

# A Long Lost Sister of Proto-Austronesian? Proto-Ongan, Mother of Jarawa and Onge of the Andaman Islands

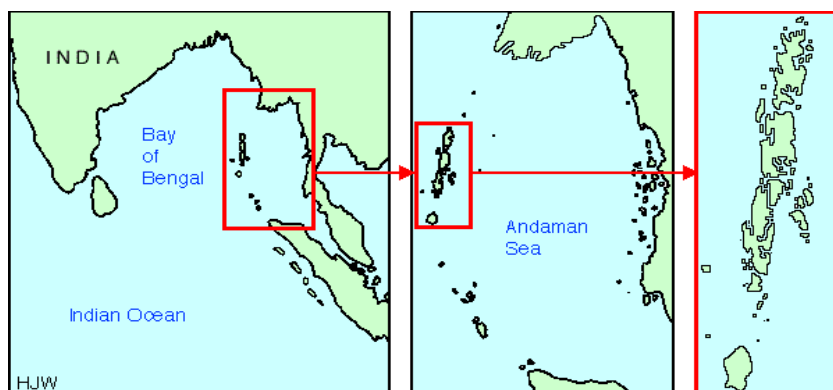
Juliette Blevins

MAX PLANCK INSTITUTE FOR EVOLUTIONARY ANTHROPOLOGY

This paper applies the comparative method to two related languages of the southern Andaman Islands, Jarawa and Onge, leading to the reconstruction of a protolanguage termed “Proto-Ongan” (PON). The same method is used to argue that Proto-Ongan may be related to Proto-Austronesian (PAN). Lexical and grammatical evidence suggests that Proto-Ongan and Proto-Austronesian are sisters, daughters of a Proto-Austronesian-Ongan (PAO). The implications of this discovery are wide-ranging, from potential solutions to problems in PAN grammar, to new hypotheses regarding ancient speaker migrations. While few of these implications are examined here, an extended Austronesian phylogeny is proposed in the hope that it will seed new avenues of research, and highlight the potential importance of Andamanese studies in understanding Austronesian prehistory.

**1. AN INTRODUCTION TO JARAWA AND ONGE, TWO LANGUAGES OF THE ANDAMAN ISLANDS.**<sup>1</sup> This paper makes use of the comparative method to argue that two related languages of the Southern Andaman Islands, Jarawa and Onge (a.k.a. Önge), are distantly related to Proto-Austronesian. Comparison of these two languages allows reconstruction of a protolanguage termed “Proto-Ongan” (PON).<sup>2</sup> The same method is used to suggest that Proto-Ongan is related to Proto-Austronesian

MAP 1: LOCATION OF THE ANDAMAN ISLANDS



(PAN). Before turning to the historical reconstructions, a brief introduction to these two languages of the Andaman Islands will inform the discussion that follows.

Jarawa and Onge are two endangered languages of the Andaman Islands, a cluster of over 200 islands in the Bay of Bengal between India and Myanmar (Burma) (see map 1). Jarawa and Onge are spoken in the southernmost part of the archipelago. Two other languages indigenous to this southern region are Jangil (Portman 1899), and Sentinelese, of North Sentinel Island, both unknown. Jangil is extinct, and has not been spoken for over 100 years. Sentinelese, on the other hand, is still spoken. However, the Sentinelese people do not welcome outsiders, and, to date, no one has been successful in engaging them in any dialogue, linguistic or otherwise.<sup>3</sup> According to recent counts,

1. This paper has a very unlikely origin that may interest those whose days are full of teaching and student advising. Several months ago, Mr. Pramod Kumar, a PhD student at the Max Planck Institute for Evolutionary Anthropology, passed along a draft of his phonology chapter on Jarawa for comment. I took the document home, and gave it a careful read, during which time I noticed resemblances between Jarawa words and words with similar sounds/meanings in Austronesian languages, or Proto-Austronesian itself. To convince myself that this was a figment of my imagination, I spent the next two days combing Jarawa wordlists. After this proved surprisingly fruitful, Mr. Kumar was kind enough to lend me his copy of Dasgupta and Sharma (1982) on Onge. When I saw that the Onge word for 'rain' was *ujeye* and 'communal house' *berai*, *berai*, I realized that the matter was serious enough to deserve more time and attention. Preliminary results of this effort are summarized here.

