

The Northeastern Luzon Subgroup of Philippine Languages

Laura C. Robinson and Jason William Lobel

UNIVERSITY OF CALIFORNIA, SANTA BARBARA
AND UNIVERSITY OF HAWAI‘I AT MĀNOA

This paper presents a survey of the languages of the northeastern part of the large northern Philippine island of Luzon—Dupanigan Agta, Pahanan Agta, Casiguran Agta, Nagtipunan Agta, Dinapigue Agta, Paranan, and Kasiguranin—the first five of which are spoken by Negrito Filipino groups. With the exception of Kasiguranin, these languages compose a subgroup called Northeastern Luzon. Evidence is presented to determine the internal and external relationships of these languages, including historical phonology, functors, and lexicon. It is argued that they are not members of the Northern Cordilleran subgroup, as has been previously suggested, but instead form a primary branch of the Northern Luzon (Cordilleran) subgroup.

1. INTRODUCTION.¹ The eastern coast of northern Luzon, from Casiguran, Aurora to the small island of Palaui off the northern tip of Luzon near Santa Ana, Cagayan, is home to a world rather different from the rest of the modern Philippines. Although Ilokano migrants settled in this area during the twentieth century, the narrow strip of lowlands sandwiched between the Sierra Madre mountains and the sea has historically been home to a unique mix of Agta (that is, phenotypically Negrito²) and a minority of non-Agta Filipinos who have apparently borrowed some or most of the forms in their languages from their Agta neighbors. All of these groups speak Austronesian languages today, although the Agta must have spoken non-Austronesian languages at some point in the past, since their ancestors were present in the Philippines probably tens of thousands of years before the arrival of Austronesian-speaking peoples (cf. Reid 1987, 1994; Bellwood 1997; Blust 2005). Considering the isolation of this part of the Philippines, it should come as no surprise that—with the exception of Kasiguranin

1. Many thanks to all of our language consultants and those who facilitated our research in this region, as well as to Dr. Robert Blust and two anonymous referees whose feedback contributed to the revision of this paper. Any errors are ours alone. A complete list of abbreviations may be found in appendix 1. ISO 639-3 codes are given in square brackets immediately following the first mention of their language names.

2. The term *Negrito* is used primarily in western academic literature, and we will use it here, although we would have preferred to promote the much more neutral term “Black Filipino” to avoid preconceived notions associated with the diminutive “Negrito” (literally ‘small black person’ < Spanish), and to remind readers that, as Thomas Headland states, “these are people who have evolved right along with the rest of us into the 20th century” (1997:607).

(KAS) [ksn], the descendant of an early Tagalog dialect that borrowed heavily from Agta languages—the languages spoken in this area form a genetic subgroup, called Northeastern Luzon (NELUZ), with at least five members: Dupaninan Agta (DUP), Pahanan Agta (PAH, sometimes referred to as Palanan Dumagat or Palanan Agta), Dinapigue Agta (DIN), Casiguran Agta (CAS, sometimes called Casiguran Dumagat), and Nagtipunan Agta (NAG).³ A sixth language, Paranan (PAR), spoken by non-Agta, is tentatively included in the NELUZ subgroup due to the overwhelming lexical similarities to Pahanan Agta, although key differences in its functors point to a possible origin outside of the subgroup.

1.1 RELATIONSHIPS TO OTHER NEGRITO FILIPINO GROUPS.

A number of other Negrito Filipino groups exist in areas adjacent to where NELUZ languages are spoken (see map 1), yet none of their languages appear to have any immediate relationship to NELUZ. These neighboring languages include Central Cagayan Agta (or “Labin” Agta, CCAgTA), which belongs to the Cagayan Valley (CV) subgroup.⁴

To the immediate south are the languages Northern and Southern Alta [aqn and agy, respectively], as well as Umiray Dumagat [due]: all are spoken by Negrito Filipino populations, but none shares much with NELUZ save for some lexical items that most likely were borrowed, and a couple of phonological innovations: Low Vowel Fronting in Umiray Dumagat (with a similar phenomenon in Northern and Southern Alta), which is also found in Inagta Alabat [dul] in Quezon Province and Manide [abd] in the northern Bikol Peninsula (Lobel 2010); and *j > d in Northern and Southern Alta, Arta, Ilokano, Umiray Dumagat, Manide, Inagta Alabat, and the Greater Central Philippine languages.

A chain of Negrito Filipino groups actually continues almost uninterrupted from Dupaninan Agta in the north through Alta, Arta [atz],⁵ Umiray Dumagat, and Remontado Dumagat [agv], down through Inagta Alabat in Quezon Province, and Manide in Camarines Norte. These latter groups do not appear to share any innovations with the former groups, yet it is likely that they were at least in casual contact at some early point, as—with the mountain range serving as a formidable obstacle to land travel—travel along the eastern coast of Luzon would have had to be primarily by sea (or at least along the coast, possibly by walking), and there would have been few other groups for residents of eastern Luzon to interact with other than other Negrito Filipinos. This has been the case at least since the time of the Spanish occupation of the Philippines: “By the time of the Spanish

3. Two or three other Agta dialects are reported to exist in the inland areas of Palanan, Dinapigue, and San Mariano towns, but we were unable to reach those areas. It is expected that they are closely related to the languages covered in this study. Note that the abbreviations given in this paragraph are the ones used throughout this paper. The ISO 639-3 codes for these languages are: Dupaninan Agta [duo], Pahanan Agta [apf], Casiguran Agta [dgc], and Paranan [prf]. Dinapigue Agta and Nagtipunan Agta do not have ISO 639-3 codes but are closely related to Pahanan Agta and Casiguran Agta, respectively.

4. The Cagayan Valley subgroup includes Gaddang [gad], Itawis [itv], Central Cagayan Agta [agt], Ibanag [ibg], Atta [att], Yogad [yog], and Isnag [isd], and possibly Malaweg (no ISO 639-3 code, erroneously conflated with Itawis), with different authors varying slightly as to which languages they include.

5. Although we were unable to locate any Arta in their former site in Aglipay and Maddela towns, Quirino Province, Lawrence Reid (pers. comm., October 10, 2010) reports that he has recently discovered another group of Arta speakers in the same general area.

arrival in the archipelago, most of the Philippine Negrito groups had already been ‘minoritized’ and driven into remote areas by the Malay ethnic groups. By contrast, in the southeastern region of Luzon (present Quezon), the Aetas [that is, Negrito Filipinos] and other Negrito groups were still a majority compared to the Malay people⁶ when the Spanish first came to the area (c. 1571, according to a Spanish document)” (Goda 2003:183–84).

MAP 1. NEGRITO FILIPINO GROUPS IN AND AROUND NORTHEASTERN LUZON



6. The use of the term “Malay people” to refer to non-Negritos reflects a widespread misunderstanding in the Philippines, dating from early twentieth-century wave migration theories that claimed that the ancestors of most non-Negrito Filipinos came from Malaysia and Indonesia. Although these theories have long since been debunked by historians (for example, Scott 1984, 1992) and are contradicted by the linguistic and archaeological evidence, they persist in the Philippines and, in fact, are still being taught in the Philippines’ national elementary and high school curriculum.

1.2 HISTORICAL LINGUISTIC GEOGRAPHY OF EAST COAST NORTHERN LUZON. The NELUZ languages occupy the area between the small island of Palaui, just off the northeastern coast of Luzon, and Dinalungan town (see map 1). Until recently, this area served as a buffer zone on the east coast of northern Luzon between Tagalog and Ilokano (or earlier, Ibanag), as illustrated in map 2. Assuming that the east coast of northern Luzon was traditionally home mainly to Agta who spoke NELUZ languages, we can formulate a hypothesis about what must have taken place

**MAP 2. NORTHEASTERN LUZON AS A BUFFER ZONE
BETWEEN TAGALOG AND ILOKANO**



with the early arrival of non-Agta in the area. In the northernmost area of northeastern Luzon, there do not appear to be any long-standing non-Agta communities. To the south, people on the east coast of Luzon east from Laguna de Bay (that is, the towns of General Nakar, Infanta, Real, and Mauban) speak one of the Quezon dialects of Southern Tagalog. North of this area, however, the oldest non-Agta town is Kasiguranin-speaking Casiguran, founded in 1607. The oldest true Tagalog-speaking town in this area, Baler, was reportedly founded even more recently than Casiguran.

At present, Tagalog is spoken as the majority language as far north as Baler, Dipaculao, and Dinalungan. The native language in the town center of Casiguran is Kasiguranin (the descendant of an early Tagalog dialect that borrowed lexicon heavily from NELUZ Agta languages). To the north of Casiguran, other than Palanan town, are four small, primarily Ilokano-speaking towns—Dilasag, Dinapigue, Maconacon, and Divilacan—which do not appear to be particularly old as non-Agta settlements (and which, with the exception of Dilasag, have less than 5,000 residents each, making them the three least-populated coastal towns on any major island in the entire Philippines).⁷ The town of Casiguran is probably one of the oldest non-Negrito Filipino settlements on the entire east coast of northern Luzon. Kasiguranin is the result of an early Tagalog dialect that borrowed a large amount of lexicon from the surrounding Agta languages. Paranan, the language of the non-Agta residents of Palanan town, is similar to the Pahanan language spoken by the Agta of Palanan town, save for the *r > /h/ change in Pahanan Agta (an areal feature in the Agta languages of northeastern Luzon), a small percentage of lexical differences, and a non-Agta substratum in Paranan.⁸

It appears that the first non-Agta Filipinos to permanently settle in the Casiguran area were Tagalogs migrating northward from Quezon Province,⁹ who traveled up the coast, finding Casiguran, likely already a locally important, relatively large Agta settlement whose bay is conveniently protected from storms by the San Ildefonso Peninsula. With relatively few other non-Agta to interact with in their new home, the early Casiguran Tagalogs began to borrow heavily from the language of the Casiguran Agta (which itself was the descendant of a language that had been adopted from non-Agta Austronesians centuries or millennia earlier). As they still would have had some communication with Tagalogs to the south, the Casiguran Tagalogs had enough need for Tagalog that they retained a significant portion of their native language, while borrowing considerably from Casiguran Agta and the other local Agta languages. The resulting Kasiguranin language is a unique mixture that, after several generations, became incomprehensible to Tagalogs and, until recently, was reportedly incomprehensible even to the Casiguran Agta whose ancestors contributed approximately half of the Kasiguranin lexicon. The first author calculated 68 percent cognates between Casiguran Agta and Kasiguranin based on a 280-

7. Note, however, that the ancestors of the modern Agta groups have almost certainly been in the region for much longer, and the fact that the names of many of these towns, as well as of the barangays thereof, begin with the Agta locative marker *di* suggests that they were probably founded on the sites of earlier Agta settlements.

8. Less than 2 percent of non-Tagalog, non-Agta lexicon is shared exclusively between Kasiguranin and Paranan, implying that there was little if any significant contact between the non-Agta residents of Paranan town and those of Casiguran town that did not also include considerable numbers of Agta.

9. Until 1979, Aurora Province was a part of Quezon Province, which was previously called Tayabas.

word list (see table 1), while Headland (1975) calculated 77 percent cognates. Kasiguranin data are included in this paper for the sake of comparison, even though it is not a member of the NELUZ subgroup. The authors are preparing a separate paper on this subject. (Lobel and Robinson n.d.). Like Kasiguranin, the Paranan language is spoken by non-Agta, but Paranan is a genetic member of the NELUZ subgroup along with the five Agta languages. The first author calculated 81 percent cognates between Pahanan Agta and Paranan (see table 1), while Headland (1975) found 85 percent cognates.

TABLE 1. COGNATE PERCENTAGES

PAH	65%				
PAR	54%	81%			
CAS	60%	78%	71%		
NAG	58%	72%	60%	77%	
KAS	49%	72%	78%	68%	55%
	DUP	PAH	PAR	CAS	NAG

1.3 PREVIOUS REFERENCES TO THE NORTHEASTERN LUZON LANGUAGES. A number of previous works have investigated the subgrouping of one or more of the NELUZ languages, but often with very little data, and with as few as only one of these languages included.

Tharp (1974a) classifies Casiguran Agta as a primary branch of Northern Cordilleran, but does not include any of the other NELUZ languages in his study. Headland (1975) considers Casiguran Agta, Kasiguranin, Pahanan Agta, Paranan, and Dupanangan Agta on the bases of shared cognate percentages and intelligibility testing, and is thus the first author to define a NELUZ subgroup similar to the one we are proposing. However, the methodology used is problematic, as (i) lexicostatistics does not differentiate retentions, borrowings, and shared innovations, which is especially important in the case of contact languages like Kasiguranin and Paranan; and (ii) intelligibility testing has never been demonstrated to be a reliable basis for subgrouping, and can give misleading results when used to determine the relationships between the languages of two communities that have always historically interacted and traded with one another, as is the situation with Kasiguranin and Casiguran Agta. Furthermore, as Blust (2000:327) points out, “we cannot tell when [lexicostatistics] gives valid results and when it does not. . . . Since we now know that languages vary widely in retention rate of basic vocabulary over lengthy intervals of time, lexicostatistics must be seen as an unreliable foundation for subgrouping hypotheses that are not independently confirmed by the evidence of exclusively shared innovations.”

Most modern works (for example, Reid 2006, Lewis 2009) assume that the NELUZ languages form a subgroup within Northern Cordilleran, which is in turn a subgroup of Northern Luzon (Cordilleran),¹⁰ but there are no published works justifying this subgrouping.

Of the NELUZ languages described in this paper, only Casiguran Agta (Headland and Headland 1974, Headland and Healey 1974, Headland and Wolfenden 1967) and Dupanangan Agta (Robinson 2011) have been described in any detail. Vanoverbergh

10. This subgroup was previously referred to as “Cordilleran,” but we have adopted the name “Northern Luzon” to reflect the current usage of Lawrence Reid, the leading scholar on the languages of the northern Philippines.

(1937) includes wordlists of Kasiguranin (his “Casiguran”) and Casiguran Agta (his “Casiguran Negrito”). Pahanan Agta and Paranan are included in *Ethnologue* and other surveys, although there are no published data available on either. The Nagtipunan Agta and Dinapigue Agta varieties discussed here have not been previously mentioned in the literature.

1.4 THE DATA AND ELICITATIONS. The data used in this study were collected in September and October of 2006 by the current authors. For each speech variety, a 1,000-item wordlist was elicited along with between 100 and 200 sentences, usually with a small group of native speakers. Much less data were elicited for Dinapigue Agta, for which the authors were only able to spend one afternoon with a single native speaker. Data were usually elicited via Tagalog (which was said to be better understood by the residents of these areas than Ilokano), except for Dupaningan Agta, where the medium was Ilokano, and Nagtipunan Agta, where a mixture of Tagalog and Ilokano was used (since Ilokano is much more prevalent on the western side of these mountains than on the eastern side).

2. THE EVIDENCE FOR A NORTHEASTERN LUZON SUBGROUP.

As mentioned above, previous authors have argued or assumed that most or all of the languages of northeastern Luzon form a single subgroup, but without presenting phonological, morphological, or lexical evidence to support this subgrouping. The goal of this section is to present such evidence.

2.1 PHONOLOGICAL EVIDENCE. This section will outline the various phonological shifts that have occurred in the NELUZ languages. When they are similar to the NELUZ forms, Kasiguranin forms are given in parentheses for the sake of comparison.

2.1.1 Consonant reflexes. Table 2 illustrates the phoneme inventory of Proto-Northeastern Luzon as we reconstruct it.

TABLE 2. PROTO-NELUZ PHONEME INVENTORY

CONSONANTS				VOWELS		
*p	*t	*k	*ʔ	*i		*u
*b	*d	*g		*e	*ə	*o
	*s		*h		*a	
*m	*n	*ŋ				
	*l					
	*r					
*w	*y					

2.1.1.1 Reflexes of PMP *R. PMP *R is reflected as /g/ in the NELUZ languages, as illustrated in forms (1)–(9).¹¹

- (1) *qaRta ‘outsiders, alien people’ > DUP, PAR, CAS, NAG ágta? (KAS ágta?)
 ‘Negrito Filipino person’¹²

11. Reconstructions are from Blust (1972, 1983–84, 1987, 1999, 2006, 2009) and Blust and Trussell (in progress) unless otherwise noted. In all numbered examples in this paper (except for [95]–[121]), all protoforms are Proto-Malayo-Polynesian (PMP) unless otherwise specified.

- (2) *qabaRa ‘shoulder’ > DUP abigí?, PAR abagá? (KAS abagá?)
- (3) *daRaQ ‘blood’ > PAH, CAS digé?, DUP, NAG digí? (KAS digí?)
- (4) *daRaT ‘littoral sea’ > DUP, PAR, CAS digét, PAH, NAG digit (KAS digét) ‘sea’
- (5) *baqəRu ‘new’ > DUP, PAH, PAR, CAS, NAG bigú? (KAS bigú?)
- (6) *diRuS ‘bathe’ > DUP dégus, PAH, PAR, CAS, NAG dígus (KAS dígus)
- (7) *hulaR ‘snake’ > DUP, PAH, CAS, NAG ulág (KAS ulág)
- (8) *niuR ‘coconut’ > DUP, PAH, PAR, DIN, CAS, NAG niyóg (KAS niyóg)
- (9) *bəŋəR ‘deaf’ > DUP bəŋŋág, PAH, PAR, DIN bəŋŋóg

The /g/ reflex of PMP *R in both the Cagayan Valley and Northeastern Luzon languages is one reason why previous authors have suggested that they subgroup together. The NELUZ languages share the *R > /g/ shift with the Cagayan Valley languages, but it is worth noting that languages further south on the Pacific coast of Luzon also reflect *R as /g/, including Manide (Lobel 2010), Inagta Alabat (Lobel 2011), and Central Philippine languages like Tagalog and Bikol.

