Contact or inheritance? New evidence on the Proto-Philippines hypothesis

A central issue in Austronesian higher-order subgrouping concerns the linguistic position of Philippine languages. Despite considerable debate in recent work, it remains unsettled whether these languages constitute multiple intersecting Malayo-Polynesian primary branches forming a linkage (Ross 2020; Liao 2020; Reid 2020) or descend from a single shared ancestor that is a daughter of Proto-Malayo-Polynesian (Zorc 1986; Blust 2005 et seq.). We offer three new lines of evidence supporting the former interpretation: first, the absence of a true *d/z merger in Central Luzon and Minahasan languages, which undermines the sole phonological basis for positing a Proto-Philippine (PPh) branch; second, the geographical distribution and semantic categories of the purported PPh-defining lexicon, which favor a diffusion-based explanation; third, the distribution of an understudied morphological innovation in Philippine languages, which further suggests frequent contact among 'mainland' languages excluding those in interior and peripheral islands. Not only do these findings indicate the absence of defining innovations for a cohesive PPh, but they also suggest significant horizontal transmission across the Philippine archipelago beyond lexical exchange and the possibility that the *d/*z merger is an areal drift rather than a diagnostic innovation. We conclude that Philippine languages are better understood as representing intersecting MP primary branches, aligning with recent perspectives (Ross 2020; Smith, to appear). Moreover, the high number of lexical innovations attributed to PPh likely reflects contact-driven processes—diffusion, borrowing, and linkage histories—rather than retention from a unified ancestor.

Contents

1 Introduction

Although robust comparative evidence supports the view that all Austronesian languages spoken outside Taiwan descend from a single common ancestor (Dahl 1973; Mills 1975; Blust 1977, 1993; Ross 2002)—Proto-Malayo-Polynesian (PMP), the first-order division of this branch remains contested in recent literature. A focal question concerns the linguistic position of the languages spoken at the first major landing site of the Austronesian diaspora and surrounding islands. Whether or not these languages descend from a single offshoot of PMP or represent multiple primary branches spoken in the Philippines remains a point of contention in the literature (Reid 1982; Blust 2005, 2019 et seq.; Zorc 1986, 2019; Reid 2020; Ross 2005, 2020; Liao 2020, among others).

This debate highlights the limitations of the Neogrammarian approach for classifying closely related languages under mutual contact. Although there is a lack of comparative evidence that Philippine languages subgroup with any other MP subgroup spoken elsewhere (except for Gorontalo-Mongondow, Sangiric, and Minahasan languages of northern Sulawesi) (Blust 1995, 2005; Smith 2017), only a single phonological merger suggests their potential descent from a shared ancestor distinct from PMP. It is therefore challenging to determine their affiliation using the traditional approach to linguistic subgrouping.

What makes Philippine languages particularly theoretically intriguing is the sheer volume of their lexical innovations on the one hand and the absence of phonological evidence for their affinity on the other—over 1,000 shared lexical items have been identified across these languages (albeit with varying patterns of distribution), none of which appear in MP languages outside the Philippines. Several researchers have thus argued for the existence of a shared common ancestor of all modern Philippine languages, Proto-Philippines (PPh), as a distinct primary branch of MP (Blust 2005, 2019, 2020, 2022; Zorc 1986, 2020). This affinity, as proposed in Blust (2005 et seq.), is the result of a historical leveling event that eliminated other PMP descendants in the Philippine island group. However, the lack of non-lexical evidence for this subgroup has led to the alternative view that the shared vocabulary is the outcome of lexical diffusion (Reid 2020; Ross 2020; Liao 2020; see also Smith 2017 for a similar view).

From a theoretical point of view, the challenge lies in identifying true cases of shared innovation among closely related languages with extensive contact and early borrowings. Given Luzon's proximity to the Visayas and Mindanao island groups, one would expect a subgrouping scenario like (1), where PMP speakers gradually diversified into distinct linguistic communities through the Austronesian expansion across the Greater Philippine region (alongside migrations out of the Philippine islands). In principle, these languages therefore represent several independent primary branches of PMP, parallel to ex-Philippine branches. It is also possible that some of the ex-Philippine branches derived from one of the branches that diversified within the Philippine islands, as exemplified by the hypothetical Branch K.

(1) *Hypothetical subgrouping scenario of MP higher-order languages*
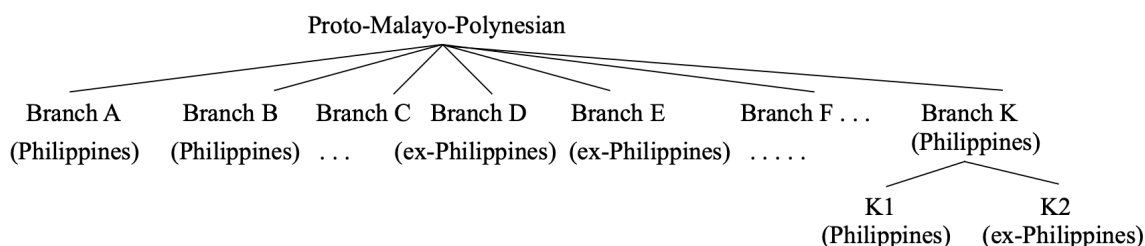
How solid is the evidence for the Philippine microgroups to be subsumed by a common ancestor excluding other MP languages? In this paper, we reappraise previous arguments for and against the

Proto-Philippine hypothesis with new evidence concerning specific aspects of the phonology, lexicon, and morphosyntax of Philippine languages. Three central questions (2a–c) are investigated.

(2) *The key questions for clarifying the PPh hypothesis*
   a. How robust is phonological evidence for a purported Philippine subgroup?
   b. What history can be reconstructed for the distribution of the proposed lexical innovations across the Philippine subgroups?
   c. Can the domain of morphosyntax tell us something new about the relationship of the Philippine languages?

The paper is structured as follows. We begin with a review of the Proto-Philippine hypothesis in Section 2. Section 3 reassesses the validity of PMP *d/z merger as a defining innovation of a purported PPh. Section 4 revisits the claimed lexical evidence for PPh, examining the distribution of these innovations in terms of semantic fields, following Haspelmath and Tadmor's (2009) quantitative study on borrowing. Section 5 analyzes the derivations of Circumstantial Voice morphology across Philippine languages and discusses how it informs the contact history of these languages. Section 6 concludes with the claims that (i) Philippine languages are best viewed as deriving from multiple PMP offshoots that interact as a linkage and (ii) the linguistic features present in modern Philippine languages reflect multiple layers of change, containing retentions from PAn/PMP at the deepest level, overlapping innovations from a linkage history, and borrowed features from later contact among groups.

2   Proto-Philippines: Puzzles and debates

Since Blake (1906), linguistic similarities among Philippine languages have led many to view them as descending from a single protolanguage (Blake 1906; Scheerer 1918; Charles 1974; Llamzon 1975; Paz 1981; Blust 2005, 2019 et seq.; Himes 2012, a.o.). This proposed branch, illustrated in (3), includes several microgroups defined by phonological innovations (Blust 1991, 2005; Himes 2012; Lobel 2013; Reid 2018).

(3) *The interrelationships of Philippine languages after Blust (1991)*
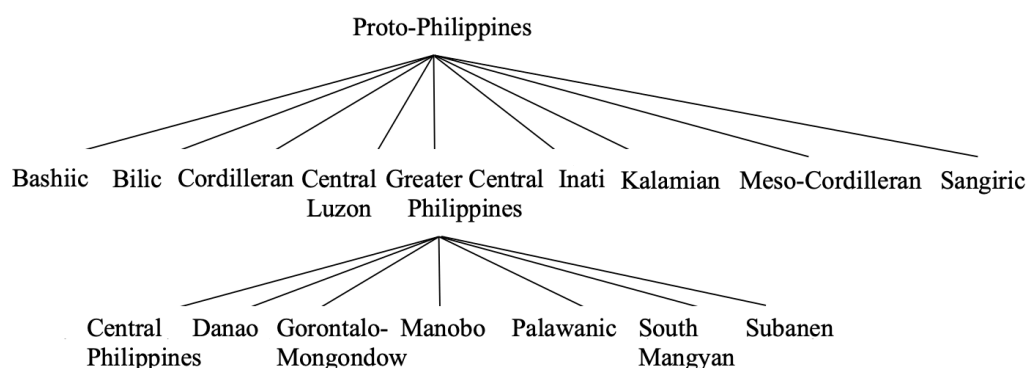
However, it remains unclear whether these microgroups share an ancestor distinct from Proto-Malayo-Polynesian (PMP). Reid (1978, 1982) highlighted the absence of exclusively shared phonological and morphosyntactic innovations for a putative Proto-Philippine branch, suggesting that these languages may not descend from a single source. In response, Zorc (1986) and Blust (2005) presented 327 lexical innovations as evidence for PPh, arguing that these replacements support a common ancestry post-PMP.

4

This proposal has been challenged on two fronts. First, the phonological and grammatical systems of PPh show little innovation and remain largely identical to PMP (REFs). Second, the quality of the proposed PPh-defining lexical innovations is low (Smith 2017:472). Blust (2019) revisited the debate by expanding the list to 1,259 lexical innovations and one phonological innovation—the merger of PMP *z and *d. While Zorc (2020) concurs with Blust, Liao (2020), Reid (2020), and Ross (2020) argue that the overlapping distribution of these innovations points to a linkage history rather than a neatly bifurcating family.

Ross (2020:370) notes that nearly all PPh-defining lexical innovations are not uniformly present in all Philippine languages and may stem from early borrowings, typical of linkage scenarios. Similar signals can arise from language contact, complicating the reconstruction of Philippine linguistic history. Recent Bayesian phylogenetic subgroupings (Gray et al. 2009; King et al. 2024) further support a clustering of Philippine languages without strong evidence for PPh, reinforcing the linkage hypothesis.

A linkage hypothesis implies a rapid expansion of PMP speakers into island Southeast Asia, which left behind an early dialect network in the Philippines. This swift expansion precluded the development of clear innovation-defined subgroups, resulting instead in overlapping innovation patterns (e.g., François 2014:170–171; Ross 1995:45–46).

While Blust acknowledges the presence of linkage-like patterns (2019:183–184, 2020:453–454), the core disagreement remains: Ross (2020:369) views PMP as the ancestor of an early Philippine dialect network, whereas Blust (2020:452) argues for a distinct PPh, noting that its innovations are confined to Philippine languages and absent in those spoken south of the Philippines.

In essence, the competing hypotheses differ in interpreting overlapping innovations. Blust (2020) supports a single, albeit internally diverse, proto-language, while proponents of the linkage history (e.g., Ross 2020) attribute the innovations to PMP, with Philippine languages forming an early dialect chain that differentiated rapidly after other MP speakers exited the region.

The goal of this paper is to evaluate the significance of shared lexical innovations for subgrouping, particularly in the absence of phonological evidence. This is crucial for Austronesian diachronic studies, as many primary branches of MP lack robust phonological support. Establishing clear criteria for using lexical or syntactic innovations in subgrouping is therefore essential.

In what follows, we revisit the debate with three new lines of evidence: phonology, lexicon, and morphosyntax. In Section 3, we highlight the previously overlooked absence of the *d/z merger in several Philippine microgroups. We then examine the semantic categories and geographical distribution of PPh-defining lexical items, showing that they likely reflect borrowing rather than inheritance. Finally, we investigate an understudied variation in Circumstantial Voice morphology, presenting new evidence for multiple layers of borrowing across Philippine subgroups.