I thank Mr. Kumar for sharing his work with me, and for lending me books on Jarawa and Onge language and culture. Where my analysis of Jarawa differs in any way from his, it is noted in the text, or in footnotes. I am also grateful to Bernard Comrie, Director of the Linguistics Department, Max Planck Institute for Evolutionary Anthropology, for supporting this research in visible and invisible ways, and for seeming to have every article or book I needed that was not in the library. I am greatly indebted to colleagues whose superb work on comparative Austronesian provides the basis of this all, especially Bob Blust, John Lynch, Andy Pawley, Laurie Reid, Malcolm Ross, and Laurent Sagart. Additional thanks to Bob Blust, Bernard Comrie, Victor Golla, John Lynch, Yoram Meroz, Andy Pawley, and Søren Wichmann for comments on earlier drafts, and to George Weber and the Andaman Association for permission to reprint the maps.

2. The term "Ongan" (pronounced with initial unrounded central vowel) is based on Jarawa *əŋ*, Onge *əŋe* 'person, human being, one belonging to the group'. I follow Ken Hale's practice for the Pama-Nyungan language family of Australia, where *pama* and *nyunga* are different words for 'man, person' in languages of Cape York and southwestern Australia respectively. For cultural notes on the meaning of Jarawa *əŋ*, see Sreenathan (2001:6–7). Abbi (2006:94) uses the terms "Ang" for the same language group. Because "Egan" and "Angan" are both used for languages or language groups of Papua New Guinea, the term 'Ongan' is used here to reduce potential confusion.

Language abbreviations used in this paper are, in alphabetical order by abbreviation: ATY, Atayal; ISN, Isneg; ITB, Itbayat; JAR, Jarawa; ONG, Onge; PAI, Paiwan; PAN, Proto-Austronesian; PAO, Proto-Austronesian-Ongan; PAZ, Pazeh; PEO, Proto-Eastern Oceanic; PMP, Proto-Malayo-Polynesian; PNNG, Proto-North New Guinea Cluster; POC, Proto-Oceanic; PON, Proto-Ongan; PPH, Proto-Philippines; PWMP, Proto-Western Malayo-Polynesian; RUK, Rukai.

Proto-Austronesian (PAN), PWMP, PMP, and PPH reconstructions are from Blust's (1995) *Austronesian comparative dictionary* unless noted otherwise. Non-Formosan Austronesian living language data are taken from Tryon's (1995) *Comparative Austronesian dictionary* [CAD] unless noted otherwise. Formosan language data are taken from the following sources: for Pazeh, Li and Tsuchida (2001); for Paiwan, Ferrell (1982) and Egli (2002); for Atayal, Egerod (1980). Ross, Pawley, and Osmond (1998), volume 1 of *The lexicon of Proto-Oceanic*, is abbreviated LPO1. The electronic Austronesian Basic Vocabulary Database is abbreviated ABVD.

Where Jarawa or Onge forms are attested in a single source, or are for some other reason questionable, that source is given in the cognate sets, making use of the following abbreviations: for Jarawa, Abbi (2006) = [AA], Directorate of Tribal Welfare (2004) = [DTW], Kumar (2005) [PK1], Kumar (2006) [PK2], Senkuttuvan (2000) = [RS], Sreenathan (2001) = [MS]; and for Onge, Cogoy (2004) = [EC], Dasgupta and Sharma (1982) = [D&S], and Portman (1887) = [P]. In addition, Man (1883) = [EHM].

Jarawa has approximately 250 speakers, and Onge 94 speakers (Abbi 2006:7). Nearly all of these speakers are monolingual.

North of the area once covered by these four languages was a group of approximately 10 different languages, together termed “Great Andamanese,” spoken by the Great Andaman Tribes: Aka-Cari, Aka-Bo, Aka-Kora, Aka-Jeru, Aka-Kede, Aka-Kol, Akar-Bele, Aka-Bea, Oko-Juwoi, and A-Pucikwar. Map 2 shows the approximate locations of these tribes and their speakers in the 1800s in relation to Jarawa, Onge, Jangil, and Sentinelese, based on government officer reports of the mid- to late-nineteenth century, and their approximate distributions today. Currently, there are no more than 40 speakers of “Great Andamanese,” and most describe it as a “mixed language” or koine, involving a melange of the once 10 distinct languages, with dominant features of Aka-Jeru (Abbi 2006:20–22). Most of these speakers are bilingual in Andamani Hindi.