2.1.1.2 Reflexes of PMP *j. PMP *j is reflected as /d/ in the NELUZ languages, with some exceptions. Among the other Northern Luzon languages, a /d/ reflex of *j only occurs in Arta, Northern and Southern Alta, and Ilokano (Reid 2006:5–6). Elsewhere in the Northern Luzon subgroup—that is, Cagayan Valley and South-Central Cordilleran—PMP *j is reflected as /g/. Forms (10)–(18) illustrate the /d/ reflex of *j in the Northeastern Luzon languages.

- (10) *pusəj ‘navel’ > DUP pusád, PAH, PAR, CAS, NAG pusád (KAS pusád)
- (11) *huaji ‘younger sibling’ > DUP, PAH, PAR, CAS, NAG wadí? (KAS wadí?)
- (12) *qapəju ‘bile, gall’ > DUP, PAH, PAR, CAS ápdu?, NAG apdú?
- (13) *ə[n]juŋ ‘nose’¹³ > PAR əddúnj (KAS əddúnj)
- (14) *qaləjaw ‘day’ > DUP, PAR, CAS, NAG aldéw DIN, PAH áldew (KAS aldéw) (also ‘sun’ in DUP, PAH, PAR, DIN, KAS)
- (15) *quləj ‘maggot’ > DUP, PAR urád, PAH, CAS uhád (KAS urád) ‘worm’ (PNELUZ *urəd)¹⁴
- (16) PAN *maja ‘dry’ > DUP, PAH, PAR, DIN, NAG madi?
- (17) *Rawəj ‘betel leaf’ > DUP gíwad, PAH, PAR giwád, DIN, CAS, NAG gawád (KAS gawád)
- (18) *laja ‘weave’ > DUP, PAH, PAR, CAS ladí? (KAS ladí?)

In several forms, such as (19)–(23), *j is reflected as /t/ in the NELUZ languages, or as /t/ in some languages and as /g/ in others. This could reflect a sporadic shift of *j > *d > /t/, but independent evidence for the intermediate step of *d > /t/ is lacking.

12. Glottal stop is not written word-initially because it is not phonemic in this position in any of the NELUZ languages.

13. This etymology was attributed to Proto-Western Malayo-Polynesian (PWMP), but since this is not a widely accepted subgroup, we have regularized this and other similar etymologies in this paper to PMP.

14. Note that the Agta languages (except Dupanangan) reflect the shift *r > h.

- (19) *ŋajan ‘name’ > DUP nagen, PAH ŋahán, PAR ŋarán, CAS, NAG ŋahén (KAS ŋarán)
- (20) *sujud ‘fine-toothed comb for delousing’ > DUP, NAG súgod, DUP, PAH, DIN, CAS suród, PAR surúd (KAS suród) (note DUP suród ‘comb’ vs. súgod ‘fine-toothed comb for delousing’)
- (21) *pajay ‘rice in field’ > DUP, PAR paráy, PAH, CAS, NAG paháy (KAS paráy)
- (22) *qikəj ‘cough’ > DUP ikár, PAH, DIN, CAS, NAG ikáh, PAR ikár (KAS ikár)
- (23) PNLUZ *pajəs ‘wind’ > DUP parás, PAH pahás, PAR parəs, CAS, NAG pahəs (KAS parəs)

There is at least one item, (24), in which *j appears to be reflected as /g/ in all of the NELUZ languages for which we have data.

- (24) *qujiŋ ‘charcoal’ > DUP, PAH, PAR, CAS, NAG ugiŋ (KAS ugiŋ)

However, David Zorc (pers. comm., August 31, 2011) suggests that there is evidence for a northern Philippine doublet *u:Riŋ alongside PMP/PPH *qujiŋ, so these reflexes would be regular.

Although the Cagayan Valley languages generally have a /g/ reflex of *j, there are also a few exceptions in which these languages have a /d/ or /t/ reflex, such as forms (25)–(30).

- (25) *palaj ‘palm (of hand)’ > Ilokano palad, Gaddang pallad
- (26) *huaji ‘younger sibling’ > Ilokano adi
- (27) *pusəj ‘navel’ > Central Cagayan Agta pusad (cf. DUP pusád)
- (28) *qapəju ‘bile, gall’ > Central Cagayan Agta apdu (cf. DUP apdí?)
- (29) *qikəj ‘cough’ > Central Cagayan Agta ikar (cf. DUP ikár)
- (30) *qaləjaw ‘day’ > Central Cagayan Agta araw (cf. Tagalog araw)¹⁵

Reid (2006:9) notes that the reflex of *palaj has a final /d/ in all of the languages of the northern Philippines, so it seems likely that this was either a borrowing or a sporadic shift in the protolanguage. Central Cagayan Agta, especially the Gattaran dialect from which forms (27)–(30) are taken, has been in extensive contact with Dupanangan Agta, so it is likely that forms (27)–(29) have been borrowed from the latter. Form (30) looks like a borrowing from Tagalog.

Thus, all of the NELUZ languages share the shifts *R > /g/ and *j > /d/, a combination of changes that is not found in any other language in the northern Philippines (Ilokano and Arta have /d/ reflexes of *j, but *R split to /g/ and /t/ without apparent conditioning in both languages; see section 3), but which is found further south in Manide, Inagta Alabat, and the Greater Central Philippine languages. Given the importance previous authors have placed on *R and *j for subgrouping purposes, this combination of phonological innovations is considered strong evidence supporting the existence of a NELUZ subgroup distinct from the other branches of Northern Luzon.

15. Thanks to an anonymous reviewer for bringing to our attention these forms from Mayfield’s (n.d.) unpublished dictionary of the Gattaran dialect of Central Cagayan Agta, which has been in close contact with Dupanangan Agta.

2.1.1.3 Reflexes of PMP *z. PMP *z is regularly reflected as /d/ in the NELUZ languages, as illustrated in forms (31)–(35). Thus, PMP *z, *j, and *d all merged as /d/ in these languages.

- (31) *zalan ‘road, path’ > DUP, PAH, PAR, CAS, NAG dilán (KAS dalán)
- (32) *zaRum ‘needle’ > DUP dágum, PAH, PAR, CAS, NAG digúm (KAS digúm)
- (33) *tazəm ‘sharp’ > DUP tadám, PAH, CAS, NAG tadəm
- (34) *quzan ‘rain’ > DUP, PAR, NAG udén, PAH udín, CAS udén (KAS udén)
- (35) *azani ‘near’ > PAH, PAR, adéni? (KAS adéne)

2.1.1.4 Reflexes of PMP *r. PMP *r is reflected as /l/ or /r/ in the NELUZ languages, without apparent conditioning, much like the Cagayan Valley languages, which also reflect a sporadic *r > l shift (Tharp 1974a). Evidence to support this is limited, however, because most of the PMP etymologies with *r have relatively obscure meanings that are not included on the second author’s 1,000-item wordlist. Therefore, we have largely had to rely on the more extensive lexical data available for Dupanigan Agta (Robinson 2011) and Casiguran Agta (Headland and Headland 1974). Forms (36)–(40) illustrate the /r/ reflex of *r:

- (36) *rəbaq ‘to collapse (house)’ > PAH nahabá?, PAR narbá? (from *marəbba?)
- (37) *rəpuk ~ *rəput ‘rotten to brittleness, of wood’ > DUP rəpótak ‘rotten, of wood’
- (38) *burik ‘speckled, of hens, etc.’ > DUP burék ‘spotted, striped, speckled, multi-colored’
- (39) *qarimaw > DUP sarimaw ‘civet cat’
- (40) PPH *warak > DUP warék ‘scatter’; CAS wahak-wahak ‘to be continually dropping large pieces of things being carried over the trail (e.g., sweet potatoes, pencils, books)’ (Headland and Headland 1974)

Compare these with forms (41) and (42), which illustrate the /l/ reflex:

- (41) *rəpag > DUP ləpág ‘slap’
- (42) PPH *haprus > DUP əplos ‘massage, rub’

2.1.1.5 Reflexes of PMP *q. PMP *q was usually lost in NELUZ, as illustrated in forms (43)–(46).

- (43) *baqəRu ‘new’ > DUP, PAH, PAR, CAS, NAG bigú? (KAS bigú?)
- (44) *taqi ‘feces’ > DUP, PAH attáy, PAR, NAG əttáy, CAS ətáy ~ əttáy (KAS əttáy)¹⁶
- (45) *tuqəlanj ‘bone’ > DUP, PAH, PAR, CAS, NAG tulánj (KAS tulánj)
- (46) *bituqən ‘star’ > DUP, PAH, PAR, CAS bitón, NAG bitún

In at least one case, (47), some of the languages do show a medial glottal stop. More data are needed to determine whether this is an exception or reflects some as yet undetermined regularity.

16. After the loss of medial *q in this form, a schwa was epenthesized to preserve the preferred disyllabic syllable structure.

- (47) *liqəR ‘neck’ > DUP le:g, PAH laʔég, PAR allig, DIN liʔég, CAS, NAG li:g (KAS əllég)

2.1.1.6 Reflexes of PMP *h. PMP *h was lost in initial position, as illustrated in forms (48)–(51).

- (48) *hadiRi ‘house post’ > DUP, PAH, PAR, CAS, NAG adigiʔ (KAS adigiʔ)
 (49) *huRas ‘wash’ > DUP, PAH, CAS, NAG ugés, PAR ugás (KAS ugás) ‘wash (general, or of hands)’
 (50) *hulaR ‘snake’ > DUP, PAH, CAS, NAG ulág (KAS ulág)
 (51) *huaji ‘younger sibling’ > DUP, PAH, PAR, CAS, NAG wadiʔ (KAS wadiʔ)

In medial position, however, PMP *h is sometimes reflected as a glottal stop in PAH, PAR, DIN, and CAS, while it disappears in NAG and DUP, as illustrated in forms (52)–(54).

- (52) *buhək ‘hair’ > DUP, NAG bu:k, PAH, PAR buʔók, DIN, CAS buʔúk (note that KAS *buhók* continues the Tagalog form)
 (53) *dahun ‘leaf’ > DUP, NAG do:n, PAH dáʔon, PAR dúʔun, CAS dəʔón (KAS duʔón)
 (54) *bahaR ‘loincloth’ > DUP be:g, PAH, PAR, DIN baʔeg, CAS, NAG bi:g (KAS baʔeg)

On the other hand, there are quite a few words in which *h was lost in all of the NELUZ languages, such as (55)–(61).

- (55) *kahiw ‘wood’ > DUP, PAH, PAR, CAS, NAG, kayúʔ (KAS kayóʔ)
 (56) *qihu ‘shark’ > DUP, NAG iyúʔ, PAH ayyúʔ, CAS iyóʔ (KAS iyúʔ)
 (57) *duha ‘two’ > DUP, PAH, PAR, DIN dúwa, CAS əduwá (KAS dúwa)
 (58) *unahik ‘climb’ > DUP, PAH, PAR, DIN, CAS, NAG unék (KAS unék)
 (59) *anahaw ‘palm tree’ > DUP, PAH, PAR, DIN, CAS anáw (KAS anáw)
 (60) *di lahud ‘downstream’ > DUP, PAH, CAS, NAG dilód (KAS dilód)
 (61) *luhaq ‘teardrop’ > DUP, PAH, PAR luwáʔ, CAS, NAG lówa

2.1.1.7 *s > /h/. The sporadic and unconditioned shift of *s to /h/ is found in all members of the NELUZ subgroup, except Casiguran Agta.¹⁷ Table 3 lists selected forms in which there is variation between /s/ and /h/ in the Northeastern Luzon languages. The *s > /h/ forms are shaded, and dashes indicate that no cognate was found in that language.

TABLE 3. VARIATION BETWEEN /s/ AND /h/

	DUP	PAH	PAR	DIN	CAS	NAG
‘thick bamboo’	—	buwaŋhina	buwaŋhina	—	buwaŋsina	—
‘lie on stomach’	hakab	—	—	hakəb	sakəb	hakəb
‘close eyes’	kisap	kihəp	kisəp	kihəp	kisəp	—
3PL.NOM	hidi	hide	hidi	hide	side	sidi
1INCL.DUAL.NOM	hikitá	hikitá	sikitá	hikitá	sikitá	sikitá
‘hiccup’	hálduʔ	sálduʔ	sálduʔ	sálduʔ	sálduʔ	sálduʔ

17. Since most of Kasiguranin’s non-Tagalog vocabulary is from Casiguran Agta, this shift is not found in Kasiguranin either.

2.1.2 Vowel reflexes. PNELUZ had a six-vowel system /a e i o u ə/, as illustrated in table 2 above. PMP *i, *u, and *ə were reflected as *i, *u, and *ə, respectively, in Proto-Northeastern Luzon. PMP *a was reflected as PNELUZ *a in most environments, but as *e after a voiced stop /b d g/, due to Low Vowel Fronting (cf. 2.1.2.2). The source of PNELUZ *o is unclear. This section will outline the various shifts affecting vowels in the NELUZ languages.

2.1.2.1 Schwa and gemination. In all of the NELUZ languages except Casiguran Agta, consonants geminate after a schwa, which has been characterized as “inherently short” (Blust 2009:548), as illustrated in forms (62)–(68). We do not have enough data to determine whether this gemination also occurred in Nagtipunan Agta.

- (62) *təlu ‘three’ > DUP, PAH, PAR, DIN talló, CAS ətəlu (KAS təlló)
 (63) *əpat ‘four’ > DUP, PAR appát, PAH, DIN əppát, CAS əpát (KAS əppát)
 (64) *ənəm ‘six’ > DUP ánnam, PAH ənnəm, PAR annám, CAS ənəm (KAS ənnəm)
 (65) *bəŋəR ‘deaf’ > DUP bəŋŋág, DIN, PAH, PAR bəŋŋég
 (66) *qəlad ‘wing’ > DUP, PAH, NAG allád ‘feather’
 (67) *ləsuŋ ‘mortar’ > DUP, PAH, PAR lassóŋ, CAS lúsoŋ, NAG lasúŋ (KAS ləssóŋ)
 (68) *təkən ‘pole, usually of bamboo, used to propel a boat or raft’ > DUP takkán, PAH, PAR təkán, CAS təkən (KAS təkən)

Occasionally, gemination does not occur where it is expected, as in the Dupanangan Agta reflex of PMP *bəRas ‘uncooked rice’ and the Pahanan Agta and Paranan reflexes of PMP *kəzut ‘pinch’, examples (69) and (70), respectively.

- (69) *bəRas ‘uncooked rice’ > DUP bagáh, PAH, PAR baggés, DIN baggés, CAS, NAG bəgís, (KAS bəgés ~ bəggés)
 (70) *kəzut ‘pinch’ > DUP kaddút, PAH, CAS kədút, PAR, NAG kadút

Gemination after the schwa also occurs in the Cagayan Valley languages and Ilokano (see, for example, Tharp 1974a), as well as in a number of other Philippine languages (for example, some Manobo and Sama-Bajaw languages, and phonetically in Maranao). It is likely that this was also a phonetic feature of PNELUZ but was subsequently lost in Casiguran Agta. There is even some evidence that gemination after the schwa was still present in early twentieth-century Casiguran Agta, as it appears that Vanoverbergh (1937) documented geminate variability in Casiguran Agta on his trip through northern Luzon in 1936 and 1937, such as his *kəttihek* ‘small’ and *ənnám* ‘six’, which we recorded as *kəttihek* and *ənəm*, respectively. In most cases, however, Vanoverbergh recorded a singleton following a schwa. In our own data, there was at least one instance of a geminate alternating with a singleton in Casiguran Agta, *ətáy* ~ *əttáy* ‘feces’. Note also that gemination after the schwa does occur in the Kasiguranin forms, suggesting that Kasiguranin borrowed these forms from Casiguran Agta before the loss of gemination in the latter language.

2.1.2.2 Low Vowel Fronting. Low Vowel Fronting (LVF), the shift of /a/ to a mid or high front vowel after a voiced stop, occurs sporadically in all of the NELUZ languages.¹⁸

18. See Blust (2000) for a discussion of this phenomenon in Sarawak.

This process is found in all of the NELUZ languages, as well as in Southern Alta (Reid 1991),¹⁹ Umiray Dumaget (Himes 2002, Lobel 2012), Manide (Lobel 2010), and Inagta Alabat (Lobel 2011). Interestingly, this shift is not found in the same words in all languages, and there is no apparent conditioning to explain which words will have LVF and which will not. For some forms, LVF is found in all of the NELUZ languages, while for others, no language reflects LVF. For many other forms, LVF is found in some languages but not others. Note that all the LVF forms in KAS are also found in CAS, suggesting that LVF may not have ever existed as a process in KAS, but was simply the result of lexical borrowing. All but one of the LVF forms in PAR are also found in PAH, suggesting that LVF in PAR is also primarily the result of lexical borrowing, but also reflecting the greater influence of Agta languages on PAR than on KAS. Examples of LVF in NELUZ languages are given in table 4, where the cells containing forms reflecting LVF are shaded. Kasiguranin forms are given in the last column for comparison. Since we have much less data for Dinapigue Agta, we have included it in the Pahanan Agta column and noted where the Dinapigue forms are the same (=D) or different. Where no Dinapigue form is listed, we lack data.