3   Is *d/z merger a defining innovation for Proto-Philippines? A reappraisal

We begin by reexamining the sole phonological innovation for the purported Philippine branch: the merger of PMP *d (e.g. *duSa 'two') and *z (e.g. *zalan 'road'). Blust (2019) notes that no language in the Philippines maintains this distinction in stable forms, citing only one exception to this generalization: in Ayta Abellen (Central Luzon), *d and *z remain unmerged in two cases: *qudang > *ulang* 'shrimp' vs. *quzan > *udan* 'rain'. However, he notes that this instance is ambiguous, because "other words [in Ayta Abellen] show a merger of *z and *d word-initially, as in PAn *depa > *depah* 'armsbreadth' and *zaRami > *dayami* 'rice husk'" (Blust 2019: 156–157).

Our investigation into the Austronesian Comparative Dictionary (ACD) and recent literature reveals many more such exceptions in Central Luzon, Minahasan, and Bashiic languages. In what follows, we discuss how these cases constitute strong evidence for *d and *z as remaining unmerged at the proto-level of these microgroups. We then revisit the absence of this merger in Chamorro—the so-called 'western Austronesian outlier' that shares similar phonological innovations with Greater Central Philippine languages—and discuss how the presence of a *d/z distinction these languages informs the linguistic prehistory of the Philippines.

3.1 Unmerged *d and *z in Central Luzon languages

There are three lower-order branches within Central Luzon: Kapampangan, Sambalic, and Hatang Kayi[1], as seen in (4). Out of those subgroups, there are four languages for which a useful amount of data are readily available: Kapampangan (KP), Ayta Abellen (AA), Ayta Mag-Antsi (AM) and Botolan (BT).[2] The interrelationship of these languages is illustrated in (4). The latter three languages are part of the Sambalic subgroup. Due to a lack of ready data, the Hatang Kayi branch is not represented in this paper.

(4)   *Central Luzon subgrouping*
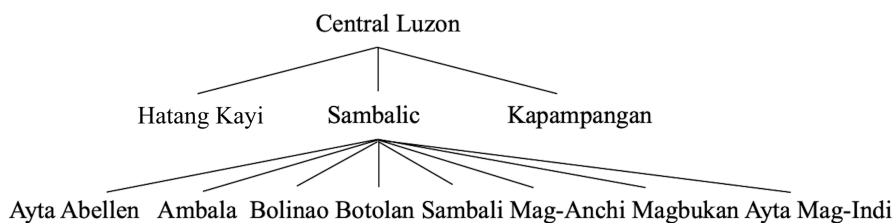
Our examination of the ACD identifies unmerged *d and *z not only in the two Ayta Abellen examples mentioned in Blust (2019), but also in other Central Luzon languages.[3] Importantly, these exceptions are not sporadic or random, but reveal a pattern. The key observations are summarized in (5).

(5)   a. *z is realized as /d/ word-initially in all four languages. Word-medially, it is realized as /d/ in Ayta Abellen and /r/ in the other three languages.

b. In many instances, *d is indeed realized identically to *z, as either /d/ or /r/.

c. Contra Blust's generalization, instances of exceptional *d > l are found in many stable forms (e.g. PAn *daNum > AA *lanom* 'water', PAn *duSa > BT *lua* 'two', and PMP *dateng > AA *lateng* 'to arrive; come').

d. The (unexpected) *d > l sound change occurs vowel-medially in Kapampangan. In the Sambalic languages, it occurs more freely: vowel-medially and word-finally in Ayta Mag-Antsi, and in all environments in Ayta Abellen and Botolan.

e. In Ayta Mag-Antsi, *z and *d appear wholly unmerged. However, there are no instances in that data of word-initial *z. If Ayta Mag-Antsi is in line with the other Central Luzon languages examined in this paper, *z and *d will show partial merging word-initially.

---

[1] Also known by the exonyms Sinauna and Remontado.

[2] More data would be valuable, particularly more instances of *z. A full dictionary of Ayta Mag-Antsi is already available at https://philippines.sil.org/resources/online$_r$esources/sgb, produced by SIL. However, its contents have not yet been sorted and indexed to their proto-forms.

[3] The absence of a true *d/z merger in Ayta Mag-antsi has been briefly discussed in Liao (2020) as well as in Himes (2012). Smith (2017b: 471) also discussed a few cases in other languages: "multiple languages have different reflexes of *z and *d: for example, Itbayaten *doha* 'two' from *duha but *rahan* 'road, path' from *zalan, and Kakilingan Sambal *olang* 'shrimp' from *qudang but *odan* 'rain' from *quzan." Our attempt here is to take a more comprehensive view at the absence of this merger within Central Luzon and across Philippine microgroups. Himes (2012) also independently concludes that *z and *d remain unmerged in Proto-Central Luzon, although not connecting the discussion to the PPh debate. See the appendix for more details.

The table below summarizes the data. The second column lists regular sound change rules which affect both *z and *d in each language; the third column outlines instances where *d diverges and is realized as /l/.[4] In the appendix, we present a hypothesis that the sound change *d > l / V_V occurred in Kapampangan and the Sambalic languages, while *d and *z were unmerged, and the flow-on effects from this sound change resulted in the split reflexes of *d.[5]

(6) *Summary of the reflexes of *d and *z in Central Luzon languages*

Language	Identical reflexes of *d/z	Instances of *d realized as /l/
Kapampangan (64 instances total)	*d > r / V_V (4 instances) *d > d (46 instances) *z > r / V_V (5 instances) *z > d (5 instances)	PMP *badas > *balás* 'sand' PMP *ida > *ila* '3PL' PWMP *ludem > *ma-lúlam* 'cloudy, about to rain' PAN *tuduq > *tulu(ʔ)* 'to drip, flow, spill'
Ayta Abellen (47 instances total)	*d > d (30 instances) *z > d (5 instances)	PAN *daNum > *lanom* 'water' PMP *dateng > *lateng* 'to come' PWMP *sidem (or *silem) > *hilem* 'afternoon' PAN *dengeR > *lenge* 'to hear' PMP *dingding > *lingling* 'wall of a house' PWMP *di hipaR > *lipay* 'the other side of a body of water' PAN *duSa > *lowa* 'two' PAN *qañud > *anol* 'to be carried on the current' PPh *sápad > *sapal* 'hand of bananas' PPh *tadék > *talek* 'dance; to dance' PAN *tuduR > *toloy* 'to sleep' PWMP *pu(n)dut (or *pu(n)zut) > *polot* 'to pick up'
Ayta Mag-Antsi (9 instances total)	*d > r / V_V (1 instance) *d > d (1 instance) *z > r / V_V (3 instances)	PMP *badas > *balah* 'sand' PMP *qalad > *alal* 'fence' PAN *qudang > *ulang* 'squid' PWMP *tadu > *talo* 'beeswax'
Botolan (33 instances total)	*d > r / V_V (4 instances) *d > d (17 instances) *z > r / V_V (2 instances)	PPh *agud > *ágol* 'to moan, agonize' PAN *daNum > *lánum* 'water' PAN *duSa > *lua* 'two' PAN *SateD > *atel* 'escort' PPh *tadék > *talék* 'dance; to dance' PWMP *tadu > *tálo* 'beeswax' PMP *tidaq > *tílaʔ* 'remainder' PMP *tudung > *tolóng* 'head cover' PAN *tuduR > *túluy* 'to sleep' PAN *tuduS > *tóʔol* 'knee' (metathesis)

At first glance, the realizations of *d and *z in the sampled Central Luzon languages seem mixed and unpredictable, but a closer look reveals that *d and *z are realized following their own distinct patterns. First, *z is always realized *consistently*[6], appearing as /d/ in one environment and /r/ in another (see the table above). Second, *d is realized *inconsistently* – in an identical environment, it can either merge with *z, appearing as /d/ or /r/ respectively, or can appear as /l/. Now, if *d and *z were merged at the level of Proto-Central Luzon, we would expect to see the same pattern of realization in their reflexes throughout all the Proto-Central Luzon descendant languages. However, that is not what the data suggest. In all four Central Luzon languages for which sufficient data is available, the reflexes of

---

[4]An anonymous reviewer also provides PWMP *damay 'peace, sympathy' > AA *lamay* 'to attend an all-night wake for a dead person', alongside AA *damay* 'show sympathy'. They suggest that *lamay* is the native Ayta Abellen word, while *damay* is a loan from Tagalog.

[5]All cited data is reproduced from the ACD entries, so some words will be listed as descending from PPh, despite the fact that this paper is arguing against its existence.

[6]PWMP *puzut > AA *polot* seems to be an exception. But *puzut has a disjunct form, *pu(n)dut, both meaning 'to pick up with the fingers'. If *polot* is a reflex of *pu(n)dut there is no inconsistency.

*d show an inconsistent pattern of realization, while the reflexes of *z show a consistent pattern. The conclusion is that *d and *z are in fact unmerged in these languages, strongly suggesting an underlying separation of *z and *d to the level of Proto-Central Luzon.

Two objections may be made against the data presented: first, that the inconsistency in the realization of *d is a result of loans from other Philippine languages; second, that the inconsistency is a result of sporadic change. The first of these objections is the stronger of the two, as the Central Luzon languages, particularly Kapampangan, have been influenced by surrounding languages—for example Tagalog. Due to the overall phonological similarity of Philippine languages, it may also be difficult to identify early loans borrowed prior to certain sound changes.[7] However, both objections fail in the face of the data on *z. If *z and *d were merged in Kapampangan, and the inconsistency in the reflexes of *d is solely a result of borrowing and/or sporadic change, probability demands we ought to see a comparable level of inconsistency in the realization of *z. In other words, if *z and *d had already merged into a united phoneme in Proto-Central Luzon, then we would expect words with *z or *d to be replaced by borrowing and/or sporadic change at roughly equal rates. However, all the reflexes of *z across the four languages studied are realized according to a regular rule, while 23% of the reflexes of *d (71% if only counting vowel-medial reflexes) appear as the divergent reflex /l/.[8] These facts strongly suggest that the difference in the realizations of *z and *d is not the result of borrowing or sporadic change, but the result of an underlying division between *z and *d in Proto-Central Luzon.

If *z and *d were already merged in Proto-Central Luzon, we would expect the patterning of the reflexes to be identical for *z and *d, but that is not the case. This difference in reflex patterning strongly suggests an inherited separation between *z and *d in Proto-Central Luzon, undermining the viability of the *d/z merger as a defining sound change for PPh. See Appendix I for a detailed discussion of how the unmerged *d and *z in Central Luzon may be explained in the broader diachronic view of the linguistic prehistory of the Philippines.

3.2   Unmerged *d and *z in Minahasan languages

Minahasan languages provide further evidence against the view that *d and *z are undistinguished across Philippine languages. Sneddon (1978, 1989) was the first to point out unmerged reflexes of PMP *d and *z in Minahasan languages. Mead (ms.) further investigates Sneddon's findings and identifies strong evidence that PMP *d and *z were originally unmerged in Proto-Minahasan (PMin)—the reflex of PMP *z is PMin *d in all environments, showing 14 examples and no counterexamples. The reflexes of PMP *d, on the other hand, are split between PMin *d and *r, being realized as *d in 38 of 52 instances (or *r in 14 out of 52 instances). In the intervocalic environment this split is particularly evident: PMP *d is realized as *r in 7 of 10 instances, the three exceptions being PAN *sədu 'hiccup' > *səduʔ* 'hiccup', PAN *ludaq 'saliva' > *ludaʔ* 'spit', and PMP *tudaq 'throw' > *tudaʔ* 'stab'.

Based on this high proportion of *r reflexes, Mead posits a systematic change rather than borrowing or sporadic change: PMP *d > PMin *r intervocalically, and PMP *d > PMin *d everywhere else, as also suggested in Sneddon's early work.

