Batanic languages*. Tokyo: Department of Linguistics, Faculty of Letters, University of Tokyo.

- Tweddell, Colin E. 1958. The Iraya (Mangyan) language of Mindoro, Philippines. Unpublished PhD diss., University of Washington.
- Usup, Hunggu Tajuddin. 1980. Rekonstruksi Protobahasa Gorontalo-Mongodow. Unpublished PhD diss., University of Indonesia.
- Vanoverbergh, Morice. 1933. *A dictionary of Lepanto Igorot or Kankanay*. Anthropos Linguistic Monograph, No. 12.
- . 1972. *Isneg-English Vocabulary*. Oceanic Linguistics Special Publication No. 11. Honolulu: University of Hawai'i Press.
- Wimbish, John. 1986. The languages of the Zambales mountains: A Philippine lexicostatistic study. *Work Papers of the Summer Institute of Linguistics, University of North Dakota Session* 30:133–42.
- Wurm, S. A., and B. Wilson. 1975. *English finderlist of reconstructions in Austronesian languages (post-Brandstetter)*. Canberra: Pacific Linguistics.
- Yamada, Yukihiro. 1965. Phonology of Itbayaten. *Philippine Journal of Science* 94(3):373–94.
- . 1967. Fishing economy of the Itbayat, Batanes, Philippines with special reference to its vocabulary. *Asian Studies* 5(1):137–219.
- Yamashita, Michiko. 1992. *Kakilingan Sambal texts with grammatical analysis*. Cecilio Lopez Archives of the Philippine Languages and the Philippine Linguistics Circle. Quezon City.
- Yap, Fe Z. Aldave. 1977. *A comparative study of Philippine lexicons*. Manila: Institute of National Language.
- Zorc, R. David. 1974. Internal and external relationships of the Mangyan languages. *Oceanic Linguistics* 13:561–600.
- . 1986. The genetic relationships of Philippine languages. In *FOCAL II: Papers from the Fourth International Conference on Austronesian Linguistics*, ed. by Paul Geraghty, Lois Carrington, and S. A. Wurm, 147–73. Canberra: Pacific Linguistics.
- . 1987. Austronesian apicals (*dDz) and the Philippine non-evidence. In *A world of language: Papers presented to Professor S. A. Wurm on his 65th birthday*, ed. by Donald C. Laycock and Werner Winter, 751–61. Canberra: Pacific Linguistics.
- . 1994. Austronesian culture history through reconstructed vocabulary (an overview). *Austronesian terminologies: Continuity and change*, ed. by A. K. Pawley and M. D. Ross, 541–94. Canberra: Pacific Linguistics.