2.2 MORPHOLOGICAL EVIDENCE. This section presents an analysis of the pronouns and case markers of the NELUZ languages and discusses their implications for subgrouping. Pronouns and case markers, as well as other types of functors, often serve as strong evidence for subgrouping (cf. Zorc 1977, 1978; McFarland 1974; Lobel 2012), especially where lexical evidence is ambiguous or even misleading due to heavy borrowing from more prestigious languages. For example, Zorc (1978:510) argues that “a language is more readily defined by its grammar than by its lexicon,” and that functors have “obvious importance within any given speech variety” due to their “high text frequency” and a “tendency towards stability and a low rate of replacement.” In most instances, the Kasiguranin forms derive primarily from the Tagalog substratum, and so were not used for reconstructing PNELUZ. Note that the full sets of pronouns, case markers, and other functors are presented in appendix 2.

2.2.1 Pronouns. Table 5 presents the reconstructions of the PNELUZ pronouns. A few general comments are in order. First, note that the bases are the same for the Topicalized Nominative and the Oblique sets, with the PMP Nominative bases having replaced the PMP Oblique bases, which is not at all uncommon in Philippine-type languages (see Lobel 2012). The Topicalized Nominatives are formed by attaching *si- to the base, while the shorter Nominative pronouns are usually an enclitic form of the base, with the exception of 1SG and 2SG, which have different forms, and also 3SG and 3PL, which are identical to the Topicalized forms and do not appear to be enclitic. The Genitive set is composed of enclitic monosyllabic pronouns largely identical to the reconstructed PPH set.

Various phonological shifts have affected the pronouns. The sporadic shift of *s > h in some of the languages was discussed in 2.1.1.7. Second, Low Vowel Fronting has affected the 3PL form, resulting in the base *-di from earlier *-da. Third, monophthongization has taken place in the 2SG.TOP/OBL base *-kaw, resulting in the base *-ko in all of the languages except Dupanigan Agta. Fourth, the *a of the base *akən ‘1SG’ is lost in

19. In Northern Alta, *a became /ə/ rather than /i/ or /e/ in this environment.

all but the southern languages (Casiguran Agta and Nagtipunan Agta), resulting in the form *hikən* from PNELUZ *siyakən. In CAS and NAG, on the other hand, the vowel of Topicalized formative *si- is lost in the 1SG.TOP form, resulting in the form *sakən*. In the Oblique form, however, neither segment is lost in CAS and NAG, resulting in the form *diyakən*. It should also be noted that reduction in the 1SG.TOP/NOM form and not in the other persons is common throughout the Philippines (cf. Lobel 2012).

TABLE 4. LOW VOWEL FRONTING IN NELUZ LANGUAGES

PMP		DUP	PAH, DIN	PAR	CAS	NAG	KAS
*baqəRu	‘new’	bigú	bigú	bigú	bigú	bigú	bigú
*quzan	‘rain’	udén	udín	udén	udén	udén	udén
*daRaQ	‘blood’	digí?	digé?	digí	digé?	digí?	digí
*daRat	‘sea’	digét	digit	digét	digét	digit	digét
*qaləjaw	‘day, sun’	aldéw	áldew	aldéw	aldéw	aldéw	aldéw
*haRəzan	‘ladder, stairs’	ágden	agdénan	agdénan	agdénan	ágen	agdénan
*tabəq	‘fat’	tabí?	tabí?	tabí?	tabí?	tabí?	tabí?
*laja	‘weave’	ladí?	ladí?	ladí?	ladí?	—	ladí?
*azani	‘near’	—	adéni?	adéni?	—	—	adéne
*maja	‘dry’	madí?	madí?	madí?	—	madí?	—
*balay	‘house’	biláy	biláy (=D)	biláy	bile	—	baláy
*danum	‘water’	dinóm	dinúm	dinóm	dinúm	dinúm	danóm
*zalan	‘road’	dilán	dilán	dilán	dilán	dilán	dalán
*balu	‘widow’	bilú?	biló?	bilú	biló?	bilú?	baló?
PNELUZ *madəggá?	‘heavy’	daggi?	daggi?	dəggi	dəgi?	dəggi	dəggá
*basəq	‘wet’	bésa	bisá?	bisá?	bisá?	bisá?	basá
*sida	‘3PL.NOM’	hidí	hide	hidi?	side	sidi?	silá?
*dakəp	‘catch’	dakkap	dikəp	dikəp	dikəp	dikəp	dakəp
*zaRum	‘needle’	dágum	digúm	digúm	digúm	digúm	digúm
*qabaŋ	‘boat’	abán	abiŋ	abéŋ	abéŋ	abiŋ	abéŋ
*bəRas	‘uncooked rice’	bagah	baggés	baggés	bəgis	bəgis	bəg(g)és
*dakəl	‘large’	dakal	dakəl (=D)	dikəl	dikkəl	dikkəl	dikkəl
*huRas	‘wash’	ugés	ugés	ugás	ugés	ugás	ugás
*Rawəd	‘betel leaf’	giwad	giwəd	gawəd	gawəd	gawəd	gawəd
*batuh	‘stone’	bitú?	bitú? (D bitó?)	batú	bitú?	bitú?	báto
*balun	‘provisions’	bilón	bilún (D bilón)	balón	bilón	bilón	balón
*Ramut	‘root’	ramót (< ILK)	gimút (=D)	gamót	gimót	gimút	gamót
*bayu	‘pound rice’	biyu?	báyyu? (D bayú?)	báyo	biyú?	biyú?	bayó
*gatəq	‘coconut milk’	gittá?	gatá?	gatá?	gatá?	gatá?	gatá?
*bahaR	‘loincloth’	beeg	baʔeg	baʔeg	biig	biig	baʔeg
*qabaRa	‘shoulder’	abigi?	—	abagá?	—	—	abagá?
*dahun	‘leaf’	doon	dáʔon	dúʔun	dəʔón	doon	dúʔón
*baRiuh	‘storm’	bágyo	bágyo	bágyu	bágyo	bágyo	bágyo
PPH *sidaq	‘main course’	hidá	sidá	sidá	sidá	sidá	sidá

TABLE 5. PROTO-NORTHEASTERN LUZON PRONOUNS

	TOP	NOM	GEN	OBL	NOM & OBL BASE
1SG	*si-akən	*=ək	*=ku	*ni/di-akən	*-akən
2SG	*si-kaw	*=ka	*=mu	*ni/di-kaw (>*ko)	*-[i]kaw
3SG	*si-ya	*si-ya	*=na	*di-ya; *ni/di-ko-na*	*-iya
1EXCL	*si-kami	*=kami	*=mi	*ni/di-kami	*-kami
1INCL.DU	*si-kita	*=kita	*=ta	*ni/di-kita	*-kita
1INCL.PL	*si-kitam	*=kitam	*=tam	*ni/di-kitam	*-kitam
2PL	*si-kam	*=kam	*=muy	*ni/di-kam	*-kam
3PL	*si-di	*si-di	*=di	*ni/di-di	*-[i]di

* CAS and NAG retain PPH *diya.

DUP, PAH, and DIN innovated the 3SG.OBL form *nikuna*, which appears to be analyzable as Oblique formative *ni-, 3SG.GEN base *-na, and a medial segment *-ku-. The exact origin of this medial segment *-ku- is unclear, but a similar medial segment *-ko- followed by the genitive base is found in the oblique pronouns of Paranan and one of the three competing oblique pronoun sets of Kasiguranin (see table 6). Dupaningan Agta further innovates the Topicalized Nominative form *hikuna* ‘3SG.NOM’, consisting of the Nominative formative *hi- (< earlier *si-) plus the same base *-kuna consisting of the combination *-ku- + *-na ‘3SG.GEN’ as found in the oblique form *nikuna*.²⁰

Paranan has a quite different oblique set analyzable as oblique formative *di-, followed by *-ko- plus the genitive base (see tables 5 and 6).

Tables 7a–d list the PNELUZ, Proto-Northern Luzon (Reid 1979a), and Proto-Northern Cordilleran²¹ pronouns (PNCORD, Tharp 1974a). The Proto-Central Cordilleran (PCCORD, Reid 1974, 1979a) and Proto-Southern Cordilleran (PSCORD, Reid 1979a, Reid 2009) forms are also included for comparison, where available.

As can be seen in tables 7a–d, the reconstructed PNELUZ pronouns have changed relatively little from the Proto-Northern Luzon forms reconstructed by Reid (1979a), and there are very few if any exclusively shared innovations with Proto-Northern Cordilleran (that is, Proto-Cagayan Valley, if the NELUZ languages are not included). If the PNELUZ

TABLE 6. OBLIQUE PRONOUNS IN PARANAN AND KASIGURANIN

	PARANAN		KASIGURANIN		PNELUZ	PNELUZ
		“Tagalog” set	*-ko- set	*kaʔo- set	GEN	OBL
1SG	dikóku?	sa ákin	sakóko?	kaóko	*=ku	*ni/di-akən
2SG	dikómu?	sa iyú?	kómo?	kaómo?	*=mu	*ni/di-kaw (>*ko)
3SG	dikóna?	sa kanyá?	kónya, koniyá?	—	*=na	*ni/di-ko-na, SNEL *diya
1EXCL	dikómi?	sa ámin	sakóme?	—	*=mi	*ni/di-kami
1INCL.DU	dikóta?	sa áta?	sakóta?	kootá?	*=ta	*ni/di-kita
1INCL.PL	dikótam	sa átam	sa kótam	—	*=tam	*ni/di-kitam
2PL	dikómoy	sa inyo	sa ómoy/ komóy	—	*=muy	*ni/di-kam
3PL	dikódi?	sa kanilá	—	kaónila, sakaóde?	*=di	*ni/di-di

20. Note that Yogad (cf. fig. 2) has *takuna* ‘3SG.OBL’, *takura* ‘3PL.OBL’, *akuna* ‘3SG.POSS’, and *akura* ‘3PL.POSS’.

21. Note that Tharp included Casiguran Agta in his Northern Cordilleran subgroup, which likely influenced his reconstructions.

pronouns do derive from PNLUZ forms as reconstructed by Reid (1979a), then there appears to have been a reduction in the topicalized formative of *siqi- to *si-, which, according to Tharp (1974a), also took place in NCORD. Note, however, that Topicalized Nominatives beginning with a simple *si- (and, therefore, identical to the reconstructed PMP nominative case marker *si) are found in many Philippine subgroups, and Reid's PNLUZ reconstructions appear to have double marking, with *si- followed by *qi-. It

TABLE 7. PRONOUN FORMS

	PNELUZ	PNCORD	PNLUZ	PCCORD	PSCORD
a. LONG-FORM NOMINATIVE PRONOUNS					
1SG	*si-akən	*si akən	*siyakən	*siyakən	*siyak
2SG	*si-kaw (>*ko)	*si kaw	*siqikaw	*siqika ~ *sikqa	*siqika
3SG	*siya	*iya, *V[n]su	*siya	*siya	*siya
1EXCL	*si-kami	*si kami	*siqikami	*dakami ~ *dikami	*siqikami
1INCL.DU	*si-kita	*si kita	*siqikita	*daqita ~ *dita ~ *data	*siqikita
1INCL.PL	*si-kitam	*si kitam	*siqikitam	*datakayú ~ *dita-kayú*	*siqikitayú
2PL	*si-kam	*si kamu, *si kayu	*siqikamuyu	*dakayu ~ *dikayu	*siqikayu
3PL	*sidi	*ida	*siqida	*daqida ~ *dida	*siqida
b. SHORT-FORM NOMINATIVE PRONOUNS					
1SG	*=ək	*ak	*-ak	*-ak	
2SG	*=ka	*ka	*-ka	*-ka	
3SG	*si-ya	*∅	*∅	*∅	
1EXCL	*=kami	*kami	*-kami	*-kami	
1INCL.DU	*=kita	*kita	*-kita	*-ta	
1INCL.PL	*=kitam	*kitam	*-kitam	*-takayú	
2PL	*=kam	*kamu, *kayu	*-kamuyu	*-kayu	
3PL	*si-di	*ida	*-da	*-da	
c. GENITIVE PRONOUNS					
1SG	*=ku	*ku ~ *-k	*-ku ~ *-k, *-ta	*-ku ~ *-k	
2SG	*=mu	*mu ~ *-m	*-mu ~ *-m	*-mu ~ *-m	
3SG	*=na	*na	*-na	*-na	
1EXCL	*=mi	*mi	*-mi	*-mi	
1INCL.DU	*=ta	*ta	*-ta	*-ta	
1INCL.PL	*=tam	*tam	*-tam	*-taku	
2PL	*=muy	*muy, *yu	*-muy	*-yu	
3PL	*=di	*da	*-da	*-da	
d. OBLIQUE PRONOUNS					
1SG	*ni/di-akən	*kani akən	*kanyaken		
2SG	*ni/di-kaw (>*ko)	*kani kaw	*kanikaw		
3SG	*diya, *ni/di-ko-na	*kani kua na	*kanya		
1EXCL	*ni/di-kami	*kani kami	*kanikami		
1INCL.DU	*ni/di-kita	*kani kita	*kanikita		
1INCL.PL	*ni/di-kitam	*kani kitam	*kanikitam		
2PL	*ni/di-kam	*kani kamu, *kani kayu	*kanikamuyu		
3PL	*ni/di-di	*kani kua da	*kanida		

* Reid (1974, 1979a) reconstructs PCCORD *-taku for the 1INCL.PL nominative pronouns, but this is revised to *-takayú with final stress in Reid (2009), and we have updated the tables here to reflect that latter source.

therefore seems likely that, even if there was a set of PNLUZ pronouns marked by the formative *siq-, there was also a simpler set with the formative *si-.

There has also been a reduction of the 2PL.TOP/OBL base from *-kamuyu (or likely a simpler form like *-kamu, as Tharp reconstructs for PNCORD, and which can also be reconstructed for PMP) to *-kam. The use of the 3rd person Topicalized forms *siya and *sidi has been extended to the Nominative set. In Dupaningan Agta, the 3SG short-form nominative is actually null, but the topicalized nominative is used when an overt pronoun is needed for disambiguation (Robinson 2011:81–83). Since the Proto-Northern Luzon 3SG short-form nominative is also reconstructed as null, it is likely that PNLUZ extended use of the topicalized pronouns in much the same way as DUP does. We reconstruct PNLUZ *siya ‘3SG.NOM’, however, because there is no evidence that a null is used synchronically in any of the languages except DUP.

PNELUZ languages do not reflect the shortened monosegmental enclitic forms *=k ‘1SG.GEN’ and *=m ‘2SG.GEN’, which Reid (1979a) reconstructs as PNLUZ allomorphs of *=ku and *=mu, respectively, following vowel-final forms. These allomorphs do sometimes occur in Dupaningan Agta, but appear to be loans from Ilokano, since they rarely if ever occur when speakers are consciously trying to speak “pure” Dupaningan Agta, as opposed to the colloquial way of speaking that includes a great deal of code-switching with Ilokano (cf. Robinson 2011).

There has been a reduction of PNLUZ *=muyu ‘2PL.GEN’ to PNLUZ *=muy. A reduction of the oblique formative *kani- to *ni- has taken place in DUP, PAH, and DIN, which also occurs in some Batanic/Bashiic languages. Note that the presence of the oblique formative *di- in PAR, CAS, and NAG, also found in Batanic/Bashiic, Sabahan, and even some Central Philippine languages, suggests that the rest of the Northern Luzon languages lost *di- as an oblique pronoun formative, just as the vast majority of Central Philippine languages did.

2.2.2 Case markers. Table 8 lists the case markers for the languages of Northeastern Luzon.

The Paranan case markers pose a particular problem. Most of the Paranan forms are similar to forms found in the Agta languages (for example, *i NOM, *ti GEN/OBL), usually Pahanan Agta, but a few of the forms (*en* ‘NOM.DEF’, *nen* ‘GEN.DEF’, and *ten* ‘OBL’) do not appear to have an origin in PNLUZ. Taking into consideration that *en*, *nen*, and *ten* are likely from earlier *in, *nin, and *tin (since Paranan [e] often corresponds to Central Philippine *i in closed syllables), the first two forms (*in and *nin) both have cognates in Central Philippine languages: Old Bikol had both *in ‘NOM.NONREF’ and *nin ‘GEN.NONREF’, while *in also has cognates in Waray-Waray and other Warayan languages, in Tausug, and in the Kamayo dialect of Barobo town. Genitive *nin has cognates in most Bikol languages and in Romblomanon. The *tin form could be from Pahanan Agta *ti*, with the final *-n being the result of analogy with the *in and *nin forms. Note that the *ʔ- : *n- : *t- contrast (where *ʔ- corresponds to the phonemically vowel-initial form) is also found in Southern Ibanag, whose case markers are *iC* ‘NOM’, *nəC* ‘GEN’, and *təC* ‘OBL’ (with the final segment being a copy of the first consonant of the following word). However, the vowels do not match, as Paranan /e/ is not cognate with Ibanag /ə/.