PMP *d and *z therefore have distinct reflexes intervocalically in PMin but merge in other environments. As there are no examples of PMP *z > Proto-Minahasan *r, this merger of *d and *z must have occurred after the split of *d, so that *z and *d were still distinct phonemes at the time *d split.[9] With

---

[7]Though in other cases it is possible to determine whether a word is borrowed or native – when it contains a phoneme which that has undergone a unique sound change in a given language, e.g. PMP *R > Ayta Abellen <y>.

[8]There are 20 total instances of *z in the data. There are 131 instances of *d, 30 of which appear as /l/.

[9]It is interesting to note the parallels between the situations of P-Central Luzon and Proto-Minahasan with regard to *d and *z. In both cases, there is both a split in *d, significantly more prominent in the intervocalic environment, and a merger of one the reflexes of *d with *z. It is possible that both languages followed similar pathways in the chronological change of these phonemes.

PMin also presenting unmerged *d and *z, the evidence against a PPh *d/z merger becomes stronger. Two PPh subgroups do not align with the proposed sound change.

3.3 Unmerged *d and *z: Implications for linguistic prehistory

3.3.1 The conservative tendencies of languages on the periphery

Many of the languages discussed in this section are on the periphery of the Philippine languages. The Minahasan languages are spoken outside of the Philippine mainland, in Sulawesi. Ayta Abellen and Ayta Mag-Antsi are spoken by Negrito groups living in the highlands, so are both culturally and physically separated from the lowland Philippine majority (in the case of Negrito languages, the archaeological evidence suggests that contact between the Negritos and Austronesian speakers happened early on (Reid 1994), implying that the maintained contrast of *z and *d represented in these languages originates from an early stage of the settlement of the Philippines (closer to PMP), as suggested in Reid 1987). In being peripheral, both Minahasan and Negrito languages are less likely to be subject to diffusionary forces through the dialect chain, and so are more likely to retain features which diffusion has obliterated from other Philippine languages. If the *d/z merger spread through contact, we therefore have an explanation as to why *d and *z are merged throughout most of the Philippines, while remaining unmerged in only these particular languages. On the other hand, no other peripheral language we have investigated has so far shown any split in *d and *z, except for Chamorro, which Blust does not actually include in his proposed PPh (Blust 2005, p. 40).[10]

3.3.2 The distribution of *d/z merger in western Austronesia and its implication for the prehistory of the Philippines

What we can see from the observation of unmerged *d and *z in Central Luzon and Minahasan is that the merger may be better understood as a common areal feature that has occurred independently in the majority of MP branches spoken in the Philippines. This observation aligns well with existing discussions in the literature that the d/z merger is a well-documented phenomenon in western Austronesian (Blust 1999; Liao 2020) and is therefore not particularly convincing as an innovation for subgrouping purposes. The table below summarizes attested mergers in various western Austronesian languages. The wide geographical and subgroup representation suggests that this merger is not an innovation that can reliably define specific subgroups.

---

[10]Inati, as presented in Pennoyer 1987, possibly has unmerged *d/z, for example PAN *huzan > *oʤen*, but the data is not clear enough to come to any conclusions. Thanks to the anonymous reviewer for bringing this language to our attention.

(7) Sample list of western Austronesian languages manifesting *d/z merger outside the Philippines

	name of branch	affiliation	geographic region
1	Atayalic	Atayalic	Taiwan
2	Basay	East Formosan	Taiwan
3	Puyuma	Puyuma	Taiwan
4	Papora	Western Plains	Taiwan
5	Saaroa	Tsouic	Taiwan
6	Kelabit	North Sarawakan, Malayo-Polynesian	Borneo
7	Kiput	North Sarawakan, Malayo-Polynesian	Borneo
8	Toba Batak	Sumatran, Malayo-Polynesian	Sumatra
9	Rejang	Western Indonesian (?), Malayo-Polynesian	Sumatra
10	Balaesang	Celebic, Malayo-Polynesian	Sulawesi
11	Javanese	Western Indonesian, Malayo-Polynesian	Java
12	Tetun	CEMP, Malayo-Polynesian	Lesser Sunda
13	Rotinese	CEMP, Malayo-Polynesian	Lesser Sunda

As the table shows, the Formosan languages that have undergone this merger span five of the ten first-order subgroups of the Austronesian family. This further indicates that the *d/z merger is a widespread and independent development rather than a shared innovation for subgrouping purposes.

The Central Luzon subgroup appears increasingly phonologically distinct within the proposed Philippine family. It uniquely preserves both the *d/z* and *n/ñ* mergers, as seen in languages such as Kapampangan (e.g., PMP *buñi* > *buñi* celebrated, acclaimed) and perhaps Hatang Kayi (e.g., PMP *ñamuk* > *yamuk* mosquito). This raises the possibility of proposing a new Philippine subgroup that excludes Central Luzon languages, using the *d/z* and *n/ñ* mergers as defining innovations. However, this scenario is considered highly improbable due to the lack of supporting evidence. Central Luzon languages share the lexical innovations supporting a unitary Philippine family and are not otherwise deviant. Thus, these two phonological features are insufficient to define the Central Luzon subgroup in opposition to all other members of the proposed Philippine family.

Finally, the absence of the *d/z merger in Chamorro is noteworthy. There is a general consensus in the literature that the speakers of Chamorro originated in the Philippines. The main evidence, as discussed in Blust (2001), lies in the presence of a regular reflex, *pakyo*, of the Proto-Austronesian word for 'typhoon,' *baRius, which suggests that the speakers migrated to the Marianas from within the typhoon belt. This hypothesis aligns with the fact that Chamorro shares the *R/g merger with Greater Central Philippine (GCP) languages (Blust 1991).

Assuming that the Philippines was the homeland of Chamorro speakers and that Chamorro was part of the GCP subgroup (setting aside the fact that Chamorro is rarely classified as a Philippine language despite its potential affinity with Philippine languages), the absence of the *d/z merger in Chamorro suggests that the merger occurred after the split between Chamorro and Proto-Greater Central Philippine. This undermines the proposal that the *d/z merger was a defining shared innovation prior to the divergence of modern Philippine languages.

To conclude, the absence of the d/z merger in two of the major Philippine microgroups (Central Luzon and Minahasan) challenges its status as the sole defining phonological innovation for Proto-Philippines. Moreover, the fact that some, but not all, Central Luzon languages exhibit this merger suggests that it may represent a recurrent drift—or potentially the result of diffusion—common among Philippine languages. The key implication of this finding is that the proposed Proto-Philippine branch lacks a definitive phonological basis. Consequently, this branch remains defined solely by lexical innovations, the validity of which is explored in the following section.

4 Revisiting the lexical evidence for a Philippine branch

We turn now to a reconsideration the lexical evidence for PPh. A key part of this evidence is the list of 1,259 lexical items with no external cognates outside the Philippines (Blust 2019, 2020). It is claimed that (a) these lexemes are present in high quantity, so as to imply descent from a single protolanguage, and (b) that a core of them (37 items) represent the strongest type of lexical evidence for common descent, i.e. replacement innovations of PMP equivalents. Therefore, they must be descended from a distinct common ancestor, i.e. Proto-Philippines, which spread throughout the archipelago (and, as Blust proposed, replaced the descendant languages of Proto-Malayo-Polynesian which had been spoken there since the initial MP expansion). Much robust discussion has already taken place on the lexical evidence for PPh, with the proposal that these lexical innovations may be the outcome of lexical diffusion (see, for example, Smith 2017; Liao 2020; Reid 2020; Ross 2020; Zorc 2020). We summarize below the key arguments against the lexical data as evidence for PPh and expand on these with a few new findings and remarks.

Several authors (Smith 2017; Ross 2020) have already mentioned short and long-distance trading links as a vector for the diffusion of loanwords through the archipelago. Smith (2017: 464), in his review of the strength of PPh etyma, points to the word for Manila hemp (*Musa textilis*), an important resource for the production of fiber for rope and weaving, as being widely distributed (in the Batanes, Northern Luzon, Central Luzon and Greater Central Philippine microgroups) and as fitting Blust's criteria for a strong witness to PPh. However, Smith suggests that since Manila hemp is a widely traded item, this may be a loan. Blust (2019: 215) counterargues that since Manila hemp is native to the Philippines, PPh speakers presumably had a word for this item, and as such it cannot be a loanword. This issue is also highlighted in Liao (2020).

This example is representative of the debate over the lexical evidence for PPh as a whole. We would make two important points here: first, assuming that speakers of some ancestral Philippine protolanguage had a term for Manila hemp which was retained by its daughter languages, it is in principle not possible to distinguish whether this language was PPh or whether these lexemes arose in and spread through the Proto-Malayo-Polynesian dialect network. This aligns with the fact that the putative Proto-Philippine node is not supported by any discrete phonological innovations, as shown in section 3. In both the retention or diffusion scenarios the simplest explanation excludes the possibility of PPh.

4.1 Borrowing processes and the sociocultural history of the Philippines

A key issue is that of distinguishing loanwords from common inheritance, and some mention of this is made in each of the papers cited above. We argue that the distribution of these etyma does not constitute sufficient evidence for common descent, as it is often not possible to distinguish borrowings from etyma inherited from a common ancestor. In addition, it is highly likely that a dialect linkage formed during the initial MP expansion, and even widely distributed PPh etyma may in fact be items which spread through this dialect network early on.

This ties into the second aspect, which highlights that dismissing the likelihood of diffusion via trade so readily may not be a prudent approach. Reid (2018) argued for this possibility, but Blust (2020: 185) countered that lexical borrowing in trade should favor the types of items which are exchanged in commercial transactions, and that the majority of items in the 1,259 item list are not of this type. We believe this point is unpersuasive, as the term "trade" is not necessarily limited to commercial transactions. Long distance trade-type relationships which do not involve narrowly commercial exchange are attested for in many linguistic areas, such as the Kula circle in western Melanesia (Malinowski 1922). Long distance interaction of other types, such as intermarriage and multilingualism, is also an important factor

in loanword diffusion and may give rise to borrowings across wider domains (Haspelmath & Tadmor 2009).

Archaeological evidence and early Chinese documents have shown that long-distance interaction networks are of great antiquity in the Philippines, that chiefdoms of considerable political complexity were already well developed in 11th Century C.E., and that the initial period of complex society formation spanned the period from the late first millennium B.C.E. to the early first millennium C.E. (Junker 1998). Spanish accounts from the early colonial period (Loarca 1903; Junker 1998) also give us a glimpse into a world of great political complexity, with smaller and larger polities, speaking mutually unintelligible languages, engaging in constant warfare, alliance building, and tributary relationships. Additionally, maritime trade within the archipelago is well established in the archaeological record for the Philippines, both during and prior to the initiation of "foreign" trade late in the first millennium A.D. (Hutterer 1977; Junker 1990a, 1990b).

Given this complex sociopolitical history, we would argue that the presence of a large amount of loanword diffusion within the Philippine archipelago after the breakup of a higher-level dialect network is more likely than not, and that due to the difficulty alluded to above of distinguishing borrowings from common inheritance in Philippine lexicons, we should be circumspect about accepting arguments for PPh which turn on the quantity of lexical evidence.

Although we cannot completely discount the possibility of common inheritance as the source of these etyma, their semantic properties seem to indicate that they could just as easily be loanwords, and in the absence of diagnostic shared sound change we cannot discard either possibility. In our view, this renders the proposed lexical evidence for PPh problematic.