Table 9 compares the reconstructed PNELUZ case markers with those reconstructed for Proto-Northern Cordilleran (Tharp 1974a), Proto-Philippines (Reid 1979b), and Proto-Central Cordilleran (Reid 2006). No reconstruction of the Proto-Northern Luzon case markers is available. The PNELUZ case markers are markedly different from those reconstructed for the other protolanguages. The nominative common case marker *i reflects the protoform rather unproblematically, but of the reconstructed PNELUZ genitive common case markers (indefinite *ti and *ta, and definite *nu and *tu), only *nu is reconstructed elsewhere (for PPH), and none of these is reconstructed for Proto-Northern Cordilleran. On the other hand, in Dupanangan Agta and Casiguran Agta (the two lan-

TABLE 8. CASE MARKERS

		DUP	PAH	DIN	PAR	CAS*	NAG	PROTO-NELUZ
COMMON, SG	NOM (INDEF)	ø	—	i	i	i, tu	i	*i
	(DEF)	i	u, tu	u	en	—	—	(*u)
	GEN (INDEF)	(na/di)	ti	ta	ti, nen	no, na	—	*ta, *ti
	(DEF)	—	nu	nu, tu	nen	—	na, tu	*nu, *tu
COMMON, PL	OBL	ha	ti	ta	ti, ten	to, ta	tu, ta	*ta
	NOM (INDEF)	—	—	—	—	du, di	—	—
	(DEF)	—	—	—	—	—	—	—
PERS, SG	GEN (INDEF)	—	—	—	—	du, di	—	—
	(DEF)	—	—	—	—	—	—	—
	OBL	—	—	—	—	du, di	—	—
PERS, PL	NOM	ni	ti	ti	ti	ti	ti	*ti
	GEN	(na) ni	ni	ni	ni	ni	ni	*ni
	OBL	ha ni	ni	ni	kənni, kən	ni	ni	*[ka]ni
PERS, PL	NOM	di	di	di	di ~ de	de	—	*di
	GEN	(na) di	di	—	di	de	—	*di
	OBL	ha di	kad	kəd	kəndi	de	de	*ka[n]di

* From Headland and Healey (1974), reprinted in Headland and Headland (1974).

TABLE 9. COMPARISON OF RECONSTRUCTED CASE MARKERS

		PNELUZ	PNCORD	PCCORD	PPH
COMMON	NOM INDEF.	*i	*i, *ia, *iu	*=y, *ø	*ʔi, *su, *ʔu
	DEF.	(*u)	—	*nan	—
	GEN INDEF.	*ta, *ti	*na	*=n, *ø	*na, *nu
	DEF.	*nu, *tu	—	—	—
	OBL	*ta (PPAR *ti)	*sa	*=s, *si*	*di, *sa
PERS., SG.	NOM	*ti	*si	*=s, *si	*si
	GEN	*ni	*ni	—	*ni
	OBL	*[ka]ni	*kani	—	*ka ni, *kay (<*ka ʔi)
PERS., PL.	NOM	*di (<*da)	*da	*da	—
	GEN	*di (<*da)	*da	—	—
	OBL	*ka[n]di (<*da)	*kada	—	—

* Reid (2006) also includes topic (*sa), dative (*=n/*=y, *kan/*kay) and locative (*=d, *ʔidi/*ʔudi/*di).

guages for which we have more extensive data), *na*—which is reconstructed for both PPH and PNCORD—is used for at least some of the genitives.²²

The southern NELUZ languages use *ta* for the oblique, while PAR and PAH have *ti*, and DUP retains a reflex *ha* of the earlier **sa*. The form *ta* also occurs in Atta, Central Cagayan Agta, and Ibanag (Tharp 1974a), where the change of **s* > *t* is regular. However, Tharp reconstructs **sa* for PNCORD based on external evidence. Therefore, the form *ta* in the southern NELUZ languages could be a borrowing from one of these languages, or could reflect an inherited **sa* having undergone a shift of **s* > /*t*/ limited to the case markers.

For the singular nominative personal case marker, the NELUZ languages all have **ti*, except DUP, which has *ni*. The latter appears to be an extension of the genitive personal case marker, where all of the languages reflect the widespread **ni*, which is reconstructed for PPH, PMP, and PAN. Although most of the languages seem to reflect *ni* for singular personal oblique as well, we tentatively reconstruct *[*ka*]ni on the basis of the Paranan reflex, since the PPH form was **kani*. It is, therefore, assumed that the **ka*-element was lost in the other NELUZ languages. It is also possible that PAR *kənni* was a borrowing from Ilokano or one of the Cagayan Valley languages. For the personal plural oblique form, we have more evidence for a **ka*-formative, as it is reconstructed for Proto-Northern Cordilleran and would be an expected form for PPH based on the other reconstructions given by Reid (1979b). Finally, the personal plural for both the nominative and genitive is **di*, which is a regular reflex of the reconstructed **da* with Low Vowel Fronting.

2.3 LEXICAL EVIDENCE FOR A NELUZ GROUP. There are a number of lexical innovations that appear to be NELUZ innovations, for which we have not found cognates after searching the lexical resources available for other languages. It is likely, however, that some of these forms will turn up in other languages as the documentation of northern Philippine languages becomes more widespread. In order to minimize this problem, we have excluded any forms with meanings that do not generally appear on Swadesh lists. If we had included the more rare forms, this list of proposed innovations would have been much longer, but it would have increased the likelihood that many of those proposed innovations would have undocumented cognates outside the NELUZ subgroup.

The following forms (71)–(76) are proposed unique PNELUZ lexical innovations that occur in all of the NELUZ speech varieties we surveyed.²³

- (71) PNELUZ *ləbbút ‘boil water’ > DUP labbút, PAH, PAR, KAS ləbbút, CAS, NAG ləbút
- (72) PNELUZ *ladú? ‘fever’ > DUP, PAH, PAR, CAS, KAS, NAG ladú? (also ‘sick’ in PAH, KAS, CAS, NAG)
- (73) PNELUZ *putát ‘full’ > DUP, PAH, PAR, CAS, KAS, NAG putát

22. In Dupanigan, *na* is a pronoun that agrees with singular genitive noun phrases (Robinson 2011:55). The Dupanigan case marking system may be in transition, as Liao (2005) describes for Central Cagayan Agta (a Cagayan Valley language), in which the genitive pronominal forms are in the process of losing their status as clitics and becoming agreement features instead.

23. Since our lexical survey of Dinapigue Agta was far less complete than for the other languages, we lack a DIN cognate for many of these forms.

- (74) PNELUZ **madəggáʔ* ‘heavy’ > DUP, PAH *madaggiʔ*, PAR, NAG *madəggiʔ*, CAS *madəgiʔ*, KAS *madəggá*
- (75) PNELUZ **démət* ‘arrive’ > DUP *démat*, PAH, PAR, CAS, KAS, NAG *démət*
- (76) PNELUZ **pilás* ‘muscle’ > DUP, PAH, CAS, KAS, NAG *pilás*, PAR *pilá*²⁴

Item (77) reflects a unique semantic shift from ‘buttocks’ to ‘vulva’.

- (77) PNELUZ **ubət* ‘vulva’ > DUP, PAR *ubát*, PAH *ubbát*, CAS, KAS, NAG *ubət*

Items (78)–(80) seem to be unique, but there are suspiciously similar-looking forms in other languages in the Philippines.

- (78) PNELUZ **lupás* ‘rice husk’ > DUP *lupás*, PAH, PAR, CAS, KAS, NAG *lupás*; cf. Ilianen Manobo *qupis* (Reid 1971); Western Subanon *pis*, Molbog *upis* (Yap 1977)²⁵
- (79) PNELUZ **[k]e[n]nam* ‘taste, try’ > DUP *énnam*, PAH *ínnam*, PAR *ennəm*, DIN *ennám*, CAS, KAS *kenam*, NAG *kennám*; cf. Central Cagayan Agta *naa-naamam* (Oates and Oates 1955)
- (80) PNELUZ **masanikíʔ* ‘shy, ashamed’ > DUP, PAH, NAG *masanikíʔ*, PAR *mansəŋkiʔ*, CAS, KAS *masanikéʔ*; cf. Ivatan *masnik* (Reid 1971, Yap 1977)

The following five items (81)–(85) are innovations that we tentatively reconstruct for PNELUZ, but which are not found in all of the languages. Based on our subgrouping argument presented in section 3, we only list words that are found in DUP, at least one of the other northern languages (PAH, PAR, DIN), and at least one of the southern languages (KAS, CAS, NAG).

- (81) PNELUZ **sánig* ‘hear, listen’ > DUP, PAH, PAR, DIN, KAS *sánig* (**na-sánig* ‘hear’, **mag-sánig* ‘listen’)
- (82) PNELUZ **tóglad* ‘push (to transport)’ > DUP, CAS, KAS *tóglad*, PAR *túglad*
- (83) PNELUZ **bakál* ‘stab’ > DUP *bakál*, PAH, DIN, CAS²⁶ *bikál*
- (84) PNELUZ **rəktat* > DUP *huméktat* ~ *ruméktat* ~ *ruméttat* ‘start a journey, commence’, PAH, CAS *huméktat* ‘leave’²⁷
- (85) PNELUZ **ləddís* ‘crush lice’ > DUP *laddís*, PAH *ləddís*, CAS, NAG *lədís*

3. INTERNAL SUBGROUPING OF THE NORTHEASTERN LUZON LANGUAGES. This section presents the evidence for the internal subgrouping of the Northeastern Luzon languages from phonology (3.1), morphology (3.2), and lexicon (3.3).

3.1 PHONOLOGICAL EVIDENCE

3.1.1 Metathesis. Casiguran Agta and Nagtipunan Agta share an unusual metathesis involving the secondary patient voice verbs, in which the first consonant and vowel of the

24. Possibly PAR **pilás* > **piláh* > *pilá*.

25. Yap (1977) also lists the initial /l/ in CAS as optional, making the cognates stronger, but there is no evidence for a morpheme boundary after the /l/ in CAS.

26. This CAS form is from Headland and Headland (1974). Our field notes contain *disón*.

27. The CAS gloss ‘leave, depart from a place’ (Headland and Headland 1974) suggests that the definitions for all three languages are probably more similar than is suggested by ‘leave’.

root metathesize if the first vowel of the root is a schwa (that is, $i- + Cə- > iɛC-$), as illustrated in forms (86)–(88). Headland and Healey (1974) note that this is a synchronic process in Casiguran Agta, and that it only occurs when the first syllable of the root is open. Since we generally elicited only one form for each verb in our own data, we can only assume that the Nagtipunan Agta process is also synchronic.

- (86) PNELUZ *lə[b]bəŋ ‘to bury’ > CAS, NAG iyəlbəŋ ‘bury (PV2.INF)’ (cf. CAS root *labəŋ*)
- (87) PNELUZ *dətton ‘to put, place’ > CAS iyəgton, NAG iyədtón ‘put, place (PV2.INF)’ (cf. CAS root *dəton*)
- (88) PNELUZ *bəttən ‘hang by rope’ > CAS niyəbətən ‘hang by rope (PV2.PAST)’

3.1.2 Monophthongization. Casiguran Agta and Nagtipunan Agta also share a monophthongization rule, in which *aw and *ay became /o/ and /e/, respectively, as illustrated in examples (89)–(94). There are no diphthongs in penultimate syllables in these languages, so all examples occur in the ultima. Note that a glottal stop is added phonetically after what would otherwise be a word-final vowel, a development shared by all of the NELUZ languages.

- (89) *anay ‘termite’ > CAS, NAG anéʔ (cf. PAH, PAR, DIN, KAS anáy)
- (90) *balay ‘public building’ ‘house’ > CAS, NAG biléʔ (cf. DUP, PAH, PAR, DIN biláy, KAS baláy)
- (91) *piray ‘crippled’ > CAS, NAG piléʔ (cf. DUP, PAH, PAR, DIN, KAS piláy)
- (92) *takaw ‘steal’ > CAS, NAG takóʔ (cf. DUP, PAH, PAR, DIN, KAS takáw)
- (93) *lanjaw ‘housefly’ > CAS, NAG lanjóʔ (cf. DUP, PAH, PAR, DIN, KAS lanjáw)
- (94) PPH *buŋaw ‘testicles’ > CAS, NAG buŋóʔ (cf. DUP, PAH, PAR, KAS buŋáw)

Headland and Healey (1974) describe this monophthongization rule for CAS as involving lax mid vowels, that is, *aw and *ay shift to /e/ and /o/, respectively. Because of the short amount of time that we worked on CAS and NAG, we can only assume that forms (89)–(94) should have lax vowels rather than tense ones.

3.1.3 *s > /h/. PNELUZ *s is reflected sporadically as /h/ in all of the NELUZ languages except CAS, which would seem to suggest that CAS split off from the other languages before this innovation occurred. However, as we will see below, there is stronger evidence that DUP is a first-order subgroup of NELUZ, and very strong evidence that CAS should subgroup with NAG. Therefore, it is assumed that the sporadic *s > /h/ shift spread according to the wave model after the breakup of PNELUZ.

3.2 MORPHOLOGICAL EVIDENCE. All of the NELUZ languages except DUP innovated a shift of *aw > /o/ in the 2SG topicalized and oblique pronouns. Note that this is an expected shift in CAS and NAG, but not in DIN, PAH, or PAR, which also reflect it. Since this shift is limited to a single pronominal base, it is unclear whether this was a separate innovation in DIN, PAH, and PAR, or if this occurred under the influence of CAS and/or NAG.

Pahanan Agta and Paranan are unique in having *ti* as a genitive and oblique case marker for common nouns, whereas the other NELUZ languages (except DUP) use *ti* as a

singular nominative case marker for personal names. It is difficult to determine whether this is an innovation, however, since *ti* is also the common genitive case marker in Ilokano, and the Ilokano common oblique case marker *iti* is also often shortened to *ti*. Therefore, this could be a borrowing from Ilokano or, more likely, a functional shift of the singular personal nominative case marker to instead mark the genitive and oblique of common nouns.

The fact that all of the languages except DUP use *ti* as the singular personal nominative marker is also evidence for separating Dupanangan from the other Northeastern Luzon languages.

DUP, PAH, PAR, and DIN share the innovated 3SG.OBL pronoun *ni/di-ku-na, while Casiguran Agta and Nagtipunan Agta both reflect *diya, a form that is widespread in the Philippines. It is likely that this innovation spread according to the wave model after the breakup of NELUZ.

3.3 LEXICAL EVIDENCE FOR INTERNAL SUBGROUPING. This section presents lexical innovations shared by two or more of the NELUZ languages. As our goal here is to determine the internal subgrouping of NELUZ, we restrict ourselves to presenting clusters of languages (although not necessarily genetic subgroups) that share at least two lexical innovations. All reconstructions in this section (that is, [95]–[121]) are for unspecified nodes below Proto-Northeastern Luzon. Lacking further evidence, we assume only that they belong to a node that includes all of the languages in which the forms are reflected.

Six innovations (95)–(100) were found in all of the NELUZ languages except DUP.

- (95) *apérit ‘short (length)’ > PAH, CAS apéhit, PAR, KAS apérit, NAG apíhit
- (96) *hágad ‘chase’ > PAH, PAR, CAS, KAS, NAG hágad
- (97) *táblal ‘tasteless’ > PAH, PAR, CAS, KAS, NAG táblal
- (98) *talád ‘sugarcane’ > PAH, PAR, CAS, KAS, NAG talád
- (99) *burák ‘blind’ > PAH, CAS buhák, PAR, KAS burák ‘blind’; KAS naburák, NAG nabuhák ‘having dirt in the eye’
- (100) *tukóy ‘know (facts)’ > PAH, PAR, CAS, KAS, NAG tukóy (also CAS, NAG ‘know how’; KAS, CAS, PAR, DIN (na)tukuy(an) ‘learn’; PAH, CAS, KAS, NAG ‘remember’)

Nine lexical innovations (101)–(109) were found that are unique to CAS and NAG.

- (101) *dulóy ‘lungs’ > CAS, NAG dulóy
- (102) *bidút ‘deer’ > CAS, NAG bidút
- (103) *mediñát ‘red’ > CAS, NAG mediñát (cf. DUP digkat)
- (104) *kəbíl ‘carry, bring, hold in hand’ > CAS, NAG kəbíl
- (105) *ləbúg ‘fight’ > CAS, NAG ləbúg (cf. Yogad labu, Tiruray lifut) (Davis and Mesa 2000; Yap 1977)
- (106) *lisó? ‘hide’ > CAS, NAG lisó? (cf. Isneg siru, Yap 1977)
- (107) *mag-ayág ‘play’ > CAS mag-iyag (with LVF from the prefix), NAG mag-ayág (cf. DUP kayag, Waray uyag) (Yap 1977)

- (108) *sahát ‘happy’ > CAS, NAG saháť
 (109) *dinsón ‘stab’ > CAS, NAG dinsón

There were four innovations (110)–(113) unique to PAH and PAR, although in this case, one language could have easily borrowed the term from the other.

- (110) *dukót ‘cook rice’ > PAH, PAR dukót (cf. DUP ‘build a fire’; semantic shift from PPH[?] ‘burn’)
 (111) *masibót ‘tight’ > PAH, PAR masibót (cf. Yogad siggat ~ sigat ‘tighten’) (Davis and Mesa 2000)
 (112) *mad(iə)ŋət ‘dirty’ > PAH madəŋət, PAR madiŋət
 (113) *habág ‘have mercy’ > PAH nahabág, PAR kahabágan

There were three innovations (114)–(116) unique to PAH, PAR, CAS, and KAS (but surprisingly absent from NAG).

- (114) *bulón ‘heel’ > PAH, PAR, CAS, KAS bulón (cf. Manide, Inagta Alabat bu-lúng ‘knee’)
 (115) *bulibuli ‘lie, untruth’ > PAH, PAR, CAS bulibuli?, KAS bulibulé?
 (116) *sapsap ‘nipa tree (living, as opposed to processed)’ > PAH, PAR, CAS, KAS sapsap

There were three innovations (117)–(119) unique to CAS and KAS.

- (117) *pəknit ‘rip, tear’ > CAS pəknit, KAS pəknet
 (118) *mapérəŋ ‘noisy’ > CAS məpəhəŋ, KAS məpérəŋ
 (119) *ma-kelagíp ‘ask’ > CAS məg-pa-kelágip, KAS na-kélagip (cf. Ilokano lagip ‘remember’)

There were two innovations (120) and (121) unique to DIN, CAS, and NAG.

- (120) *mudít ‘face’ > DIN, CAS, NAG mudít
 (121) *mag-eplək ‘thirsty’ > DIN məg-íplək, CAS məg-eplək, NAG mag-eplək

3.4 SUMMARY OF EVIDENCE FOR INTERNAL SUBGROUPING.

All of the NELUZ languages except Dupaningan Agta share the singular personal case marker *ti* and the monophthongization of the diphthong in the second person singular topicalized and oblique pronouns. This contributes to the evidence for placing Dupaningan Agta as a primary branch of NELUZ.

Casiguran Agta and Nagtipunan Agta share a unique metathesis in the secondary patient voice, and the monophthongization of diphthongs. We calculated 77 percent similarity between the two varieties on a 200-word Swadesh list (cf. table 1), very close to the threshold for considering them to be dialects of a single language. Moreover, it is unlikely that the similarities between Casiguran Agta and Nagtipunan Agta are due to recent borrowing or influence from common trade languages, for although the languages are separated by relatively few miles, the mountains separating the two are formidable, and both groups claimed not to travel across them. Additionally, while Casiguran Agta is most influenced by Tagalog and Kasiguranin, Nagtipunan Agta is surrounded by Ilokano, which likely accounts for many of its lexical differences with the other more coastal

NELUZ languages. We, therefore, conclude that Casiguran Agta and Nagtipunan Agta are very closely related, if not dialects of a single language.

Pahanan and Paranan share the innovation of using the case marker *ti* for the genitive and oblique of common noun phrases. The first author calculated the lexicon of Pahanan Agta and Paranan to be 81 percent cognate (cf. table 1).

Evidence from lexical innovations as described in 3.3 also suggests that Dupaningan forms a primary branch of the Northeastern Luzon subgroup vis-à-vis the other five languages. Of the remaining five languages, Casiguran Agta and Nagtipunan Agta form one branch together, and Pahanan Agta and Paranan form another branch together, while the evidence for the exact position of Dinapigue Agta is lacking, due to the smaller amount of data available.

Figure 1 illustrates the proposed internal subgrouping of the Northeastern Luzon languages. Based on the data presented here, the position of Dinapigue Agta remains unresolved, as it shares some features with Pahanan Agta and Paranan, and others with Casiguran Agta and Nagtipunan Agta. Given the limited amount of data on this variety, it is not surprising that it is difficult to resolve its place within the tree. Kasiguranin is not included in this tree, even though it shares several lexical innovations with Casiguran Agta, which are most likely early borrowings, as is much of the non-CPH lexicon of Kasiguranin. However, Kasiguranin's Tagalog-heavy substrata clearly point to its genetic relationship with Tagalog (Lobel and Robinson 2012), and the similarities to Casiguran Agta can be explained as the result of borrowing after it separated from the core Tagalog dialects.

4. EXTERNAL RELATIONSHIPS OF THE NORTHEASTERN LUZON LANGUAGES. Having discussed the unity of the NELUZ subgroup and its internal structure, we will now address the question of how this subgroup is related to other languages in the northern Philippines, considering the evidence for or against each of the various possibilities.

To date, it has been argued that the NELUZ languages subgroup with the Cagayan Valley languages, as two coordinate branches of a Northern Cordilleran subgroup, which is in turn one of the primary branches of the Northern Luzon group. Figure 2 illustrates the internal structure of the Northern Luzon subgroup according to Reid (2010). In order to support this hypothesis, it would be necessary to find exclusive innovations shared by both Proto-Cagayan Valley and Proto-NELUZ. However, we find no innovations shared exclusively by these two subgroups, as will be discussed below.

FIGURE 1. PROPOSED NELUZ INTERNAL SUBGROUPING

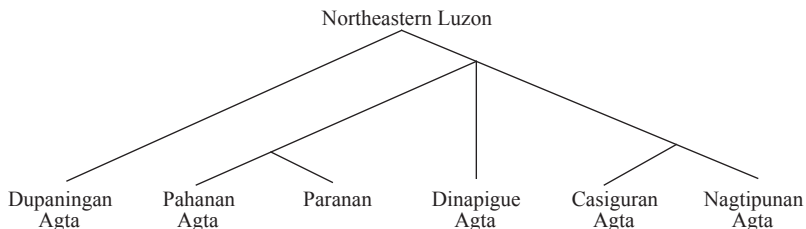
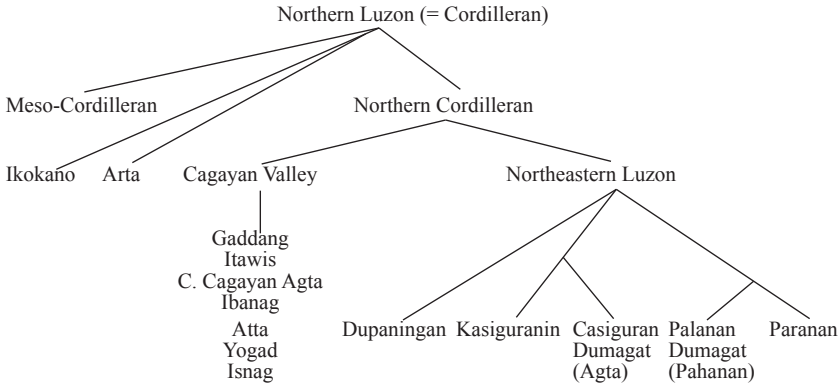


FIGURE 2. THE NORTHERN LUZON LANGUAGES (REID 2010)



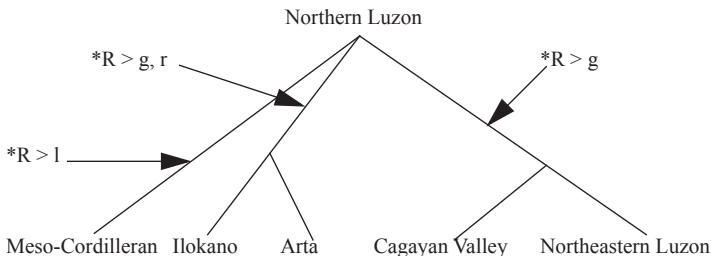
If an exclusive link with the Cagayan Valley languages cannot be demonstrated, then the next possibility is that NELUZ still belongs in the Northern Luzon subgroup, but is not a part of NCORD, in which case NELUZ would be coordinate with NCORD, C/SCORD, Ilokano, and Arta. To prove this, we would need evidence that the NELUZ languages share exclusive innovations with the other Northern Luzon languages. The evidence for this possibility is stronger than the evidence for a link with the Cagayan Valley languages.

A third possibility is that the NELUZ languages are coordinate with Northern Luzon, in a higher subgroup. Any similarities shared exclusively with Northern Luzon would be due to a heavy Northern Luzon overlay from early contact. Since very few innovations for Northern Luzon have been identified, this latter hypothesis awaits further study.

4.1 PHONOLOGICAL EVIDENCE. This section will discuss whether the phonological innovations shared by the NELUZ languages (cf. section 2) can be used to subgroup the NELUZ languages vis-à-vis the other languages of the northern Philippines.

4.1.1 *R > /g/. As in the Cagayan Valley languages, the NELUZ languages all reflect the shift of *R > g. The reflexes of *R have generally been considered quite important in subgrouping the languages of the Philippines, and the current hypotheses for subgrouping the languages of the northern Philippines are no exception. Figure 3 presents a tree of the Northern Luzon languages taking into consideration only the *R > g shift.

FIGURE 3. INTERNAL STRUCTURE OF NORTHERN LUZON BASED ON *R



Note that this tree is nearly identical to the one proposed by Reid (2010), except that it subgroups Ilokano and Arta together, which Reid (1989) leaves open as a possibility. NELUZ and Cagayan Valley are the only subgroups within Northern Luzon to reflect an unconditioned $*R > g$ shift. In Ilokano and Arta, $*R$ split to $/g/$ and $/r/$ without clear conditioning (Tharp 1974b, Reid 1989).

4.1.2 The reflexes of $*j$ and $*d$. In the NELUZ languages, $*j$ merged with $*d$ as $/d/$. Within Northern Luzon, only NELUZ, Arta, Ilokano, and Northern and Southern Alta reflect the merger of $*j$ with $*d$, where the other Northern Luzon languages have merged $*j$ with $*g$ instead (Reid 2006:5–6). Problematically, this suggests a different subgrouping than the one suggested by the reflexes of $*R$, as illustrated in figure 3. If we were to base the subgrouping of the NELUZ languages primarily on the reflexes of $*j$, we would get the tree presented in figure 4.

The reflexes of $*R$ are also indicated for the sake of comparison. Note that if this is the correct subgrouping, then the mergers of $*R$ and $*g$ in NELUZ and the CV languages were independent of one another, as was the shift of $*R > l$ in Northern and Southern Alta, on the one hand, and South-Central Cordilleran, on the other. Certainly, it would be ideal to minimize the instances of convergence in any subgrouping hypothesis, so this is not an ideal tree. On the other hand, if we revisit the reflexes of $*R$ in figure 3, this time including the $*j$ reflexes, we also run into problems, as can be seen in figure 5 below.

To reconcile the reflexes of $*j$ with the tree based on $*R$, more convergent changes are required. The shift of $*j > d$ must have occurred independently three separate times, while the $*j > g$ shift must have occurred twice. In terms of parsimony, then, the tree based on the reflexes of $*j$ is preferable. But in order to support the $*j$ tree, we would need exclusively shared innovations between NELUZ, Ilokano, and Arta on the one hand, and between CV

FIGURE 4. INTERNAL STRUCTURE OF NORTHERN LUZON BASED ON $*j$

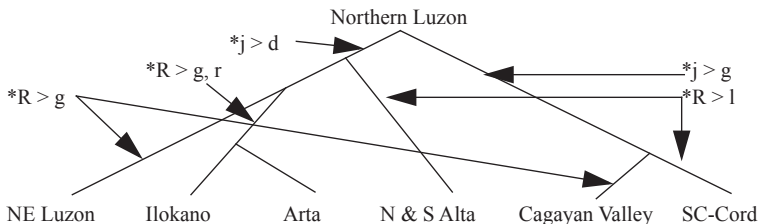
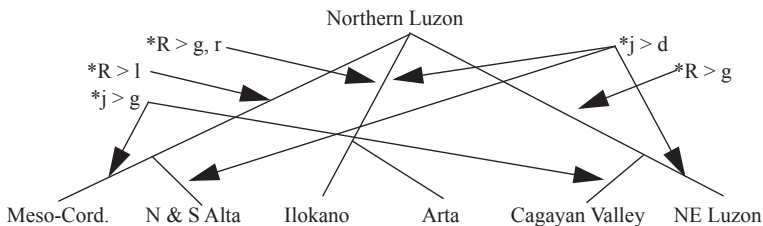


FIGURE 5. SUBGROUPING OF NORTHERN LUZON WITH $*R$ AND $*j$



and Meso-Cordilleran on the other. We have not specifically sought out such evidence because we believe that there is a simpler explanation for the reflexes of *j and *R.

It is important to note that there are very few modern reflexes of *j and *R in the Philippines. In all cases, both *j and *R have merged with some other phoneme in all known Philippine languages, “making none of them what Dempwolff called ‘Test-Sprachen’ for *j or *R” (Bob Blust, pers. comm., February 13, 2013). Furthermore, none of the modern reflexes of *j or *R continue the reconstructed phonetic values attributed to PMP: *j as a “palatalized velar stop” (Blust 1991a:134) or a velar fricative (Ross 1992), and *R as a uvular trill or fricative (see Blust 1991a and Wolff 2003, among others). Indeed, throughout the Philippines, there is a surprising lack of diversity in the reflexes of both *R and *j, likely due to the contact-based influence of a handful of powerful languages in relatively recent times, something that has been alluded to in various contact and leveling hypotheses presented by Reid (1987, 1994) and Blust (1991b, 1999).

It should also be noted that, although the NELUZ languages are the only Northern Luzon languages to reflect both *R > g and *j > d, this same pair of shifts is found in the Greater Central Philippines subgroup. We suggest, then, that the phonological shifts of *R > /g/ and perhaps also *j > /d/ in PNELUZ happened under influence of speakers of more prestigious Greater Central Philippine languages who were sailing along the east coast of northern Luzon, where the NELUZ languages are spoken. The various northern Philippine languages may have retained a distinct *R phoneme until relatively recently, but then came under the influence of languages whose populations were much more powerful or at least much more advanced in terms of trade and regional connections, so their pronunciation of *R and *j assimilated to that of the prestige language. If they were under the influence of a GCPH language such as Tagalog or Bikol, this would have been *R > /g/.

This possibility is consistent with the idea of the “stereotyped Philippine /g/ reflex” of *R (Conant 1910), which was revisited by Blust (1991b) who proposed that the /g/ reflex of *R in Ilokano, and the presence of *mag-* < *maR- in northern Philippine languages that do not otherwise reflect *R as /g/, may have happened under the influence of a prestigious Greater Central Philippine language.

This may not be such a stretch after all, since recent historical research suggests that the most important centers of long-distance trade in the Philippines were Ma-i (Mindoro) in the late tenth century AD (Scott 1984), Butuan in northeastern Mindanao at the opening of the eleventh century AD (Scott 1984, Hontiveros 2000), and then Jolo in the Southwest Philippines beginning in the 1300s, all areas where Greater Central Philippine languages are spoken (Butuanon in Butuan, Tausug in Jolo, and either Tagalog or a Bisayan or Mangyan language in Mindoro).

It is also possible that *R > /g/ and *j > /d/ developed in the NELUZ languages under the influence of both Ilokano to the north and Greater Central Philippine languages to the south. If speakers of NELUZ languages were surrounded by two very different languages that both happened to share these reflexes, it is not difficult to believe that they would have adjusted their pronunciation of *R and *j to sound similar to the two surrounding, more prestigious, groups.

In short, then, while the reflexes of *R and *j are particularly salient and have been frequently used for subgrouping, they constitute fairly weak evidence for classifying the NELUZ languages.

4.1.3 *t/s* metathesis. The main feature defining Proto-Northern Luzon is *t/s* metathesis, which is also found in NELUZ. The sequence *tVs metathesized to *sVt in the Northern Luzon languages. There is also the long-distance metathesis of *tVcVs to *sVCVt. These processes were first noted by Dyen (1972), and both are reconstructed by Reid (2006) for Proto-Northern Luzon. In our data, the metathesis appears to be sporadic for all of the languages in which it occurs. The Cagayan Valley languages (except Malaweg) reflect the *s > /t/ shift, so the metathesis, assuming it occurred in the ancestor of the modern Cagayan Valley languages, is obscured. Based on our data, Umiray Dumaget does not reflect this metathesis. Forms (122)–(128) illustrate the *t/s* metathesis.

- (122) *Ratas ‘milk’ > DUP, PAH, CAS, NAG gisát ‘breast’²⁸
- (123) *tanjis ‘cry’ > DUP, PAR sanjét PAH, CAS, NAG sanjít (KAS sanjét)
- (124) *tədis ‘crush lice’ > DUP saddit, PAR səddít
- (125) *bitiəs ‘calf of leg’ > DUP bisat
- (126) *di-taqs > DUP disat ‘high ground’, CAS disat ‘term for the upper part of a hillside’ (Headland and Headland 1974)
- (127) *tiRis ‘decant’²⁹ > DUP sigít ‘dip for liquid with a small bucket’, CAS sigít ‘to pour water into a cup (from another container, or from a faucet)’ (Headland and Headland 1974)
- (128) *təRas ‘hardwood, hard’ > DUP sagget ‘narra (a kind of valuable hardwood)’

In form (129), the NELUZ reflexes do not reflect metathesis.

- (129) *taqəmis > PAH, PAR, CAS, NAG támʔis (KAS tamís) ‘sweet’ (DUP sámʔit is clearly an Ilokano loan, since Dupanangan Agta does not retain medial glottal stop in the native stratum)

Moreover, the process is no longer productive, as there are words with the sequences *tVs* and *tVcVs*, such as (130).

- (130) PNELUZ *tulós ‘continue’ > DUP, PAH, CAS tulós, NAG tulús (KAS tulós)

4.1.4 *r > /l/. As mentioned in 2.1.1.3, the NELUZ languages reflect the split of *r > /r/ and /l/ without any apparent conditioning. Tharp (1974a) notes that the Cagayan Valley languages and Ilokano also reflect sporadic *r > /l/. Note that the *r > /l/ shift is fairly common in the Philippines, and is also reflected in Tagalog, Cebuano, Ilonggo, and many other Central Philippine and Greater Central Philippine languages.