4.2   Geographical and subgroup distribution of PPh etyma

A cursory analysis of the PPh etyma in the ACD also shows that the argument from amount of cognate sets supposedly reconstructable to PPh cannot be valid. We conducted a basic search for geographically defined clusters of cognate sets supposedly reconstructed to PPh as follows: first, the full list of PPh protoforms and their reflexes was retrieved from https://acd.clld.org/; second, language location data for each language containing a reflex of each set was attached, sourced from Glottolog . Next, two geographical regions were defined (somewhat arbitrarily), northern Luzon and north Sulawesi.

Northern Luzon was defined geographically as a polygon extending northwards from a line beginning roughly at latitude 15'5" N. 90 of the cognate sets argued to be reconstructable to PPh are found only within this region. Furthermore, many of these cognate sets have reflexes in Ilocano and then in some languages with fewer speakers which have had a historical relationship with Ilocano. Even given that the ACD data is probably not exhaustive, it seems improbable that this many cognate sets of the total could have been retained only in this geographical area. We suggest that for these cases it cannot be ruled out that these cognate sets are instead the result of borrowing at the level of the PMP linkage or in the 3000 years since the divergence of the Philippine languages began, rather than common inheritance, especially in the absence of phonological evidence. For example,in cases such as *lánut, for which Blust's reconstructed PPh meaning is "abaca fiber, fiber-yielding plant, tree with a bark which yields a vine-like fiber", all of the attested reflexes of this form are present only within the Philippine archipelago, with none in the Greater Central Philippine (GCP) languages of northern Sulawesi such as Tondano and Gorontalo, whereas reflexes of this term do in fact appear in the GCP languages Central Subanon, Western Subanon, Aklanon, Mansakan, and Bisayan. In fact, a large number of the etyma proposed to be reconstructable to PPh share this pattern of attestation: they are present in the Philippines proper but not in the GCP languages of northern Sulawesi. Notably, only seven proposed PPh etyma are found in these languages: *usauR 'to go downstream', *láyug 'tall, of trees', *liqed 'footprints, tracks',

*iqit 'scraped or filed down', *habél 'weave cloth', *buál 'uproot a tree', and *butí 'smallpox; swine disease'.[11]

If it was indeed the case that these etyma were present in PPh, we are forced to assume that all reflexes of these cognate sets were lost or replaced in the GCP languages of northern Sulawesi. A more parsimonious explanation would be that these etyma represent loanwords, which spread via the Philippine maritime interaction sphere, but did not reach the GCP languages in northern Sulawesi further across the ocean. As for the seven PPh etyma which do appear in northern Sulawesi, these could represent either loans which spread throughout the early PMP dialect network, or retentions from PMP for which there are no longer any Philippine-external witnesses. We consider this to be counterevidence for the PPh hypothesis and evidence for an interaction sphere postdating the departure of the GCP North Sulawesi languages for their current location. We will discuss below an example of probable morphosyntactic borrowing which appears to show the same distribution, further reinforcing this interpretation.

The patterns of distribution of other words in the lexemes reconstructed for PPh also indicate that they are more likely to be the product of borrowing, rather than common inheritance. For example, several lexemes show a distributional pattern like that for the proposed PPh protoform *labas 'pass by, overlook something when searching; to pass by (of time)'. Consider the figure below, where black markers represent the reflexes of the protoform.
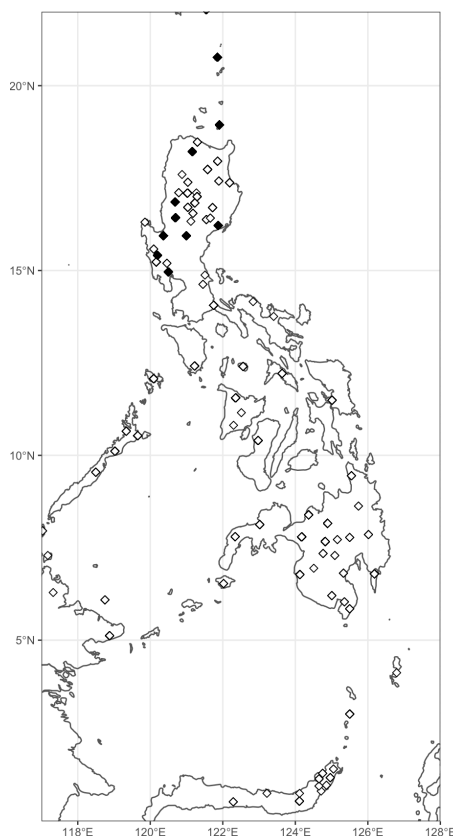
Figure 1: Reflexes of protoform *labas

In this type of distributional pattern, reflexes of a proposed PPh lexeme appear in a single large, regional language and several languages spoken by smaller groups. For *labas, the reflex *ag-pa-labás* appears in Ilocano (Northern Luzon microgroup), a major Philippine language of northwestern Luzon

[11] However, a closer examination of the Dutch dictionaries of Sulawesi may add to this list. We set this aside for future investigation.

	Reflex	Reflex Gloss	ACD Name	Speakers, approx.
1	labah	to pass by	Ayta Abellan	3500
2	lábis	excessive, more than enough (as a pole that is longer than it needs to be); excess	Casiguran Dumagat	610
3	debas-en	to make, take , etc. something too far (as house dimension beyond the specifications, bananas beyond proper ripeness)	Ibaloy	116000
4	h<om>abas	to pass by (object, day or time)	Ibatan	33000
5	ag-pa-lábas	to let pass; tolerate; be understanding	Ilocano	8100000
6	na-lábas	past	Isneg	40000
7	pa-xavas-en	to let pass	Itbayaten	3500
8	na-labás	gone; gone away; passed; passed on	Kankanaey	240000
9	labas	pass by, pass through (in the process of leaving, going out)	Kapampangan	2800000
10	on-labás	to go beyond, pass through; surplus, excess above requirements	Pangasinan	1800000
11	ni-mi-avas	passed by	Yami	4000

Table 1: Reflexes of *labas reported in the ACD

with around nine million speakers, and which has been an important regional language and lingua franca for many centuries (Rubino 1997). All of the other reflexes of *labas appear in smaller languages also spoken in northern and central Luzon such as Kankanaey, Ibaloi, Casiguran Dumagat (Northern Luzon microgroup) and several Negrito languages of central Luzon such as Kapampangan and Ayta Abellen (Central Luzon microgroup). The only other place where reflexes of this term appear is in the Batanes, in the Batanic languages Yami and Ibatan, of which Itbayat is known to have a historical relationship with Ilocano speakers (Maree 2005; Gallego 2020).

The geographical distribution, limited to northern and central Luzon, and the fact that Ilocano has historically been a prestige language in this region tentatively suggest that rather than being reconstructable to a deeper protolanguage like PMP or the purported PPh, this term originated in Ilocano and then spread as a loan among smaller languages which it was in contact with. The fact that Ilocano as a regional prestige language and a lingua franca would have been part of the language inventory of bi- or multi-lingual speakers of smaller groups in this region means that here we have a classic example of a situation with a sociological gradient through which loanwords must have spread from Ilocano into the smaller languages. Even though the meanings of the reflexes of this term do not necessarily fall into a "highly borrowable" category, we know that such situations can lead to lexical replacement in basic vocabulary and closed word classes (Thomason & Kaufman 1988; Haspelmath & Tadmor 2009).

To conclude, rather than pointing to a Proto-Philippines which replaced PMP daughter languages in the Philippines, the lexical evidence reveals an alternative scenario, one of a rich and complex history of contact and cultural change. The compilation of the 1,259 item list of Philippine-only etyma and their proposed reconstructions is a major resource for uncovering more about Philippine prehistory, and

further careful examination should reveal many more interesting details about the Philippine interaction sphere.

5 Horizontal transfer across the Philippine islands: Insights from morphosyntax

We have shown in sections 3 and 4 that the proposed phonological and lexical evidence for a purported Philippine branch does not withstand closer scrutiny. We turn now to an understudied case of morphosyntactic variation that further indicates extensive contact and borrowing across the Philippines. As will be shown below, this variation is widely attested in languages spoken on the major islands yet is absent in outlier languages within the same subgroups, reinforcing the view that the reported lexical items exclusive to Philippine languages likely result from contact.

5.1 Circumstantial voice basics

As is well-known, many Philippine languages inherit a complex voice system from Proto-Austronesian known as 'Philippine-type voice'. This system features four-way affixal morphology on the verb that controls argument-marking patterns and relativization restrictions (see, e.g., Himmelmann 2002 and Chen & McDonnell 2019). Consider below an example from Tagalog (6a–d), where the syntactic pivotal argument in each voice is italicized in the translation. To remain analysis-neutral, the case markers are glossed with abstract labels (CM_1; CM_2; P_1; P_2) as they do not affect the main evidence discussed here.

(8) Tagalog

    a. B⟨**um**⟩ili si   AJ ng   keyk mula kay   Lia para kay   Joy.
        buy⟨**AV**⟩ ‚PN.PIVOT AJ‚ ID.CM_2 cake P_1 PN.CM_2 Lia P_2 PN.CM_2 Joy
        '*AJ* bought cake from Lia for Joy.'   (ACTOR VOICE)

    b. Bi-bilih-**in**   ni   AJ ang   keyk mula kay   Lia para kay   Joy.
        CONT-buy-**PV** PN.CM_1 AJ ‚PIVOT cake‚ P_1 PN.CM_2 Lia P_2 PN.CM_2 Joy
        'AJ will buy *cake* from Lia for Joy.'   (PATIENT VOICE)

    c. Bi-bilih-**an**   ni   AJ ng   keyk si   Lia para kay   Joy.
        CONT-buy-**LV** PN.CM_1 AJ ID.CM2 cake ‚PN.PIVOT Lia‚ P_2 PN.CM2 Joy
        'AJ will buy cake from *Lia* for Joy.'   (LOCATIVE VOICE)

    d. **I**-bi-bili   ni   AJ ng   keyk mula kay   Lia si   Joy.
        cv-CONT-buy PN.CM1 AJ ID.CM2 cake P_1 PN.CM2 Lia ‚PN.PIVOT Joy‚
        'AJ will buy cake from Lia for *Joy*.'   (CIRCUMSTANTIAL VOICE)

    Not only is this four-way system found in Philippine languages, but it is also attested across higher-order Austronesian languages spoken in Taiwan, northern Borneo, and northern Sulawesi. Consider the following parallel examples of voice alternation in Paiwan (Formosan) (9) and Kimaragang (Dusunic, Malayo-Polynesian) (10).

(9) Paiwan

    a. Q<**m**>alup a   caucau tua vavuy i   gadu   tua vuluq.
        <**AV**>hunt ‚PIVOT man‚ CM_2 pig   LOC mountain CM_2 spear
        'The man hunts wild pigs in the mountains with a spear.'   (ACTOR VOICE)

    b. Qalup-**en** nua caucau a   vavuy i   gadu   tua vuluq.
        hunt-**PV** CM_1 man   ‚PIVOT pig‚   LOC mountain CM_2 spear
        'The man hunts wild pigs in the mountains with a spear.'   (PATIENT VOICE)

c. Qalup-**an** nua caucau tua vavuy a gadu tua vuluq.
hunt-**LV** CM₁ man CM₂ pig ⟦pivot mountain⟧ CM₂ spear

'The man hunts wild pigs in the mountains with a spear.' (LOCATIVE VOICE)

d. **si**-qalup nua caucau tua vavuy i gadu a vuluq.
CV-hunt CM₁ man CM₂ pig LOC mountain ⟦PIVOT spear⟧

'The man hunts wild pigs in the mountains with a spear.' (Ferrell 1979:202) (CIRCUM-STANTIAL VOICE)

(10) Kimaragang

a. **Mang**-alapak okuh do niyuw.
AV.TR-split ⟦1SG.PIVOT⟧ CM₂ coconut

'I will split the coconut(s).' (ACTOR VOICE)

b. Lapak-**on** kuh it niyuw.
split-**PV** 1SG.CM₁ ⟦PIVOT coconut⟧

'I will split the coconut(s).' (PATIENT VOICE)

c. Ong taak-**an** okuh dikau do siin [...]
if give-LV ⟦1SG.PIVOT⟧ 2SG.CM₁ CM₂ money

'If you give me money ....' (LOCATIVE/GOAL VOICE)

d. Nokuroh.tu n-**i**-lapak nuh do niyuw inoh dangol kuh?
why PST-**CV**-split 2SG.CM₁ CM₂ coconut ⟦MED.PIVOT⟧ bush.knife 1SG.POSS

'Why did you use my bush knife to split coconuts?' (Kroeger 2005: 13, 10; glosses ours)
(CIRCUMSTANTIAL VOICE)

As the data above show, each of the four voices in these languages exhibits a distinct yet unified argument-marking pattern. In sentences marked in Circumstantial Voice (CV), the instrument or bene-factive phrase carries a special marking labeled as pivot, which renders the sole phrase in the clause eligible for relativization ((8d), (9d), and (10d)). There is consensus in the literature that this four-way voice morphology can be traced back to Proto-Austronesian or at least to an early stage of Austronesian prehistory when speakers were still in the homeland, Taiwan (Ross 2009, 2012; Blust & Chen 2017). Consider the voice morphology reconstructed for Proto-Austronesian below, where each of the four voices was marked by a single affix, such as the morpheme *Si-/Sa- for Circumstantial Voice.

(11) Early Austronesian voice morphology (Blust 2009; Ross 2009, 2012; Blust & Chen 2017)

a. Actor Voice: *<um>

b. Patient Voice: *-in

c. Locative Voice: *-an

d. Circumstantial Voice: *Si-/Sa-

5.2 Innovative CV morphology in Philippine languages

The focus of the discussion here is an underexplored variation in CV morphology found in Philippine languages. In the majority of Austronesian primary branches, CV is free to combine with pivot phrases representing a wide range of thematic roles—instrument, theme, benefactive, reason, and cause. This is exemplified by examples from Seediq (12) and Paiwan (13). Each language represents a distinct Austronesian primary branch.

(12) Seediq (Formosan)

    a. [s]-hanguc=mu   ∅   sari ka   dakis.
       CV-cook=1SG.CM$_1$ CM$_2$ taro PIVOT Dakis
       'I cooked meat for Dakis.' (Chen 2017:101)   (benefactive pivot)

    b. ga=na   [s]-sebuc ∅   ricah ka   qreti.
       PROG=3SG.CM$_1$ CV-hit   CM$_2$ plum PIVOT stick
       'He/she is knocking down plums (from the trees) with a stick.' (Chen 2017:99) (instrument
       pivot)

    c. [s]-k<n>-narux   na temi ka   knrudan=na.
       CV-STAT<PFV>sick CM$_1$ Temi PIVOT age=3SG.POSS
       'Temi got sick because of her age.' (Chen 2017:79)   (reason pivot)

    d. Wada=mu   [s]-paadis ∅   dakis ka   tigami.
       PFV=1SG.CM$_1$ CV-send   CM$_2$ Dakis PIVOT letter
       'I sent Dakis a/the letter.' (Chen 2017:121)   (theme pivot in ditransitives)

    e. [s]-p-seeliq=mu   ∅   robo ka   rodux nii.
       CV-CAUS-butcher=1SG.CM$_1$ CM$_2$ Robo PIVOT chicken this
       'I asked Robo to butcher this chicken.'   (theme pivot in causatives)

(13) Paiwan (Formosan)

    a. 'u-[s<in>i]-pangul   sa a'-pungdjuq ta   kasiw ti   kapi.
       1SG.CM$_1$-CV<PFV>hit LK STAT-broken CM$_2$ wood PIVOT Kpi
       'I hit the wood broken for Kepi.' (Wu 2013:192)   (benefactive pivot)

    b. [si]-tekel ni   kapi ta   vava a   kupu.
       CV-drink PN.CM$_1$ Kapi CM$_2$ wine PIVOT cup
       'Kapi drinks wines with the cup.' (Wu 2013:31)   (instrument pivot)

    c. [s<in>i]-kan m   zepul ta   ci'aw a   za vengaLay nimadu.
       CV<PFV>eat PN.CM$_1$ Zepul CM$_2$ fish   PIVOT that pregmancy 3SG.CM$_1$
       'Zepul ate fish because of her pregnancy.' (Chang 2006:73)   (reason pivot)

    d. [si]-pavai ti   cemedas a   zua hana tjay zapul.
       CV-give PN.CM$_1$ Cemedas PIVOT that flower CM$_2$ Zepul
       'Cemedas gave that flower to Zepul.' (Chang 2006:297)   (theme pivot in ditransitives)

    e. 'u-[si]-pa-veli=anga   tjay kapi a   vatu.
       1SG.CM$_1$-CV-CAUS-sell=COS CM$_2$ Kapi PIVOT dog
       'I have caused Kapi to buy (i.e. sold Kapi) the dog.' (Wu 2013:33)   (theme pivot in
       causatives)

Unlike the Formosan pattern exemplified above, many Philippine languages exhibit a set of innovative morphemes that co-occurs with the CV affix, where the affix specifies the thematic role of the pivot phrase. In Tagalog, for example, the sequence *i-ka-* indicates reason phrases, *i-pag-* indicates instrument, and *i-pang-* indicates benefactive.[12] Regardless, there are also verbs that do not take the innovative affixes *ka-*, *pag-*, and *pang-*.

---

[12]Notably, the reason-denoting sequence *i-ka-* discussed here is formally identical to the sequence *i-ka-* in some Formosan languages such as Saisiyat and Puyuma, but is functionally distinguishablethe latter is used for all Circumstantial Voice constructions that contain a stative verb stem and do not necessarily require a reason pivot.

(14) Tagalog (Malayo-Polynesian)

    a. $\boxed{\text{i-p<in>ag}}$-luto   ni   Kyla ng   adobo si   Juan.

       CV-PAG<PFV>-cook CM$_1$ Kyla INDF.CM$_2$ adobo PIVOT Juan

       'Kyla cooked adobo for Juan.'   (CV + *pag-* for benefactive pivot)

    b. $\boxed{\text{i-p<in>ang}}$-ka-kain   ni   Kyla ang   kutsara.

       CV-PANG<PFV>-RED-eat PN.CM$_1$ Kyla PIVOT spoon

       'Kyla is eating with the spoon.'   (CV + *pang-* for instrument pivot)

    c. $\boxed{\text{i-k<in>a}}$-matay ni   Juan ang   kanser.

       CV-KA<PFV>-die CM$_1$ Juan PIVOT cancer

       'Juan died of cancer.'   (CV + *ka-* for reason pivot)

On the other hand, other types of pivot possible in CV, such as the theme phrase in three-place constructions marked in CV, do not appear with the additional affix *pang-*, *pag-*, or *ka-*. Consider (15).[13]

(15) Tagalog (Malayo-Polynesian)

    a. $\boxed{\text{i}}$-b<in>igay ni   Juan ang   pera kay   Aya.

       CV-<PFV>give PN.CM$_1$ Juan PIVOT money PN.CM$_2$ Aya

       'Juan gave Aya the money.'   (CV for theme pivot in ditransitives)

    b. $\boxed{\text{i}}$-p<in>a-kanta=ko   kay   Ivan ang   kanta.

       CV-CAUS<PFV>sing=PN.CM$_1$ PN.CM$_2$ Ivan PIVOT song

       'I asked Ivan to sing a song.'   (CV for theme pivot in causatives)

This data reveals four subtypes of Circumstantial Voice verbal morphology that are unattested in Formosan languages, summarized in (16).

(16) Four subtypes of CV morphology found in some Philippine languages

    a. CV

    b. CV + *pag-*

    c. CV + *pang-*

    d. CV + *ka-*[14]

For example, Itbayaten, a Bashiic language that does not bear a particularly close relationship with Tagalog, employs the same three-way thematic-oriented subdivision in CV morphology. Consider (17).

(17) Itbayaten (Malayo-Polynesian)

    a. $\boxed{\text{i-cha}}$-hakey

       CV-KA-like

       'A is the cause of liking.' (Yamada 2015:50)   (CV + *ka-* for reason pivot)

---

[13]An anonymous reviewer agreed with this generalization but pointed out several exceptions where *pag-* is attached to non-benefactive pivots: *i-pag-kait* 'withhold theme,' *i-pag-tanggol* 'defend theme,' and *i-pag-daos* 'celebrate theme.' This indicates that the sequence *i-pag-* is not fully productive in Tagalog, and its use is subject to native intuition. Our consultation with native speakers did not reveal any conditions that predict the presence or absence of *pag-*.

[14]An anonymous reviewer suggests that Formosan languages also use the CV morphological sequence with the prefix *ka-*, implying that the pattern in (16d) is neither innovative nor unique to Philippine languages. However, this similarity is only apparent. In Formosan languages, the sequence CV-*ka-* is obligatorily used for the CV form of stative stems; that is, the function of *ka-* is stative rather than marking a reason pivot, and is therefore distinct from the innovative function discussed here. Readers are encouraged to consult the reference grammars of Formosan languages for further details.

b. $\boxed{\text{i-pang}}$ -among
CV-PAN-fish
'to fish with A' (Yamada 2015:50) (CV + *paN-* for instrument pivot)

c. $\boxed{\text{i-pang}}$ -xap
CV-PAN-get
'to get for A' (Yamada 2015:50) (CV + *paN-* for benefactive pivot)

Importantly, like in Tagalog, not all CV-marked verbs in Itbayaten employ the innovative affix. For example, although some instrument-selecting CV verbs take the affix *pang-* (17b), there are similar verbs that do not require that affix. Consider two reported examples without the innovative affix (18) and a similar locus of variation in Yami.

(18) Itbayaten (Bashiic, Malayo-Polynesian)
a. $\boxed{\text{i}}$ -chali
CV-dig
'to dig with X' (Yamada 2015:50) (CV for instrument pivot)

b. $\boxed{\text{i}}$ -'inom
CV-drink
'to drink with X' (Yamada 2015:50) (CV for instrument pivot)

Our survey of 40 Philippine languages reveals a similar use of the additional, innovative morphology in languages across different regions of the Philippines and 5 of the 15 subgroups defined in Blust (2001): Bashiic, Central Luzon, Greater Central Philippines, Kalamian, and Cordilleran. The distribution of the languages found with this innovation is illustrated in Figure 2 and summarized in (19). All language names displayed on the map represent languages that exhibit the specific innovation.[15]
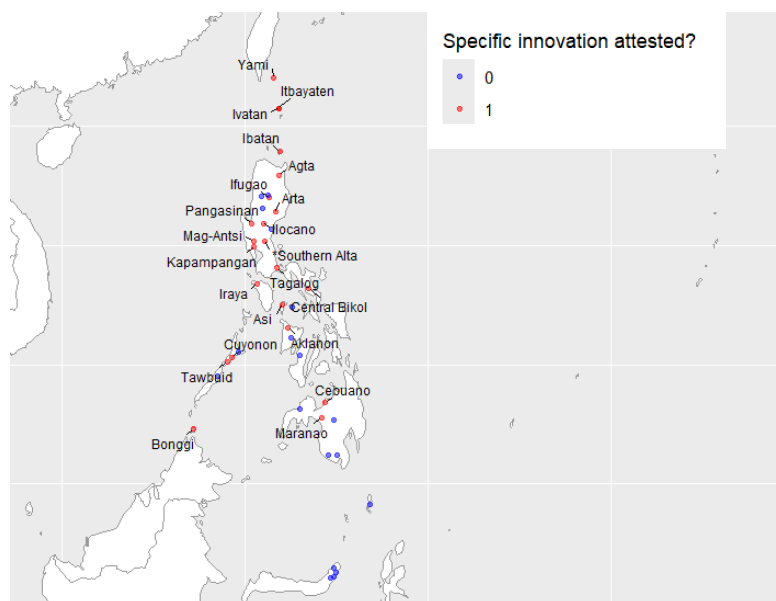
Figure 2: Distribution of languages with and without the innovative CV morphology

---

[15]Asterisks in the tablesuch as 0* and 1*stand for cases attested with other types of innovation beyond the innovative CV morphology under discussion. See also footnote 14 for an explanation of languages that lack an overt Circumstantial Voice affix *i-* for CV clauses (and why we consider affixes such as *pag-* as a valid case of the specific innovation under discussion). We remain agnostic about the subgrouping affiliation of Iraya (a language spoken by Mangyans on the island of Mindoro) and tentatively classify it as part of Central Luzon.