4.1.5 Schwa and gemination. As mentioned earlier, all of the NELUZ languages except Casiguran Agta geminate consonants after a schwa. This innovation is also found

28. The semantic shift here is not unique to the NELUZ languages. It also occurs in the CV languages, NALTA, and Umiray. It is unclear if the semantic shift is more widespread than that.

29. This and the following reconstructions are from Dyen (1972), who cites them as PAN, but since he does not use any Formosan evidence, we have changed this to PMP. Dyen does not reconstruct meanings for his protoforms, but rather lists the meanings in the various daughter languages.

in the Cagayan Valley languages and Ilokano (Tharp 1974a). This is weak evidence, however, as many other Philippine and non-Philippine Austronesian languages also geminate consonants after schwa (for example, some Manobo and Sama-Bajaw languages, phonetically in Maranao, and evidence thereof in North Borneo, cf. Blust 2010), which is hypothesized to have been phonetically shorter than the other three reconstructed PAN vowels (Blust 2009:548).

4.1.6 Low Vowel Fronting (LVF). All of the NELUZ languages have Low Vowel Fronting (discussed in 2.1.2.2 above), as do Southern Alta (Reid 1991, also citing a similar process in Northern Alta), Umiray Dumaget (Himes 2002), Manide (Lobel 2010), and Inagta Alabat (Lobel 2011), all of which are spoken by Negrito Filipino populations along the eastern coast of Luzon. Since there is no other evidence that these languages form a linguistic subgroup, we consider this an areal feature of Negrito Filipino languages that has spread from Casiguran Agta and Pahanan Agta to the non-Agta languages Kasiguranin and Paranan, respectively, through borrowing of individual lexical items. It is also worth noting that there are key differences in the implementation of LVF in the Northeastern Luzon languages, in Manide and Inagta Alabat, and in Umiray Dumaget (Lobel 2012). In the NELUZ languages, LVF occurs only sporadically, and only after voiced stops (/b d g/). In Manide and Inagta Alabat, LVF is similarly sporadic, but occurs after glides /w y/ as well as after voiced stops /b d g/. Umiray Dumaget is unique among these languages because its LVF process (which occurs after voiced stops /b d g/ and glides /w y/) is completely productive, including synchronic alternations resulting from prefixation, infixation, and suffixation.

4.1.7 /r/ ~ /h/. All of the Agta languages of northeastern Luzon have an /h/ variant of PNELUZ *r, as illustrated in forms (131)–(134). This is not found in the non-Agta languages Paranan and Kasiguranin. In Dupaningan Agta, for which the first author has much more comprehensive data, we found that /r/ and /h/ are in dialectal variation, with /r/ most commonly found in the northern dialects, and /h/ increasingly prevalent in southern dialects. Frequent movement of individuals, however, makes it difficult to define geographically based dialects for this semi-nomadic group. The Dupaningan Agta data here generally reflect the northern dialect documented in Robinson (2011) and thus preserve /r/.

(131) PNELUZ *barák ‘piglet’³⁰ > DUP barák, PAH bahák, PAR barák, CAS, NAG bahák (KAS barák) (cf. CCAGTA bahák)

(132) PNELUZ *uráy ‘wait’ > DUP, PAR uráy, PAH, CAS, NAG uháy (KAS uráy)

(133) PNELUZ *ikár³¹ ‘cough’ > DUP ikár, PAH, DIN, CAS, NAG ikáh, PAR ikár (KAS ikár) (cf. CCAGTA ikár)

(134) PNELUZ *karamáy ‘centipede’ > DUP, PAR karamáy, PAH, DIN kahamáy, CAS, NAG kahamé? (KAS karamáy)

Interestingly, this feature is also found in Central Cagayan Agta, which belongs to the Cagayan Valley subgroup. Oates and Oates (1958) note that although /h/ is more frequent stem-initially, and /r/ more frequent stem-finally, /r/ alternates freely with /h/ in stem-

30. This is from PAN *bərək and the expected PNELUZ reflex is thus *bəgək.

31. This is from PMP *ikəj and the expected PNELUZ reflex is thus *ikəd.

medial position in about half of the words in which the two phonemes occur, such as in items (135)–(138).

(135) CAGTA *sihat* ~ *sirat* ‘stomach’

(136) CAGTA *fuhaw* ~ *furaw* ‘white’

(137) CAGTA *gahut* ~ *garut* ‘roofing grass’

(138) CAGTA *pahin̄jil* ~ *parin̄jil* ‘cheek’ (Oates and Oates 1958)

In other cases, CAGTA /h/ is clearly derived from earlier *r, but the /r/ reflex does not occur synchronically, such as in form (139).

(139) CAGTA *huplaano* < Spanish *eroplano* ‘airplane’

The alternation of /r/ and /h/, and the replacement of *r with /h/, are areal features of the Agta languages of Cagayan, Isabela, and northern Aurora provinces, rather than an inherited feature in a particular innovation-defined subgroup.

4.1.8 Glottal stop. All of the Agta languages of northern Luzon underwent a process of final glottal stop epenthesis (Vanoverbergh 1937:15, Tharp 1974a:69), but while most historically vowel-final forms contain a word-final glottal stop, there are also forms today that do not. Due to its distribution, this is considered an areal feature in the Agta languages of northern Luzon.³²

The loss of word-final glottal stop is one innovation that has been considered defining of Proto-Northern Luzon (Reid 2006), but because of the aforementioned glottal stop epenthesis, any evidence of loss of word-final glottal stop has been obscured in the NELUZ languages (that is, it is unclear whether the inherited glottal stop was dropped prior to the generalization of the word-final glottal stop on all historically vowel-final roots).

4.1.9 *s > /h/. As discussed in 2.1.1.7 above, *s sporadically becomes /h/ in the NELUZ languages (though not in Casiguran Agta). This shift does not occur in the Cagayan Valley languages (Tharp 1974a), nor in Arta (Reid 1989), and is only found in one language in each of the Central Cordilleran and Southern Cordilleran subgroups: Balangaw in the former (Reid 1974) and Kalanguya in the latter (Himes 1998). Only a single form in Ilokano reflects an *s > /h/ shift: the negator *saʔan*, which has the alternate pronunciation *haʔan*.³³ Otherwise, Ilokano lacks an /h/ phoneme in the native stratum (although speakers now freely borrow Tagalog and English words with /h/).

4.1.10 Monophthongization of *aw and *ay. All of the Cagayan Valley languages share a vowel reduction rule in which *aw and *ay monophthongize to /o/ and /e/, respectively, when followed by another consonant (Tharp 1974a). Casiguran Agta and

32. Note that Manide and Inagta Alabat also reflect sporadic glottal stop epenthesis, but that word-final glottal stop remains phonemic in these languages, as it is unpredictable.

33. Note that independent *s > /h/ shifts are found in functors sporadically throughout the Philippines, including Manide, Inagta Alabat, most Waray-Waray dialects, Butuanon, Tausug, and even in a number of Sabah’s Dusunic languages, including Kujau, Dusun Kuala Monsok, Dusun Tindal, Dusun Tambunan, Coastal Kadazan, and Sonsogon. Bob Blust (pers. comm., February 13, 2013) also points out that “*S > *h > zero also happened between PAN and PMP in high-frequency morphemes, such as *Səpat > *əpat ‘four’, *Si- > *i- ‘instrumental/benefactive voice’, or *Sika- > *ika- ‘ordinal numeral prefix’, as we would normally expect all of the PMP forms here to begin with *h. This appears to be a frequency effect in historical change.”

Nagtipunan Agta share a very similar shift, in which *aw and *ay monophthongize to /ɔ/ and /ɛ/, respectively, as discussed in 3.1.2. We consider the shifts in each of these two subgroups to be independent of one another, as this is an exceedingly common innovation.

4.1.11 Merger of *ə with *a. All of the Cagayan Valley languages merged *ə and *a as /a/, but this change did not occur in Proto-Northeastern Luzon, as the only NELUZ language that consistently merges *ə and *a is Dupaningan Agta, or certain dialects thereof. In other Dupaningan Agta dialects, however, /a/ and /ə/ remain separate phonemes (Robinson 2011:4). The shift of *ə > /a/ in the northern dialects of Dupaningan Agta likely occurred under the influence of Ibanag or some other CV language, prior to the Ilokano expansion into the area.

4.1.12 *s > /t/ and *ti > /s/. Tharp (1974a) lists a set of ordered innovations (1) *s > t, (2) *ti > s / __ {V, y} that have taken place in the Cagayan Valley languages (cf. table 10) and also in Ilongot (Himes 1998:138).³⁴ The NELUZ languages do not share either change.

4.1.13 Other Cagayan Valley innovations. There are a number of other innovations in the Cagayan Valley languages (Tharp 1974a) that are not found in any of the NELUZ languages: (1) the *p > f / _u shift found in Yogad, Ibanag, Gaddang, Central Cagayan Agta, and Itawis (note that there is no [f] phone in any of the NELUZ languages); (2) the *b > h / _u shift in Central Cagayan Agta and Itawis; (3) the intervocalic rhoticization of *d (*d > r / V _V), which is found in Atta, Gaddang, Ibanag, Itawis, and Yogad (although note that the instances of *j > r in the NELUZ languages could be via an intermediate step *j > *d > r, cf. 2.1.1.2); (4) the palatalization of *d > /j/, /z/, or /h/ before /i/, which is found in

TABLE 10. *s > t AND *ti > s IN THE NORTHEASTERN LUZON AND CAGAYAN VALLEY LANGUAGES*

	PMP *asuk 'smoke'	PMP *pusuq 'heart'	PMP *diRus 'bathe'	PMP *uRsa 'deer'	PMP *tian 'stomach'
Northeastern Luzon					
Dupaningan	asók	pusú	dégus	ógsa	tiyan
Pahanan	asók	pusó?	dígus	úgsa	tiyan
Paranan	asók	pusú	dígus	úgsa	tiyan
Casiguran	asók	pusó?	dígus	ógsa	tiyán
Kasiguranin	asók	pusó?	dígus	ogsa	tiyán
Cagayan Valley					
CC Agta	atú?	futu	zigut	úgta?	—
Ibanag	atú?	futú	zigú?	uttá	sa:ŋ
Atta	atu	puttu	jigu?	—	sa:n
Isneg	atú?	(púso)	dixut	ugtá	—
Itawit	atúk	futú?	zihut	uttá	—
Malaweg	asú?	púsu	digus	úgsa	—

* Data are from our field notes, except Isneg (from Vanoverbergh 1972), the Central Cagayan Agta and Atta reflexes of 'heart' (from Tharp 1974), and the Central Cagayan Agta, Atta, and Itawit reflexes of 'bathe' (also from Tharp 1974).

34. According to the first author's field notes, these innovations are also curiously absent from Malaweg, which is supposedly a member of the Cagayan Valley subgroup, and which *Ethnologue* considers a dialect of Itawit. Note, however, that the most recent edition (Lewis 2009) includes the comment "Malaweg may possibly be reclassified as a separate language."

CCAGTA, Atta, Ibanag, and Itawis; (5) the assimilation of the final /g/ of the prefix *maR- (*mag- > maC₁/ _C₁), which is found in Atta, Gaddang, Ibanag, Itawis, and Yogad; (6) the monophthongization of *uy in final position (*uy > i / _#), which is found in Atta, Ibanag, and Itawis; and (7) the lenition of *g > Ø in Gaddang, and of *g > /h/ in Itawis.

4.1.14 Summary of phonological evidence. Table 11 summarizes the phonological evidence discussed in this section. Of the twenty aforementioned phonological innovations, four are shared by the NELUZ languages and the Cagayan Valley (CV) languages, but all four are also found in other languages of the Philippines, which suggests that there is very little to link the NELUZ languages exclusively with the Cagayan Valley languages. The first of these four shared innovations is *R > g, which is quite common and found elsewhere in the Philippines, including (sporadically) in Ilokano and Arta. The second shared innovation is *t/s* metathesis, which has been reconstructed for Proto-Northern Luzon (Reid 2006). The third shared innovation is sporadic *r > l, which is also found in many CPH and even GCPH languages. The fourth shared innovation is gemination after schwa, which is also found in Ilokano and a number of other languages in the Philippines including some Manobo languages and Sama-Bajaw, as well as phonetically in Maranao (cf. Lobel and Riwarung 2011), and historically in the North Borneo languages (cf. Blust 2010). Based on phonological evidence, then, it would seem that the NELUZ languages likely belong to the Northern Luzon subgroup, but probably should not be placed in a subgroup with the Cagayan Valley languages.

TABLE 11. SUMMARY OF PHONOLOGICAL EVIDENCE

	NE Luzon	Cagayan Valley	Others
1) *R > /g/	YES	YES	GCPH; Ilokano and Arta have /g/ and /t/
2) <i>t/s</i> metathesis	YES (sporadic)	YES (sporadic)	All Northern Luzon
3) Sporadic *r > /l/	YES (sporadic)	YES (sporadic)	CPH, GCPH
4) Gemination after schwa	YES	YES	Ilokano, others
5) *j and *d merger	YES	NO	Arta and Ilokano
6) LVF	YES	NO	Other Negrito Filipino languages
7) *r > /h/	Agta only	NO (only CCAGTA)	NO
8) Sporadic *s > /h/	YES (not in CAS)	NO	VARIOUS: Kalanguya (SCORD), Balangaw (CCORD)
9) Loss of contrastive stress	YES	NO (only CCAGTA)	Pangasinan (Zorc 1979) and SCORD (Himes 1998)
10) Diph. > Monoph.	NO (only CAS and NAG)	YES	Not in N. Phil.
11) *a and *ə merger	NO (only DUP)	YES	Not in N. Phil.
12) *s > /t/, *ti > /s/	NO	YES	Not in N. Phil.
13) *p > f / _u	NO	YES	Not in N. Phil.
14) *b > h / _u	NO	YES	Not in N. Phil.
15) *-d- > -r- / V _V	NO	YES	Tagalog and others
16) *d > j, z, h / _i	NO	YES	Not in N. Phil.
17) *maR- > *maC ₁ / _C ₁	NO	YES	Not in N. Phil.
18) *uy > i / _#	NO	YES	Not in N. Phil.
19) Lenition of *g, *j, *R / V _V	NO	YES	Not in N. Phil.
20) -ʔ# epenthesis	YES	NO (only CCAGTA)	NO

4.2 MORPHOLOGICAL EVIDENCE. Just as there is little phonological evidence for subgrouping the NELUZ languages with the Cagayan Valley languages, there is likewise little unambiguous morphological evidence for such a relationship.

The NELUZ oblique case marker **ta* is probably borrowed from a Cagayan Valley language in which **s* is regularly reflected as /*t*/. Although Tharp (1974a) reconstructs the oblique as **sa* for the Cagayan Valley languages, the form is *ta* in Central Cagayan Agta, Atta, and Ibanag. In fact, none of the Cagayan Valley languages in Tharp's study have a synchronic form *sa*, but Tharp reconstructs **sa* based on external evidence and because *ta* in these languages derives regularly from **sa*. However, **ta* should probably be reconstructed for PNCORD and **sa* for pre-NCORD.

The NELUZ languages share the reduction of the 2PL.GEN pronoun *=*muyu* to *=*muy* with CAGTA, but given the close borrowing relationship between CAGTA and DUP, this could easily be a borrowing into CAGTA.

4.3 LEXICAL EVIDENCE FOR THE POSITION OF NELUZ. Himes (n.d.) lists 326 proposed lexical innovations for Proto-Cordilleran or "Proto-Northern Philippines" (although it is not clear what the difference is between his PCORD and PNP). After excluding morphological innovations already discussed in the previous section, innovations that are repeated (for instance, **ʔagal* 'cry' and **ʔagal* 'weep'), forms that we found to occur in higher-level protolanguages, and forms for which we believe the evidence to be very weak, we considered 273 unique pair-meaning innovations, of which 18 forms (140)–(157) were shared with at least two NELUZ languages. Some of these are entirely unique lexical items, some involve unique semantic shifts, and others reflect unique phonological innovations.