(19) List of Philippine languages surveyed in the paper

#	Microgroup	language	Innovation attested?	voice morphology	pivot type	source
1	Bashiic	Yami	0	i-	instrument	Rau & Dong (2006)
		Yami	1	i-ka-	reason, manner	Rau & Dong (2006)
2		Itbayaten	0	i-	instrument	Yamada (2014)
		Itbayaten	1	i-cha-	reason, cause	Yamada (2014)
		Itbayaten	1	i-paN-	benefactive, instrument	Yamada (2014)
3		Ivatan	1	qi-pang-	instrument	Reid (1966)
4		Ibatan	0	i- ... -an	benefactive	Maree (2007)
		Ibatan	1*	pang-, pay-	instrument	Maree (2007)
5	Bilic	Tboli	0	zero	instrument	Forsberg (1992)
6		Blaan	0	zero	instrument	Bondoc (2015)
7	Central Luzon	Kapampangan	0	i-, pan-, pag-	benefactive	Forman (1971)
		Kapampangan	1*	pi-, pipag-, pipan-	instrument, location	Forman (1971)
		Kapampangan	1*	pag-, pan-, i-	instrument	Forman (1971)
8		Iraya	1*	pinaN-	instrument	Or (2018)
		Iraya	0	-an	benefactive	Or (2018)
9		Mag-Antsi	1*	paN-	benefactive, instrument	Javier et al. (2018)
		Mag-Antsi	1	ika-	reason	Javier et al. (2018)
10	GCP	Maranao	1	i-, i-pang-	instrument	McKaughan & Macaraya 1967
		Maranao	1	-an, pang-, p-i-ang-	referential	McKaughan 1967
11		Western Bukidnon Manobo	0	i-	instrument, benefactive	Elkins (1967)
12		Central Bikol	1*	i-, pang-	instrument	Mintz (1971)
		Central Bikol	1	i-, i-pig-, i-ka-	benefactive	Mintz (1971)
13		Cebuano	1	i-, gi, i-pag-, gika-	instrument	Tanangkingsing (2009), Bunye & Yap (2019)
		Cebuano	1	i-, -an, pag-...i-, pago-..-an, pang- ... -an	benefactive	Tanangkingsing (2009)
14		Hiligaynon	0	-an	benefactive	Wolfenden (1971)
		Hiligaynon	0	i-	instrument	Wolfenden (1971)
15		Romblomanon	0	ging-, in-, gina-, i-	instrument	Law (1997)
		Romblomanon	0	ging-...-an, ging-...-i	benefactive	Law (1997)
16		Cuyonon	1*	ing-pang	instrument	Castro (2023)
		Cuyonon	0	ing-...-an	benefactive	Castro (2023)
17		Asi	1	i-, ipang-	instrument	Hendrickson & Kilgour (1985-89)
		Asi	1	i-ka(ka)	cause/reason	Hendrickson & Kilgour (1985-89)
		Asi	0	a-...-an, i-	benefactive	Hendrickson & Kilgour (1985-89)
18		Tagalog	0	i-	theme	primary data
		Tagalog	1	i-pang-	instrument	primary data
		Tagalog	1	i-pag-	beneficiary	primary data
		Tagalog	1	i-ka-	reason	primary data
19		Aklanon	1*	-in-...-an, pa-...-an, ipang-	benefactive	de la Cruz & Zorc (1968)
		Aklanon	0	-in-, i-, iga-	instrument	de la Cruz & Zorc (1968)
20		Aborlan Tagbanwa	0	i-	instrument	Hussey (1965)
21		Tawbuid	1*	faN-	instrument	Fleming (2022)
		Tawbuid	0*	na-...-an, fag-...-an	reason	Fleming (2022)
22		Bonggi	0*	-an	benafactive	Boutin (2002)
		Bonggi	1*	peng-, -in-peng-	instrument	Boutin (2002)
23		Subanen	0	pig-, pog-...-on	instrument	Estioca (2020)
		Subanen	0	pig-...-an, pog-...-an	benafactive	Estioca (2020)
24	Inati	Inati	0	i-, gin-, -an	instrument, benefactor?	Manzano (2021)
25	Kalamian	Agutaynen	0	-an	beneficiary	Quankenbush et al. (2010)
		Agutaynen	0*	pag-...-an	cause	Quankenbush et al. (2010)
		Agutaynen	0	i-	instrument	Quankenbush et al. (2010)
26	Cordilleran	Nothern Alta	0	i-	instrument, causative theme	Laguia (2018)
27		Southern Alta	1*	pang-, i-	instrument	Abreu (2018)
		Southern Alta	0	-en	benefactive	Abreu (2018)
28		Ifugao	0	iN-...-an	benefactor	Hohulin (2011)
		Ifugao	1*	i-, impuN-, iN-, puN-	instrument	Hohulin (2011)
29		Balangao	0	i-	instrument	Shetler (1976)
		Balangao	0	i-..-an	benefactor	Shetler (1976)
30		BontokFinallig	0	i-	benefactor	Fukuda (1997)
		BontokFinallig	0	i-...-an	instrument	Fukuda (1997)
31		Pangasinan	0	i-	instrument	Benton (1971)
		Pangasinan	0*	impama-	reason, cause	Benton (1971)
		Pangasinan	1	pan-, inpan-, inpan-, pangi-, inpangi-	instrument	Benton (1971)
		Pangasinan	0	i-...-an, in-...-an	benefactive	Benton (2018)
32		Agta	1*	i-, pang-	instrument	Robinson 2008
		Agta	0	i-...-an	benefactive	Robinson 2008
33		Karao	0*	i-...-an	benefactor, recipient	Brainard (1997)
		Karao	0	i-	instrument	Brainard (1997)
34		Ilocano	1*	para-, paN-, pag-, pangi-	instrument	Rubino (1997)
		Ilocano	0	i-...-an	benefactive	Rubino (1997)
35		Arta	1*	paN-, i-,	instrument	Kimoto (2017)
36	Minahasan	Tonsawang	0	-i	instrument, beneficiary	Hayes (2019)
37		Tondano	0	-i	instrument, beneficiary	Snedon (1975)
38	Sangiric	Bantik	0	zero	ditr theme, benefactor, instrument	Utsumi (2022)
39		Ratahan	0	zero	theme, instrument, benefactive	Himmelmann & Wolff (1998)
40		Talaud	0	i-	ditr theme, benefactor, instrument	Utsumi (2013)

As shown in (19), this innovative morphology is found in various Philippine languages. Consider below examples below from Asi (Greater Central Philippines) and Pangasinan (Meso-Cordilleran). In both languages, the same innovative morphemes are employed for CV sentences with different types of pivot phrase.

(20) Asi (Greater Central Philippines, Malayo-Polynesian)

    a. i-pang -limpyo nako   kag suka   it mga gaha.
        F-CV-PANG-clean $1\textsc{sg.cm}_1$ $\textsc{cm}_2$ vinegar G PL   window

        'I'll clean the windows with vinegar.' (Hendrickson & Kilgour 1985:39)   (CV + *pang-* for instrument pivot)

    b. i -sandrok nako   sida   it suya.
        CV-dish.food $1\textsc{sg.cm}_1$ $3\textsc{sg.pivot}$ G viand

        'I'll dish up some of this viand for her.'   (CV for benefactive pivot)
        (Hendrickson & Kilgour 1985:40)

    c. Kag i-k<in>a -matay nida   ay   kanser.
        N   CV-KA<PFV>-die $3\textsc{sg.cm}_1$ PIVOT cancer

        'What he died of was cancer.' (Hendrickson & Kilgour 1985:40)   (CV + *ka-* for reason pivot)

(21) Pangasinan (Meso-Cordilleran, Malayo-Polynesian)

     i-mpan -katli   nen Mark iyan   katli   ed   samay papel.
     CV-PAN.PFV-cut $\textsc{cm}_1$ Mark DEM.PIVOT scissors $\textsc{cm}_2$ cm_2 paper

     'Mark used this scissors to cut that paper.' (Rosario 2017:81) (CV + PAN- for instrument pivot)

Crucially, although this innovative pattern is spread across multiple Philippine microgroups, it is not attested in all members of these groups. While the languages discussed above belong to four of the nine first-order branches under the putative Philippine subgroup, many other members of these branches do not share this innovation. This is exemplified by the data below from Yami (22) and Karao (23). Both languages either employ a single (retentive) affix *i-* for CV clauses—as do Formosan languages (12)–(13)—or employ other affixes to denote specific types of CV constructions (e.g. *i-...-an* for clauses with a benefactive pivot, as in (23b)).

(22) Yami (Bashiic, Malayo-Polynesian)

    a. i -tangtang
        CV-smash

        'to smash with X'   (CV for instrument pivot)

    b. i -vaod
        CV-tie

        'to tie up with X' (Rau & Tong 2005: 540, 547, 554)   (CV for instrument pivot)

(23) Karao (Meso-Cordilleran, Malayo-Polynesian)

    a. 'i -tegteg na   nga'ngi-'i   batho-cha 'aramdi.
        CV-flatten $\textsc{cm}_1$ child-PIVOT rock-$\textsc{cm}_2$ wire

        'The child will use the rock to flatten the wire.' (Brainard 1997:101)   (CV for instrument pivot)

    b. 'i -tongkal- an na to'o-'i   nga'nga na   'amayo.
        CV-buy-AN $\textsc{cm}_1$ person-PIVOT child $\textsc{cm}_2$ toy

        'The person will buy the child the toy.' (Brainard 1997:100)(CV + *-an* for benefactive pivot)

A closer look at the distribution of the 22 Philippine languages (out of the 40 examined) that exhibit these innovative affixes reveals an important facts: while this innovation is found across the three major islands (Luzon, Visayas, and Mindanao), including four Negrito languages that are under extensive contact with either Ilocano (Arta, Agta, Mag-antsi) or Tagalog (Southern Alta), it is attested neither in Palawan nor in members of the Greater Central Philippine group spoken in northern Sulawesi and the islands between Sulawesi and Mindanao. Consider the examples below from Talaud (24a–c).