- (140) PCORD **ʔatəd* 'to give' (semantic shift < PAN **SatəD* 'escort') > DUP *atád*, PAH, PAR, CAS, NAG *atəd*, KAS *átəd*
- (141) PCORD **dakəl* 'ancestors' (semantic shift < PPH **dakəl* 'large') > DUP *dadákəl*, PAH *dəddikəl*, NAG *dədikəl* 'parents' (cf. PNELUZ **da[k]kəl*, 'large')
- (142) PCORD **dakəl* 'flood' (semantic shift < PPH **dakəl* 'large') > DUP *dakəl* i *dinum*, PAR *dikəl* a *dinum* (cf. PNELUZ **da[k]kəl* 'large' + **dinum* 'water')
- (143) PCORD (also PNCORD) **dutdut* 'feather, body hair' > PAH *dútdut* 'body hair', PAR, KAS *dútdut* 'body hair, feather', DIN, CAS *dutdut* 'feather'
- (144) PCORD **bu(st)(əu)y* 'calf of leg' > PAH, DIN, CAS *butóy*, NAG *butúy*
- (145) PCORD **baŋləs* 'rotten' > PAH *bɨŋləs*, NAG *bəŋləs* 'rotten, of meat'
- (146) PNP **piklat* 'scar' (phonological innovation < PAN **pil(ae)k*) > DUP, PAR *péklat*, PAH *píklat*, CAS, KAS *peklát* (cf. NAG *pilá?*) (also Tagalog *peklat*, probably a borrowing from a Northern Philippine source)
- (147) PCORD **kəməl* 'squeeze' > PAH *kaməl-kaməl-ən*, PAR *kəmməl* 'squeeze, check for softness'
- (148) PCORD **lətəg* (PNCORD **ləttəg*) 'swell' > DUP *linómtag* (< ***l<um>attag*), NAG *linúmtəg*
- (149) PCORD **p(əu)dəw* (PNCORD **pudəw*) 'white' > PAH, PAR, DIN, KAS *pudéw*

- (150) PCORD **laman* ‘wild pig’ (semantic innovation < PPH ‘flesh’) > DUP, PAH, CAS, NAG *lamán*
- (151) PCORD **s(iu)gəm* (PNCORD **təggəm*, PMP **s(iə)jəm*) ‘ant’ > PAH *səggám* ‘small, red ant’; CAS *səgám* ‘ant (general)’; KAS *əgəm* ‘large, red ant’
- (152) PCORD **yupyup* ‘blow’ (innovation < PAN **iyup*) > PAR, KAS *yópyop*
- (153) PCORD **yəgyəg* ‘earthquake’ > DUP *yágyag*, PAH, DIN *yógyog*, CAS, KAS *yəgyəg*, ‘shake’
- (154) PCORD (also PNCORD) **tukak* ‘frog’ > DUP, PAH, PAR, DIN, CAS, KAS, NAG *tukák*
- (155) PCORD **s(ai)luk* (phonological innovation < PPH **siluŋ*) ‘space under house’ > PAH *sahók*, KAS *sarók*
- (156) PCORD (also PNCORD) **takdəg* ‘stand’ > DUP, PAH, PAR, CAS, KAS, NAG *taknəg*³⁵
- (157) PCORD **tubuŋ* ‘water container’ > PAR, KAS *tubúŋ* ‘bamboo water scooper’

The following form (158) appears to be shared with PCORD/PNPH, but has undergone a semantic shift unique to the NELUZ languages.

- (158) PCORD (also PNCORD) **ʔubət* ‘anus, buttocks’³⁶ > PNELUZ **ubát* ‘vulva’ > DUP, PAR *ubát*, PAH *ubbát*, CAS, KAS, NAG *ubát*

The following five PCORD/PNPH lexical innovations (159)–(163) were shared only with DUP, for which we have a much richer lexical database. We have eliminated a number of forms that appeared to be borrowings from a non-NELUZ language (either ILK or CCAFTA, the two major sources of borrowing for DUP).

- (159) PCORD **salug* (PNCORD *talug*) ‘swim’ > DUP *sulóg* (semantic innovation < ‘river’)
- (160) PCORD **sakbat* ‘carry (on shoulders)’ (phonological shift < PAN **sabat*) > DUP *sakbét* ‘carry by slinging over the shoulder’
- (161) PCORD (also PNCORD) **bukəl* ‘round’ > DUP *nag-bukal* (cf. DUP *bukal* ‘seed’; < PMP **bukəl* ‘seed’)
- (162) PCORD **tapəw* ‘float’ > DUP *tappáw*
- (163) PCORD **gudwa* (PNCORD **gədduwan*) ‘split’ > DUP *gaddúwa* ‘half’ (based on **dua* ‘two’)

It seems, then, that there is a sufficient amount of lexical evidence to support placing the NELUZ languages within the Northern Luzon subgroup, but it appears that they constitute a primary branch of Northern Luzon, and do not immediately subgroup with the CV languages.

5. CONCLUSION. This paper has attempted to fill a gap in the existing literature by presenting data on the Northeastern Luzon subgroup of languages. By examining the phonological, lexical, and functor innovations, we have shown that these languages form

35. This is one of the handful of forms where DUP retains the schwa.

36. Himes includes ‘vagina’ as one of the meanings, but this seems to be based on the CAS reflex.

an independent subgroup, as has been previously suggested. Moreover, although it has long been assumed that the Northeastern Luzon languages subgroup with the primarily non-Agta languages of the Cagayan Valley, we have shown that this is probably not the case, due to the fact that there are no exclusively shared innovations between the NELUZ languages and the Cagayan Valley languages. As such, there can be no “Northern Cordilleran” subgroup (traditionally assumed to consist of the Cagayan Valley and Northeastern Luzon languages). There is some evidence that the NELUZ languages belong to the Northern Luzon (Cordilleran) subgroup, in which case both the Northeastern Luzon subgroup and the Cagayan Valley subgroup would form primary branches of Northern Luzon. However, since relatively few innovations have been identified for Northern Luzon (Reid 2006 notwithstanding), a more definite placement of these languages awaits further research on the larger Northern Luzon subgroup.

APPENDIX 1. LIST OF ABBREVIATIONS

AV	actor voice	NCORD	Northern Cordilleran
CAS	Casiguran Agta	NEG	negative
CCAGTA	Central Cagayan Agta	NELUZ	Northeastern Luzon
CCORD	Central Cordilleran	NOM	nominative
CV	Cagayan Valley	NONREF	nonreferential
CONT	continuative	OBL	oblique
DEF	definite	PAH	Pahanan Agta (Palanan Dumagat)
DIN	Dinapigue Agta	PAN	Proto-Austronesian
DU	dual	PAR	Paranan
DUP	Dupaningan Agta	PCCORD	Proto-Central Cordilleran
EXC	exclusive	PERS	personal name marker
EXT	existential	PMP	Proto-Malayo-Polynesian
GEN	genitive	PNCORD	Proto-Northern Cordilleran
GCPH	Greater Central Philippines	PNELUZ	Proto-Northeastern Luzon
ILK	Ilokano	PNLUZ	Proto-Northern Luzon
IMPER	imperative	PNPH	Proto-Northern Philippine
INC	inclusive	POSS	possessive
INCOM	incompletive	PPH	Proto-Philippine
INDEF	indefinite	PV	patient voice
INF	infinitive	PV2	secondary patient voice
KAS	Kasiguranin	SCORD	Southern Cordilleran
LOC	locative	SPKR	speaker
LSTNR	listener	TAG	Tagalog
LV	locative voice	TOP	topic
LVF	low vowel fronting	UDGT	Umiray Dumagat
NAG	Nagtipunan Agta	+	loanword

APPENDIX 2. FUNCTOR SETS

1. PRONOUNS

	DUP	PAH	PAR	DIN	CAS	NAG	PNELUZ	KAS
TOPICALIZED NOMINATIVE								
1SG	hikán	hikón	hikón	hikón	sakøn	sakøn	*si-akøn	akú
2SG	hikáw	hikó	hikó	hikó?	sikó	sikó?	*si-kaw (>*-ko)	ikáw
3SG	hikúna	siyá	siyá	sya	siyá	siyá	*siya	siyá/sya
1EXCL	hikamí	hikamí	sikamí	—	sikamí	sikamí	*si-kami	kamí?
1INCL.DU	hikitá	hikitá	sikitá	hikitá	sikitá	sikitá	*si-kita	kitá
1INCL.PL	hikitám	hikitám	sikitám	hikitám	sikitám	sikitám	*si-kitam	táyú?
2PL	hikám	hikám	sikám	hikám	sikám	sikám	*si-kam	kayú?
3PL	hidí	hidí	hidí	hidí	sidi	sidi	*sidi	silá?
NOMINATIVE								
1SG	=ak	=æk	=æk	=æk	=æk	=æk	*=æk	akú
2SG	=ka	=ka	=ka	=ka	=ka	=ka	*=ka	ka
3SG	(hikúna)	siyá	siya	sya	siyá	siyá	*siya	Ø/sya/siyá?
1EXCL	=kami	=kamí	=kamí	=kámí	=kamí	=kamí	*=kami	kamí?
1INCL.DU	=kitá	=kitá	=kitá	=kitá	=kitá	=kitá	*=kita	kitá
1INCL.PL	=kitám	=kitám	=kitám	=kitám	=kitám	=kitám	*=kitam	táyú?
2PL	=kam	=kam	=kam	=kam	=kam	=kam	*=kam	kayú?
3PL	hidí	hide	hidí?	=hide	sidé	sidi?	*side	silá?
GENITIVE								
1SG	ko	ko	ku	ku	ku	ku	*=ku	kú?
2SG	mo	mo	mu	mo	mu	mu	*=mu	mú?
3SG	na	na	na	na	na	na	*=na	niyá?
1EXCL	mi	mi	mi	mi	mi	mi	*=mi	mé?
1INCL.DU	ta	ta	ta	ta	ta	ta	*=ta	tá?
1INCL.PL	tam	tam	tam	tam	tam	tam	*=tam	tam
2PL	moy	moy	moy	moy	moy	moy	*=muy	muy
3PL	di	di	di	di	di	di	*=di	dé?/nilá?
OBLIQUE								
1SG	nikán	nikøn	dikóku?	nikøn	diyakón	diyakøn	*ni/di-akøn	sa ákin; sakóko?; kaóko
2SG	nikáw	nikó?	dikómu?	nikó?	dikó	dikó?	*ni/di-kaw (>*-ko)	sa iyú?; kómo?; kaómo?
3SG	nikúna	nikóná ?	dikóna?	nikuná ?	diyá	diyá?	*ni/di-ko-na, SNEL *diya	sa kanyá?; kónya; koniyá?
1EXCL	nikamí	nikamí?	dikómi?	—	dikamí	dikamí?	*ni/di-kami	sa ámin; sakóme?
1INCL.DU	nikitá	nikitá	dikóta?	—	dikitá	—	*ni/di-kita	saáta?; sakóta?; kootá?
1INCL.PL	nikitám	nikitám	dikótam	—	dikitám	—	*ni/di-kitam	sa átam; sa kótam
2PL	nikám	nikám	dikómoy	nikám	dikám	dikám	*ni/di-kam	sa ingy; sa ómoy; sa komóy
3PL	nidí	nidé?	dikódi?	nidé?	didé	didí?	*ni/di-de	sa kanilá, kaónila, sakaóde?
1SG>2SG*	taka	taká	taká?	taká	taká	taká	*taka	—

* In both DUP and CAS, *taka* is used for 1SG>2SG, while *takám* is used for 1SG>2PL (Robinson 2011, Headland and Healey 1974). We did not elicit 1SG>2PL, so we lack data for this in the other languages of NELUZ.

2. DEMONSTRATIVES

	DUP		PAH	PAR	DIN	CAS	NAG	PNELUZ	KAS (TAG-based)
	SG.	PL.							
TOPICALIZED NOMINATIVE									
SPKR/ LSTNR	idé	dagínde	saiyé	(sa) iyən	saiyé	səyé?	saiyí?	*sa-ʔiye	—
LSTNR	idáy	dagínday	saéna?	saiyən(/d)	—	səiná?	saína?	*sa-ʔiná?	—
FAR	idó	dagínto	—	—	—	—	—	—	—
NOMINATIVE									
SPKR/ LSTNR	idé	dagínde	iyé	iddi	iyé	yé?	yi	*ye	itó, 'tu
LSTNR	idáy	dagínday	iná?	yúdi? <iyən>	iná?	iná?	iná?	*ina?	iyán, 'yan
FAR	idó	dagínto	itúd	itúd <iyud>	itúd	yód	—	*i[t]ud	'yun
GENITIVE									
SPKR/ LSTNR	na idé	na dagínde	niiyé	iddi	naiyé	(na)iyé?	(na)yi	*na[i]yé?	nitú
LSTNR	na idáy	na dagínday	nan-iná?	hidhaən	nainá?	(na)iná?	naína?	*nainá?	niyán
FAR	na idó	na dagínto	() itúd	—	—	(na)yód	—	*nai[]úd	—
OBLIQUE									
SPKR/ LSTNR	ihé	—	hé?	həddi	hé?	sé?	sí?	*se?	ditú
LSTNR	iháy	—	hiná?	haən, hən, had	hiná?	siná?	šiyən, siná?	*sina?	diyán
FAR	ihó	—	hud	hud, haód	hud	sá?; sod	—	*sud	doón
LOCATIVE									
SPKR/ LSTNR	(i)hé	—	—	haddi	—	taiyé?	—	—	nandító
LSTNR	(i)háy	—	—	haan	—	tainá?	kainá?	—	nandyán
FAR	(i)hó	—	—	húdi	—	taiyód; taiyá?	kayá, taiyá?	—	nandón

3. NEGATORS

	DUP	PAH	PAR	DIN	CAS	NAG	PNELUZ	KAS
VERBAL	awán	(a)wán	awán	awán	awán	awán	(*awan)	hindé?; di
IMPERATIVE	awan	dyan	diyán	dyán	diyán	diyán	*diyán	wag
EQUATIONAL	awán (bakkán)	bakən	bakən	bakən	bəkən~ bakən	bəkən	*b(ə)kən	hindé?; di
EXISTENTIAL	awán (limus)	awán	awán	awán	awán	awán	(*awan)	walá?
DON'T LIKE	madi(an)	uməd	uməd, sála	uməd	idél, sála	sála	N *uməd, S *sála	ajək, umádi
DON'T KNOW	haybas (awan ko kantandi)	awán ku tukóy	áhay	aháy	ahóy	awán ku, tukúy	*ahəy	aháy

4. VERBAL CONJUGATIONS

	DUP	PAH	PAR	DIN	CAS	PNELUZ	KAS
ACTIVE VOICE WITH <um>							
INF	-um-	-um-	-um-	-um-	-um-	*<um>	-um-
COMP	-inum-	-umm-, -imm-	-inum-	-inum-	-inum-	*<inum>	-inum-, ni-...-um-
CONT	(-um-)*	(-um-)	CumV-	(-um-)	CumV-	—	mæg-CV-...-ən
INCOM	(-um-)	(-um-)	(-um-)	(-um-)	-um-	*<um>	-um-
ACTIVE VOICE WITH mag-							
INF	mag-	mag-	mag-	mag-	mag- ~ mæg-	*mag-	mag-
COMP	nag-	nag-	nag-	nag-	nag- ~ næg-†	*nag-	minag-
CONT	(mag-)	(mag-)	(mag-)	(mag-)	(mag- ~ mæg-)	—	—
INCOM	(mag-)	(mag-)	(mag-)	(mag-)	(mag- ~ mæg-)	*mag-	—
PATIENT VOICE							
INF	-an	-ən	-ən	-ən	-ən	*-ən	-ən
COMP	-in-	-in-	-in-	-in-	-in-	*<in>	-in-, ni-
CONT	—	(-in-)	CV-...-ən	(-ən)	CV-...-ən	—	Cin-; CV-...-ən
INCOM	(-an)	(-ən)	(-ən)	(-ən)	-ən	*-ən	-ən
LOCATIVE VOICE							
INF	-an	-an	-an	-an	-an	*-an	-an
COMP	-in-...-an	-in-...-an	-in-...-an	-in-...-an	-in-...-an	*<in>... -an	ni-/in-...-an
CONT	—	(-an)	CV...-an	(-an)	CV...-an	—	CV-...-an
INCOM	(-an)	(-an)	(-an)	(-an)	-an	*-an	-an
SECONDARY PATIENT VOICE							
INF	i-	i-	i-	i-	i-	*i-	i-
COMP	ni-	in-	in-	ni-	ní-	*ni-	ni-
CONT	—	(in-)	CV-	(i-)	CiyV-/iCu-	?	i-CV-
INCOM	(i-)	(i-)	(i-)	(i-)	i-	*i-	—

* Optional reduplication is possible to emphasize the continuative in DUP actor voice (see Robinson 2011:126)

† Headland and Healey (1974) note the existence of *minag-* and *minæg-*, but these did not occur in our data.

5. EXISTENTIALS AND QUANTIFIERS

	DUP	PAH	PAR	DIN	CAS	NAG	PNELUZ	KAS
HAVE	atóy	tehud	tehód	te(t); tehúd	teəgse?	əgsé?	te N	*tehud me, mérun <terun, teruun>
DON'T HAVE	awán (límus)	awán	awán	awán	awán	awán	(*awan)	walá?
MANY, A LOT	makpál	makpól <makpal, cf. kuməppal>	makpál	makpál	meádu	meádu	N *makpal S *me?adu (cf. UDGT) (ILK adu)	marámi
FEW, A LITTLE	ballék	bádit	bádit	bəllik	kətihək	kəttihək	*b(aə)ddit *bəllik	sabəddit

6. INTERROGATIVES

	DUP	PAH	PAR	DIN
WHAT	ánya, héya	ánya <ənya>	ánya?	ánya
WHO (NOM)	héya	(ti)déya?	tidé(ya)	tidéya
WHOSE	makin; heya i makin...	kandéya	kandéya?	(tidéya?)
WHEN (PAST)	hángan	nengkán	nikán	nikasangán
WHEN (FUT)	hángan	nikán	nikán	nikasangán
WHERE (PAST/FUT)	hádyá	hidyá	hadya	hadya
WHERE (PRES.)	hénan, hádyá	hadya	hadya	hadya
FROM WHERE	taga hadya	taghídyá	tagahádyá	tagahadya
WHY	apay	bákin	bákit	bákit
HOW (MANNER)	panyan	kónya	kúdyá	(anya)
HOW MUCH (EXTENT)	kasano*	sakónya...ti ka-	sakúdyá...ti ka-	gasíno
HOW MUCH (\$)	hángan	sángan	tagsángan	sángan
HOW MANY	hángan a pirásu	sángan a pirásu?	sángan a pirásu	sángan
WHICH	(anya)	ánya...ti duwwa	hadya ti duwa'y	hadyá...ta dúwwa?(é)
HOW MANY TIMES	naminhángan	pensángan	nakasángan	nakasángan
TO WHERE		hidyá	hadya	hadyá
WHAT	CAS ánya	NAG anya	PNELUZ *anya	KAS ano
WHO (NOM)	teesiyá?	syesiyá	N *tidéya S *[]siya	sinu
WHOSE	kiniasiyá?	(syesiyá)	PP *kandéya	kanino
WHEN (PAST)	nikəsiyá?	nikasya ~ nikasye	PP = *nikan S *nikasiya	ka-ilán
WHEN (FUT)	nikəsiyá?	nakasya ~ nikasye		ka-ilán
WHERE (PAST/FUT)	(ta)ahé	ahé	*hadya	sáan
WHERE (PRES.)	ahé	ahé	*hadya	nasaán
FROM WHERE	tagaahé	tagaahé	*taga-...	tagasáan
WHY	bákit, ataáy	ataáy, atáe	—	bákit
HOW (MANNER)	pakódya	pinakódya	*pakodya	paáno
HOW MUCH (EXTENT)	kódya	kódya...ka-	*[sa]kodya	sapaáno
HOW MUCH (\$)	tígsángan	sángan	*t(ia)gsángan	təgsángan
HOW MANY	sángan	sangán a pirásu?	*sángan	sangán a mómon
WHICH	ahé ta adúwa?	ahé	=‘where’	alín sa dúwwa?
HOW MANY TIMES	nakasángan	nakasángan	*naka-sángan	nakasángan
TO WHERE	ahé	taahé	=‘where’	sáan

* *Kasano* is an Ilokano loan.

7. ADVERBS OF TIME

	DUP	PAH	PAR	DIN
DAY BEFORE YESTERDAY	ha essa nakkápon	ten essa aldew	tu issa a aldíw	tu essa aldíw
YESTERDAY	nakkápon	nennápon	tuápon	tuápon(he?)
LAST NIGHT	ha kalláp hé(ya)	nenggíbi?	tugíbi(he?)	—
EARLIER	nak(k)óya	nemmamaya?	tumámaya?	tumámaya(hé?)
TODAY, NOW	(ni)yé?in	nadíd	nadíd	nadíd
LATER	ayénan, nokkán	mamaya	mamaya?	mamaya?
TOMORROW	no?úgma	niiláw	niiláw	nugágabi?
DAY AFTER TOMORROW	mekadúwa a pamalák	puwéra niiláw	pwéra niiláw	nu éssa aldíw
MORNING	mallédum	dímadímang	gagábi?	gágábi?
NOON	pamalák	tangháli?	tangháli?	tánghali
AFTERNOON	ápon	apón	apon	ápón
NIGHT	kállap	gíbi?	gíbi?	kálláp
MIDNIGHT	lúbok na kalláp	(gíbi?)	hatinggabi	hatinggabi
EARLY MORNING	pómsag	(dímadímang)	kaldíwán	ésding aldíw ðan
DAY	pamalak, aldew	aldew	aldíw	aldíw
DAY BEFORE YESTERDAY	CAS to esá a aldéw	NAG tu éssa a aldó?	PNELUZ —	KAS nung éssang aldéw
YESTERDAY	to apón	tuápon (sé?)	*tu-?ápon	kahápon
LAST NIGHT	tukəláp(se?)	tukəláp(sé?)	—	nunggíbi?
EARLIER	tomamayasé?	tumámaya(sé?)	+tu-mámaya[se?]	kanína
TODAY, NOW	nadíd	nadíd	*nadíd	ngay-ún
LATER	mamaya?	mamaya?	+mamaya?	mamaya
TOMORROW	niiláw, tagabiyá?	tagabiyá?	*ni?iláw, S *tagabiyá?	niiláw
DAY AFTER TOMORROW	(sa)péra? niiláw	ta issa a gagábi?	+pwéra-ni?iláw	pwéra niiláw
MORNING	gagábi?	gagábi?	*gagábi?	umága?
NOON	tangháli?	tangháli?	+tangháli?	tangháli?
AFTERNOON	apón	apón	*apon	ápon
NIGHT	kəláp	kəláp	*kəláp, *gabi?	gíbi?
MIDNIGHT	hatinggabi	—	+hatinggabi	hatinggíbi?
EARLY MORNING	amulaldéw	—	—	amulaldéw
DAY	aldéw	aldó?	*aldaw	aldéw

8. PARTICLES

	TAG	ILK	DUP	PAH	PAR
ALSO	din	=met	=bi	—	—
DISCOVERY	pala	=gayam	=noman	bal	bəl
EMPHASIS	ngá?	=ngarud	=man	—	ngani?
FIRST	múna	=pay=la(e)ng	=pala*	pallá?	palla
ONLY	lang, lámang	=la(e)ng	=la	—	la?
POSSIBLE	yáta	—	=wade(n), =wan	—	(hu)wád
REQUEST	—	paki-	—	(bénid)	pəd
AS IF	párang	kas-la...	konna ha..	—	—
QUOTATIVE	daw	=kan(o)	=kan	—	kan
ALREADY	na	=en, =on	=dan ('-n, '-on)	dən	dən
AGAIN	muli	manen	=*manon	ruwáy	huwáy
EVEN	pati	uray	mensan	patí	patí
	DIN	CAS	NAG	PNELUZ	KAS
ALSO	—	béman bé?	bí?	*bi	din/rin
DISCOVERY	—	balé?	bal	*bəl[i]	balí?
EMPHASIS	—	ngani	ngáni (dən)	*ngani	ngáni
FIRST	la?	pa san	pa san	*palla	muná
ONLY	la?	san	san	*la	laang
POSSIBLE	—	wadí?	wadí?	*wad[i]	yáta
REQUEST	—	paki-...pad	pad	*p(æ)d	paki-...ned
AS IF	kumón a...	kumón sa?; kumóna?	koman ti...	S *kumón	kumón
QUOTATIVE	kan	kan	—	*kan	daw
ALREADY	dən/rən	—	—	*dən	na
AGAIN	huwáy	huwáy	huwáy	*ruwáy	ruwáy
EVEN	(kahít)	páti	patí	*pati	patí

* Although the /l/ in this form most commonly occurs as a singleton, it can occur as a geminate.

REFERENCES

- Bellwood, Peter. 1997. *Prehistory of the Indo-Malaysian archipelago*. Honolulu: University of Hawai'i Press.
- Blust, Robert A. 1972. Note on PAN *qa(R)(CtT)a 'outsiders, alien people'. *Oceanic Linguistics* 11:166–71.
- . 1983–84. Austronesian etymologies II. *Oceanic Linguistics* 22–23:29–149.
- . 1987. Lexical reconstruction and semantic reconstruction: The case of Austronesian "house" words. *Diachronica* 4:79–106.
- . 1991a. Patterns of sound change in the Austronesian languages. In *Patterns of change, change of patterns: Linguistic change and reconstruction methodology*, ed. by Philip Baldi, 129–65. Berlin: Mouton de Gruyter.
- . 1991b. The Greater Central Philippines hypothesis. *Oceanic Linguistics* 30:73–129.
- . 1999. Subgrouping, circularity and extinction: Some issues in Austronesian comparative linguistics. In *Selected papers from the Eighth International Conference on Austronesian Linguistics*, ed. by Elizabeth Zeitoun and Paul Jen-Kuei Li, 31–94. Taipei: Academia Sinica.

- . 2000. Why lexicostatistics doesn't work: The 'universal constant' hypothesis and the Austronesian languages. In *Time depth in historical linguistics*, vol. 2: *Papers in the prehistory of languages*, ed. by C. Renfrew, A. McMahon, and L. Trask, 311–31. Cambridge: The McDonald Institute for Archaeological Research.
- . 2005. The linguistic macrohistory of the Philippines: Some speculations. *Current issues in Philippine linguistics and anthropology: Parangal kay Lawrence A. Reid*, ed. by Hsiu-chuan Liao and Carl R. Galvez Rubino, 31–68. Manila: Linguistic Society of the Philippines and SIL Philippines.
- . 2006. Anomalous liquid: Sibilant correspondences in Western Austronesian. *Oceanic Linguistics* 45:210–16.
- . 2009. *The Austronesian languages*. Canberra: Pacific Linguistics.
- . 2010. The Greater North Borneo hypothesis. *Oceanic Linguistics* 49:44–118.
- Blust, Robert A, and Stephen Trussel. In progress. *Austronesian comparative dictionary, web edition*. Available at: <http://www.trussel2.com/ACD/>. Accessed December 20, 2012.
- Charles, Mathew. 1974. Problems in the reconstruction of Proto-Philippine phonology and the subgrouping of the Philippine languages. *Oceanic Linguistics* 13:457–509.
- Conant, Carlos E. 1910. The RGH law in Philippine languages. *Journal of the American Oriental Society* 31:70–85.
- Davis, Philip W., and Angela D. Mesa. 2000. *A dictionary of Yogad*. Munich: Lincom Europa.
- Dyen, Isidore. 1965. *A lexicostatistical classification of the Austronesian languages*. New Haven: Yale University.
- . 1972. Non-gradual regular phonetic changes involving sibilants. In *Langues et techniques, nature et société*, vol. 1: *Approche linguistique*, ed. by Jacqueline M. C. Thomas and Lucien Bernot, 95–99. Paris: Klincksieck.
- Fox, Robert B., Willis E. Sibley, and Fred Eggan. 1965. A preliminary glottochronology for Northern Luzon. *Asian Studies* 3(1):103–13.
- Goda, Toh. 2003. *Postcolonialism and local politics in Southeast Asia*. Quezon City: New Day Publishing.
- Headland, Thomas N. 1975. Report of Eastern Luzon language survey. *Philippine Journal of Linguistics* 6(2):47–54.
- . 1997. Revisionism in ecological anthropology. *Current Anthropology* 38(4):605–30.
- Headland, Thomas N., and Janet D. Headland. 1974. *A Dumagat [Casiguran]–English dictionary*. Canberra: Pacific Linguistics.
- Headland, Thomas N., and Alan Healey. 1974. Grammatical sketch of Casiguran Dumagat. *Papers in Philippine Linguistics* 6:1–54.
- Headland, Thomas N., and Elmer P. Wolfenden. 1967. The vowels of Casiguran Dumagat. In *Studies in Philippine anthropology*, ed. by Mario D. Zamora, 592–96. Quezon City: Alemar Phoenix.
- Himes, Ronald S. 1998. The Southern Cordilleran group of Philippine languages. *Oceanic Linguistics* 37:120–77.
- . 2002. The relationship of Umiray Dumagat to other Philippine languages. *Oceanic Linguistics* 41:275–94.
- . n.d. NPH reconstructions. MS.
- Hontiveros, Greg. 2000. *Butuan of a thousand years*. Butuan City (Philippines): Butuan City Historical & Cultural Foundation, Inc.
- Lewis, M. Paul, ed. 2009. *Ethnologue: Languages of the world*. 16th ed. Dallas: SIL International.
- Liao, Hsiu-chuan. 2005. Pronominal forms in Central Cagayan Agta: Clitics or agreement features? In *Current issues in Philippine linguistics and anthropology: Parangal kay*

- Lawrence A. Reid, ed. by Hsiu-chuan Liao and Carl R. Galvez Rubino, 346–60. Manila: Linguistic Society of the Philippines and SIL Philippines.
- Lobel, Jason William. 2010. Manide: An undescribed Philippine language. *Oceanic Linguistics* 49:480–512.
- . 2011. Inagta Alabat: Another undescribed Philippine language. MS.
- . 2012. Philippine and North Bornean languages: Issues in description, sub-grouping, and reconstruction. PhD diss., University of Hawai'i.
- Lobel, Jason William, and Labi Hadji Sarip Riwarung. 2011. Maranao: A preliminary phonological sketch with supporting audio. *Language Documentation and Conservation* 5: 31–59.
- Lobel, Jason William, and Laura C. Robinson. n.d. Kasiguranin: Notes on the position of a Philippine contact language. MS.
- Mayfield, Roy. n.d. *Gattaran Agta dictionary*. Typescript. 350 pp.
- McFarland, Curtis D. 1974. The dialects of the Bikol area. PhD diss., Yale University.
- . 1980. *A linguistic atlas of the Philippines*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.
- Nickell, Thomas L. 1985. A partial stratificational analysis of Eastern Cagayan Agta language. In *The Agta of northeastern Luzon: Recent studies*, ed. by P. Bion Griffin and Agnes Estioko-Griffin, 119–46. Cebu City: University of San Carlos Publications.
- Nickell, Thomas L., and Kristy Nickell. 1987. Eastern Agta dictionary. Unpublished ms.
- Oates, William J., and Lynette F. Oates. 1955. *A vocabulary of Central Cagayan Negrito*. Manila: Summer Institute of Linguistics.
- . 1958. The phonemes of Central Cagayan Negrito. In *Oceania Linguistic Monographs* 3, ed. by A. Capell and S. A. Wurm, 48–66. Sydney: University of Sydney.
- Reid, Lawrence A. 1974. The Central Cordilleran subgroup of Philippine languages. *Oceanic Linguistics* 13:511–60.
- . 1979a. Towards a reconstruction of the pronominal systems of Proto-Cordilleran, Philippines. In *South-East Asian linguistic studies* 3, ed. by Nguyen Dang Liem, 259–75. Canberra: Pacific Linguistics.
- . 1979b. Evidence for Proto-Philippine nominative marking. *Philippine Journal of Linguistics* 10(1–2):1–20.
- . 1982. The demise of Proto-Philippines. In *Papers from the Third International Conference on Austronesian Linguistics*, vol. 2: *Tracking the travellers*, ed. by Amran Halim, Lois Carrington, and S. A. Wurm, 201–16. Canberra: Pacific Linguistics.
- . 1987. The early switch hypothesis: Linguistic evidence for contact between Negritos and Austronesians. *Man and Culture in Oceania* 3:41–59.
- . 1989. Arta, another Philippine Negrito language. *Oceanic Linguistics* 28:47–74.
- . 1991. The Alta languages of the Philippines. In *VICAL 2, Western Austronesian and contact languages: Papers from the Fifth International Conference on Austronesian Linguistics*, ed. by Ray Harlow, 265–97. Te Reo Special Publication. Auckland: Linguistic Society of New Zealand.
- . 1994. Unravelling the linguistic histories of the Philippine Negritos. In *Language contact and change in the Austronesian world*, ed. by Tom Dutton and Darrell T. Tryon, 443–75. Berlin: Mouton de Gruyter.
- . 2006. On reconstructing the morphosyntax of Proto-Northern Luzon. *Philippine Journal of Linguistics* 37(2):1–63.
- . 2009. The reconstruction of a dual pronoun to Proto Malayo-Polynesian. In *Discovering history through language: Papers in honour of Malcolm Ross*, ed. by Bethwyn Evans, 461–77. Canberra: Pacific Linguistics.

- . 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.
- Robinson, Laura C. 2011. *Dupaningan Agta: Grammar, vocabulary, and texts*. Canberra: Pacific Linguistics.
- Ross, Malcolm D. 1992. The sound of Proto-Austronesian: An outsider's view of the Formosan evidence. *Oceanic Linguistics* 31:23–64.
- Rubino, Carl Ralph Galvez. 2000. *Ilocano dictionary and grammar*. PALI Language Texts. Honolulu: University of Hawai'i Press.
- Scott, William Henry. 1984. *Prehispanic source materials for the study of Philippine history*. Quezon City: New Day Publishers.
- . 1992. *Looking for the prehispanic Filipino and other essays in Philippine history*. Quezon City: New Day Publishers.
- Tharp, James A. 1974a. The Northern Cordilleran subgroup of Philippine languages. *University of Hawai'i Working Papers in Linguistics* 6(6):53–114.
- . 1974b. Notes on the Ilokano reflexes of Proto-Austronesian *R. *University of Hawai'i Working Papers in Linguistics* 6(6):47–51.
- Thomas, David, and Alan Healey. 1962. Some Philippine language subgroupings: A lexicostatistical study. *Anthropological Linguistics* 4(9): 21–33.
- Vanoverbergh, Morice. 1937. *Some undescribed languages of Luzon*. Nijmegen: Dekker and van de Vegt.
- . 1972. *Isneg-English vocabulary*. Honolulu: University of Hawai'i Press.
- Walton, Charles. 1979. A Philippine language tree. *Anthropological Linguistics* 21(2):70–98.
- Wolff, John U. 1973. Verbal inflection in Proto-Austronesian. In *Parangal kay Cecilio Lopez: Essays in honor of Cecilio Lopez on his seventy-fifth birthday*, ed. by Andrew B. Gonzalez, 71–91. Quezon City: Linguistic Society of the Philippines.
- . 2003. The sounds of Proto-Austronesian. In *Issues in Austronesian historical linguistics*, ed. by John Lynch, 1–14. Canberra: Pacific Linguistics.
- Yap, Fe Aldave. 1977. *A comparative study of Philippine lexicons*. Manila: Institute of National Language.
- Zorc, R. David. 1977. *The Bisayan dialects of the Philippines: Subgrouping and reconstruction*. Canberra: Pacific Linguistics.
- . 1978. Functor analysis: A method of quantifying function words for comparing and classifying languages. *The Fifth LACUS Forum*, 510–21. Columbia, SC: Hornbeam Press.
- . 1979. On the development of contrastive word accent: Pangasinan, a case in point. In *Southeast Asian linguistic studies*, vol. 3, ed. by Nguyen Dang Liem, 241–58. Canberra: Pacific Linguistics.