(24) Talaud (GCP, Malayo-Polynesian)

a. lama'a $\boxed{\text{i}}$-saraing ngimangitou.
   dish   CV-dance   3PL.CM$_1$
   'Dishes will be used by them in a dance.' (Utsumi 2022: 20)   (CV for instrument pivot)

b. inassa n-$\boxed{\text{i}}$-laha=ku   huna m-maria.
   fish   PST-CV-cook=1SG.CM$_1$ for   CM$_1$-Maria
   'I cooked fish for Maria.' (Utsumi 2022: 24)   (CV for benefactive)

c. ana? itou   n-$\boxed{\text{i}}$-luass   i-piteres [uauggu na-?angkat=te   huru   ].
   child 3SG.CM$_1$ PST-CV-be.pleased CM$_1$-Peter [because AV.PST-promote=COMP teacher]
   'Her/his child made her/him pleased because (her/his child) became a teacher.' (Utsumi 2022: 25)   (CV for reason pivot)

This distribution yields important questions concerning the prehistory of Philippine languages: Where is the center of dispersal of the innovative CV morphology, and how does its distribution inform the validity of a Philippine branch? We suggest that pinpointing the precise locus of this innovation is challenging; however, the core value of this case study lies not in its origin but in its distribution pattern—which mirrors the phonological and lexical innovations discussed above. Specifically, the innovation appears across multiple Philippine microgroups, yet only in a subset of languages. Moreover, it is absent from languages spoken on remote islands as well as in the interiors of the islands.

We argue that this empirical picture is best interpreted as the outcome of extensive contact, resulting in morphosyntactic borrowing across the three main islands of the Philippines while sparing offshore and interior languages. Such a pattern supports a borrowing scenario rather than a retention hypothesis. In other words, it does not imply that the innovation emerged at Proto-Philippine and was subsequently lost in some languages—an outcome unlikely given both its complete absence in geographically isolated areas and the fact that various Philippine languages from distinct subgroups employ the same plain pattern of CV morphology, as do languages outside the Philippines, such as Formosan languages.

This empirical picture is difficult to explain unless one assumes that the innovation of CV morphology emerged and was lost independently in all Philippine microgroups. Nor is the innovation compatible with a drift analysis, given that various languages exhibiting the innovation employ the same set of innovative morphological markers for the division of thematic roles. This distribution is instrumental in establishing the histories of Philippine languages. Tracing the source and dispersal patterns of this innovation is beyond the scope of the current paper and awaits further research. Ultimately, the evidence from morphosyntax underscores the presence of extensive borrowing and contact at the morphosyntactic level.

5.3   Implications: horizontal transfer at the morphosyntactic level

As discussed above, the observed distribution of the innovative CV morphology suggests that these patterns are unlikely to result from inheritance from the tree top (i.e. retention from the putative shared common ancestor), as that scenario would require the innovation to be lost in (i) all languages spoken outside the three major islands and (ii) the majority of languages under the branches attested with this

change. Instead, it reinforces the presence of horizontal transfer across the main islands of the Philippines. This innovative CV morphology fits well with Type III structural change (25) defined in Heine and Kuteva (2005: 124).

(25) *Type III Structural effect of contact-induced grammaticalization*

The new and the old categories coexist side by side, but the structure of the old category is redefined as a result of the presence of the new category (differentiation).

A related question to this distribution, as both anonymous reviewers queried, is the source language in which this innovation emerged and the directionality of the hypothetical borrowing. An anonymous reviewer also questioned whether the current distribution of these three affixes may be the outcome of extensive loss in extra-Philippine languages.

Given the consensus in the literature that structural borrowing presumes extensive lexical borrowing from the same language (e.g., Thomason & Kaufman 2001; Heine & Kuteva 2005; Aikhenvald 2006; Matras 2009; a.o.), the presence of structural borrowing in the innovative CV morphology strongly indicates that extensive lexical borrowing also took place across the main islands of the Philippines. This understudied locus of morphosyntactic variation thus lends further empirical support to the lexical diffusion account for the shared lexical items.

6 Conclusion

This paper has reexamined recent arguments for Proto-Philippines as the sole surviving PMP descendant on Luzon (Blust 2019, 2020, 2022; Zorc 1986, 2020) and presented new evidence for an alternative view. We argue that there is little motivation to postulate this alleged ancestor, as the claimed evidence for PPh can be explained by various layers of diffusion involving extensive contact, as suggested by Ross (2020). Three lines of new evidence support this view. First, the absence of the PMP *d/z merger in the Central Luzon subgroup undermines the merger as an innovation defining PPh. Second, the geographical distribution of these innovations, along with the semantic categories of reported lexical innovations as defined by Haspelmath & Tadmor's (2009) criteria, indicates a high likelihood of borrowing rather than inheritance. Third, the distribution of an underexplored variation of Circumstantial Voice (CV) morphology suggests multiple layers of borrowing across Philippine subgroups, lending new empirical evidence for extensive contact on the main islands of the Philippines. We therefore conclude that there is no obvious motivation to assume the existence of PPh.

7 Appendix

Himes (2012) proposes an account of the Central Luzon languages using regular sound change rules. One of the proposed rules is *d > l / V_V in Kapampangan and the Sambalic languages. As well as this, he proposes a number of other rules to explain the split realization of *d, but these are lacking in that they are highly targeted rules which often lack phonetic motivation, and that they leave much of the data still unexplained. For example, Himes suggests that in the Sambalic languages, *d > /l/ word initially, except when the next consonant is a liquid. However, there are many stable forms, stretching back to PWMP and earlier, which violate this rule. E.g. PWMP *dapít > AA *dapít 'part, place, direction', PAN *depah > AA depah 'armsbreadth', or PAN *daqaN > BT daʔan 'old (of objects)'. Additionally, there is no clear phonetic motivation for *d to change to /l/ in word-initial but not word-final position. We suggest that the sound change *d > l / V_V is sufficient to explain not only the vowel-medial instances of *d > l, but also the word-initial and word-final instances, once the possibility of back-formation from affixed forms or forms in connected speech is taken into account.

First we examine the straightforward instances of *d > l / V_V. There were 17 instances of this sound change found in the data, across all four languages, and seven exceptions. The exceptions are PWMP *ledek > AA ledek 'to pound grain by mortar and pestle'; PPh *pidek (with metathesis)$_2$ > BT/AM kirip 'eyelash' (a metathesis); PAN *kuden > KP kúran 'large cooking pot for rice', PPh *dúdun > KP durún 'locust', and PPh *katúday > KP kature 'a plant: Sesbania grandiflora'; and PMP *pudul > BT poról 'blunt, dull', plus one exception found in Himes (2012) but not the ACD data: PMP *ludaq > BT ludáʔ 'spit'. For six of these exceptions, we have good reason to suspect that the *d > l / V_V sound change did not apply to them. ledek is likely a new coinage from the root *-dek$_2$. kirip, durún and kature are all words unique to the Philippines and may be considered as loanwords or new coinages. For ludáʔ and poról, we should expect them to resist the change due to the OCP effect – if *d were to change to /l/ in these forms, it would result in identical liquids appearing in succession. In other instances, the change may go ahead in such a context: see PWMP *ludem > KP ma-lúlam 'cloudy, about to rain'. If the given explanations hold, only the Kapampangan reflex kúran from PAN *kuden remains as an unexplained exception to the rule *d > l / V_V, leaving it standing on solid ground.

In the Sambalic languages (but not Kapampangan), *d can also appear as /l/ in word-initial or -final position, though it shows up less frequently than in the vowel-medial position. In Ayta Abellen, for example, PAN *daRaq > daya 'blood' but PAN *daNum > lanom 'water'. We hypothesize that such a split realization of *d could be the result of a secondary effect of the rule *d > l / V_V, that is, the result of back-formation from affixed forms, or even from forms in connected speech. In the case where *d is in word-initial or -final position, the presence of prefix or suffix respectively would place *d in a vowel-medial position, making it subject to the rule *d > l / V_V. This would result in two competing forms of the same stem being present in the language, one (an l-stem) contained within the affixed form, the other (a d-stem) within the unaffixed stem. For example, if Proto-Central Luzon *dateng 'to come' were affixed to become *ka-dateng-an, the affixed form would change to *kalatengan under the rule *d > l / V_V, and later that form could be reinterpreted as consisting of a stem *lateng affixed with *ka-...-an. If this process occurred in Proto-Sambalic, it would mean that in some cases, the innovative l-stem would replace the original, and in other cases, the original d-stem would remain, and be re-affixed to push out the innovative affixed form. A similar scenario could play out if a word without a final consonant precedes a word beginning with *d in connected speech, or a word with a final *d is followed by a word without an initial consonant. Phonetically, *d would be an intervocalic position, and thus be realized as /l/. Again, the result would be two competing forms in the language, one with a d-stem and one with an l-stem. Such a scenario would lead to the inconsistent realisation of *d in word-initial and -final position that we see in the Sambalic languages. After this process, the remaining d-stem reflexes would merge

with *z, splitting with the l-stems in their realization. An anonymous reviewer informs us that a similar process is currently active in Tagalog, where /d/ regularly becomes /r/ between vowels (Blust 2013:623): "In roots like *damdam* 'feel', we now find a variant *ramdam* which arose from frequent prefixation in its verbal form *ma-damdam-an* 'to feel'.

8   References

Abreu, M. (2019). *A reference grammar of Southern Alta (Kabuloan Dumagat)*. DLSU, BAGCED-DEAL. Manila: De La Salle University.

Aikhenvald, A. (2006). Reflections on language contact, areal diffusion, and mechanisms of linguistic change. *Sprachbund in the West African Sahel*, 23–36.

Benton, R. (2019). *Pangasinan reference grammar*. University of Hawaii Press.

Blake, F. (1906). Contributions to Philippine grammar. *Journal of the American Oriental Society* 27, 317—396.

Blust, R. (1977). The Proto-Austronesian pronouns and Austronesian subgrouping: a preliminary report. *Working Papers in Linguistics* 9(2):115. Honolulu: Department of Linguistics, University of Hawaii.

Blust, R. (1991). The Greater Central Philippine hypothesis. *Oceanic Linguistics* 30(2), 73–129.

Blust, R. (1993). *s metathesis and the Formosan/Malayo-Polynesian language boundary. In D. Øyvind (ed.). *Language, a Doorway Between Human Cultures: Tributes to Dr. Otto Chr. Dahl on His Ninetieth Birthday*, 178–183. Oslo: Novus.

Blust, R. (1995). The position of the Formosan languages: method and theory in Austronesian comparative linguistics. In P. Li et al., (eds.), *Austronesian studies relating to Taiwan*, 585–650. Symposium Series of the Institute of History and Philology, Academia Sinica, No. 3. Taipei: Academia Sinica.

Blust, R. (1999). Subgrouping, circularity and extinction: some issues in Austronesian comparative linguistics. In E. Zeitoun and P. Li (eds.), *Selected Papers from the Eighth International Conference on Austronesian Linguistics*, 31–94. Symposium Series of the Institute of Linguistics (Preparatory Office), Academia Sinica, No. 1. Taipei: Academia Sinica.

Blust, R. (2001). Malayo-Polynesian: new stones in the wall. *Oceanic Linguistics* 40:151–55.

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

Blust, R. (2013). *The Austronesian Languages* (Revised). Australian National University.

Blust, R. (2019). The resurrection of Proto-Philippines. *Oceanic Linguistics* 58(2), 153–256.

Blust, R. (2020). Response to comments on "The resurrection of Proto-Philippines." *Oceanic Linguistics* 59(1–2), 450–479.

Blust, R. (2022) Proto-Philippine Addenda: Mind, Theory, and Data. *Oceanic Linguistics* 61(1): 322–404.

Blust, R. and V. Chen. (2017). The pitfalls of negative evidence: 'Ergative Austronesian', 'Nuclear Austronesian' and their progeny. *Language and Linguistics* 18(4):577–621.

Bondoc, I. (2015). A grammar sketch of Blaan. MA Thesis: University of the Philippines Diliman.

Brainard, S. (1997). Ergativity and grammatical relations in Karao. In T. Givón (ed.), *Grammatical relations: A functionalist perspective*, 85–154. Amsterdam/Philadelphia: John Benjamins.

Boutin, M. E. (2002). Nominative and genitive case alternations in Bonggi. In F. Wouk and M. Ross (eds.), *The history and typology of western Austronesian voice systems*, 209–239. Canberra: Pacific Linguistics.

Castro, J., C. Comeo, R. Bautista, and C. Mamucud. (2023). The Verbal Morphology of Cuyonon: Voice, Aspect, and Modality. Paper presented at The 32nd Meeting of the Southeast Asian Linguistics Society (SEALS32).

Charles, M. (1974). Problems in the reconstruction of Proto-Philippine phonology and the subgrouping of the Philippine languages. *Oceanic Linguistics 13*, 457–509

Chen, V. and B. McDonnell. (2019). Western Austronesian voice. *Annual Review of Linguistics* (5), 173–195.

de la Cruz, B.A. and R.D. Zorc. (1968). *A study of the Aklanon dialect*. Volume One: Grammar. Peace Corps.

Ebarhard, D. M., G. F. Simons, & C. D. Fennig. (2021). Ethnologue: Languages of the world.

Elkins, R. (1967). *Major grammatical patterns of western Bukidnon Manobo*. Norman: Summer Institute of Linguistics of the University of Oklahoma.

Estioca, S. (2020). A grammar of Western Subanon. Ph.D. dissertation, University of Hawai'i at Mānoa.

Fleming, H. (2022). *A Grammar of Eastern Tawbuid*. SIL International.

Forman, M. (1971). *Kapampangan grammar notes*. University of Hawaii Press.

Forsberg, V. 1992. A pedagogical grammar of Tboli. *Studies in Philippine Linguistics* 9: 1–110.

François, A. (2014). Trees, waves and linkages: Models of language diversification. In C. Bowern & B. Evans (eds.), *The Routledge Handbook of Historical Linguistics*: 161–189. Routledge.

Fukuda, T. (1997). A discourse-oriented grammar of Eastern Bontoc. *Studies in Philippine Linguistics* 10(1): 1–116.

Gallego, M. K. S. (2020). Ibatan of Babuyan Claro (Philippines) - Language contexts. *Language Documentation and Description*, 17, 87-110.

García Laguía, A. (2018). Documentation of Northern Alta: grammar, texts and glossary. PhD Thesis, Universitat de Barcelona.

Gray, R. D., Drummond, A. J., & Greenhill, S. J. (2009). Language phylogenies reveal expansion pulses and pauses in Pacific settlement. *Science* 323(5913), 479–482.

Greenhill, S. J. and Gray, R. D. (2009). Austronesian language phylogenies: Myths and misconceptions about Bayesian computational methods. In L. Sagart et al. (eds.), *Austronesian historical linguistics and culture history: A festschrift for Robert Blust*: 1–23). Australian National University.

Hammarström, H., R. Forkel, Robert, M. Haspelmath, & B. Sebastian. (2024). Glottolog 5.1. Leipzig: Max Planck Institute for Evolutionary Anthropology. https://doi.org/10.5281/zenodo.14006617

Haspelmath, M. and U. Tadmor. (2009). The Loanword Typology project and the World Loanword Database. In M. Haspelmath & U. Tadmor (eds.), *Loanwords in the world's language: A comparative handbook*: 1–34). De Gruyter.

Heine, B. and T. Kuteva. (2005). *Language contact and grammatical change*. Cambridge: Cambridge University Press.

Hendrickson, G. and Kilgour, H. (1985). Bantoanon (Asi) grammar essentials. Retrieved from http://tagmi.net/en/grammar

Himes, R. (2012). The Central Luzon Group of Languages. *Oceanic Linguistics*: 51(2), 490–537.

Himmelmann, N. (2002). Voice in Western Austronesian: An update. In F. Wouk & M. Ross (eds.), *The history and typology of Western Austronesian voice systems*, 715. Canberra: Pacific Linguistics.

Himmelmann, N. and J. Wolff. (1999). *Toratán (Ratahan).* Lincom Europa.

Hussey, S. C. (1965). Aborlan Tagbanwa: Verbal system and related topics. MA thesis, The Hartford Seminary Foundation.

Javier, J.R., F. Rosario, S. Benosa, and S. Klimenko. (2018). Ang Mag-antsi Sentence. *The Archive* 15: 73–109.

Kimoto, Y. (2017). A grammar of Arta: A Philippine Negrito language. PhD dissertation, Kyoto University.

Law, R. S. (1997). *The Romblomanon grammar essentials sketch*. Summer Institute of Linguistics. Available at: https://romblomanon.net/en/study-of-language/grammar.

Liao, H. (2011). On the development of comitative verbs in Philippine languages. *Language and Linguistics* 12, 205–237.

Liao, H. (2020). A reply to Blust (2019) The resurrection of Proto-Philippines. *Oceanic Linguistics* 59(12), 426–449.

Llamzon, T. A. (1975). Proto-Philippine phonology. *Archipel* 9, 29–42.

Lobel, J. (2013). Philippine and north Bornean languages: Issues in description, subgrouping, and reconstruction. Ph.D. dissertation, University of Hawai'i.

Malinowski, B. (1922). *Argonauts of the Western Pacific: An Account of Native Enterprise and Adventure in the Archipelagos of Melanesian New Guinea.* London: Routledge & Kegan Paul.

Manzano, D. (2021). Inati (Pancy Island, Philippines)—Language Contexts. *Language Description and Documentation* 20: 35–53.

Maree, R. (2007). *Ibatan: A grammatical sketch of the language of Babuyan Claro Island.* Linguistic Society of the Philippines Special Monograph Issue 53. Manila: Linguistic Society of the Philippines.

Matras, Y. (2009). Defining the limits of grammatical borrowing. In Angela Marcantonio (ed.), *The Indo-European language family: Questions about its status (Journal of Indo-European Studies Monograph Series 55)*: 1–25. Washington DC: Institute for the Study of Man.

McKaughan, H. P., & B. A. Macaraya. (1967). *A Maranao dictionary*. Honolulu: University of Hawaii Press.

Mead, D. (2020). Did Proto Malayo-Polynesian *d and *z merge in Proto Minahasan? - A review of the fate of PMP *d, *z, *j, *r and *l in Proto Minahasan and the Minahasan languages. Unpublished manuscript.

Mintz, M. (1971). *Bikol grammar notes*. Honolulu: University of Hawaii Press.

Or, E. (2018). A grammar of Iraya [Unpublished masters thesis]. University of the Philippines Diliman.

Pawley, A. (2006). Explaining the aberrant Austronesian languages of southeast Melanesia: 150 years of debate. *Journal of the Polynesian Society* 115, 215–258.

Paz, C. (1981). *A reconstruction of Proto-Philippine phonemes and morphemes*. Cecilio Lopez Archives of Philippine Languages and the Philippine Linguistics Circle.

Pennoyer, D. (1987). Inati: the hidden Negrito language of Panay, Philippines. *Philippine Journal of Linguistics* 18(2)/19(1), 1–36.

Rau, V. & M. Dong. 2006. *Yami Texts with reference grammar and dictionary*. Institute of Linguistics, Academia Sinica.

Reid, L. (1966). *An Ivatan Syntax*. Oceanic Linguistics Special Publication. Honolulu: Pacific and Asian Linguistics Institute, University of Hawai'i.

Reid, L. (1978). Problems in the reconstruction of Proto-Philippine construction markers. In S. Wurm & L. Carrington (eds.), *Second International Conference on Austronesian Linguistics: Proceedings* Vol. 1: 33–36). Department of Linguistics, Research School of Pacific Studies, Australian National University.

Reid, L. (1982). The demise of Proto-Philippines. In A. Halim, L. Carrington, & S. Wurm (eds.), *Papers from the 3rd International Conference on Austronesian Linguistics* vol. 2: *Tracking the travelers*: 201–216. Australian National University.

Reid, L. (1987). The early switch hypothesis: Linguistic evidence for contact between Negritos and Austronesians. *Man and Culture in Oceania* 3 (Special Issue): 41–59.

Reid, L. (1994). Unravelling the linguistic histories of Philippine Negritos. In Language contact and change in the Austronesian world, ed. by T. E. Dutton and D. T. Tryon, 443-475. Berlin: Mouton de Gruyter.

Reid, L. (2018). Modeling the linguistic situation in the Philippines. In R. Kikusawa & L. Reid (eds.), *Let's Talk about Trees*. Osaka: Senri Ethnological Studies, Minpaku.

Reid, L. (2019). Reassessing the Position of Isinay in the Central Cordilleran Language Family. In A. Lee & V. C. R. Santiago (eds.), *The Archive Special Publication* No.16, 1–32. Manila: University of the Philippines Diliman.

Reid, L. (2020). Response to Blust "The Resurrection of Proto-Philippines." *Oceanic Linguistics* 59(1–2), 374–393.

Robinson, L. (2008). Dupaningan Agta: Grammar, vocabulary, and texts. Ph.D. dissertation, University of Hawai'i at Mānoa.

Ross, M. (1995). Some current issues in Austronesian linguistics. In *Comparative Austronesian Dictionary*: 45–120. Mouton de Gruyter.

Ross, M. (2002). Final words: Research themes in the history and typology of western Austronesian languages. In F. Wouk and M. Ross (eds.), *The History and Typology and Western Austronesian Voice Systems*, 451–474. Canberra: Pacific Linguistics.

Ross, M. (2005). The Batanic languages in relation to the early history of the Malayo-Polynesian subgroup of Austronesian. *Journal of Austronesian Studies* 1(2), 1–24.

Ross, M. (2009). Proto Austronesian verbal morphology: a reappraisal. In K. A. Adelaar & A. Pawley (eds.), *Austronesian Historical Linguistics and Culture History: A Festschrift for Robert Blust*:295–326. Canberra: Pacific Linguistics.

Ross, M. (2012). In Defense of Nuclear Austronesian (and Against Tsouic). *Language and Linguistics* 13(6):1253–330.

Ross, M. (2020). Comment on Blust "The resurrection of Proto-Philippines." *Oceanic Linguistics* 59(1–2), 366–373.

Rubino, C. (2000). *Ilocano Dictionary and Grammar: Ilocano-English, English-Ilocano*. Honolulu: University of Hawaii Press.

Scheerer, O. (1918). Outlines of the history of exploration of the Philippine languages and their relatives in East and West. *The Philippine Review* 3(1), 59–67.

Shetler, J. (1976). *Notes on Balangao grammar*. Huntington Beach, California: Summer Institute of Linguistics.

Smith, A. (2017). The Western Malayo-Polynesian Problem. *Oceanic Linguistics* 56(2), 435–490.

Tanankingsing, M. (2009). A Functional reference Grammar on Cebuano. PhD dissertation, National Taiwan University.

Thomason, S. & T. Kaufman. (1988). *Language contact, creolization and genetic linguistics*. Berkeley: University of California Press.

Thomason, S. & T. Kaufman. (2001). *Language Contact*. Edinburgh University Press Edinburgh.

Yamada, Y. (2014). A grammar of the Itbayat language of the Philippines. Himeji, Japan.

Zorc, D. (1986). The genetic relationships of Philippine languages. In P. Geraghty, L. Carrington, & S. Wurm (eds.), *FOCAL II: Papers from the 4th International Conference on Austronesian Linguistics*: 147–173. Pacific Linguistics.

Zorc, D. (2019). Klata/Giangan—A new southern Philippine subgroup. *The Archive: Special Publication* 16: 33–51. Diliman: University of the Philippines.

Zorc, D. (2020). Reactions to Blust's "The resurrection of Proto-Philippines." *Oceanic Linguistics* 59(1–2), 394–425.