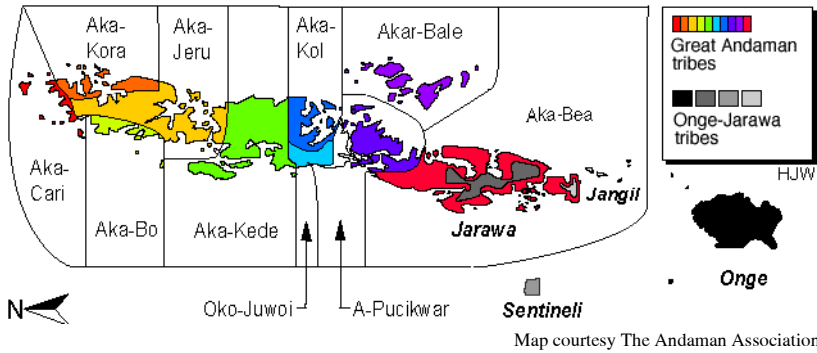
The indigenous people of the Andaman Islands, including all speakers of Jarawa and Onge, are “Negritos”—a purely descriptive term for black-skinned, frizzy-haired people of the Indo-Pacific region. They are very small people, averaging below 5 feet for both men and women, with skin as dark as that of any human group. The Andamanese have “peppercorn” hair, which grows in curled spiral tufts or cones, with naked skin visible in between. Some Andamanese women show the trait of steatopygia (or ‘fat bottom’), where large amounts of fat accumulate in the buttocks. In other characteristics, such as facial structure, for example, the Andamanese resemble other Negrito people of Southeast Asia, including the Negritos of the Philippines and the Semang of the Malay peninsula: the typical face is broad, with prominent cheek bones, and a broad, straight nose.<sup>4</sup>

The Andamanese appear to represent a preneolithic hunter-gatherer society (Man 1883, Radcliffe-Brown 1922). Agriculture was not practiced, and though stone tools were used, wood, shells, bones, and wild boar tusks were preferred as points. The Andamanese had no means of making fire, and always kept fires burning, carrying resin torches or smoldering logs through the forest. At the same time, they manufactured rafts, dugout log canoes, single outriggers, and a wide range of bows and arrows, axes, buckets, and baskets.<sup>5</sup> Shelters made of wood and woven palm leaves were of three different types, as in other parts of the Austronesian-speaking world (cf. Ross, Pawley, and Osmond 1998, chap. 3): small temporary shelters, or lean-tos, put together quickly while hunting; semipermanent individual family houses; and more solid large communal shelters. Ethnologies of the Andamanese include Man (1883) and Radcliffe-Brown (1922).

3. North Sentinel is also difficult to access. It is almost entirely surrounded by dangerous reefs. Or, at least, it was. The tsunami of December 26, 2004 put the Andaman Islands in the consciousness of many people around the world. A photograph of a Sentinelese man shooting a bow and arrow at a rescue helicopter was carried by the Associated Press. While many islands of Nicobar were submerged farther south, North Sentinel Island was actually uplifted by the earthquake, so that the reefs are now exposed well above sea level, and unlikely to survive (Weber 2006, chap. 4.) Many lives were lost in the tsunami, but the majority of Jarawa and Onge people survived. Nothing further is known about the Sentinelese.

4. Other human populations with steatopygia are the Khoi people of southern Africa and the Pygmies of central Africa. The Khoi also have peppercorn hair. Many biologists see the recurrence of certain physical features in Africa, Asia, and Oceania, including small stature and steatopygia, as local isolates resulting from the genetics of small isolated populations, rather than remnants of formerly widespread human traits.

MAP 2. TRIBAL TERRITORIES OCCUPIED IN THE 1860s



Recent archaeological studies, including radiocarbon dating of middens, can be found in Cooper (2002).

The physical world inhabited by the Andamanese is rich and varied. The flora and fauna of the islands are still being discovered, with thousands of plant species, including hardwood trees, palms, bamboos, canes, and orchids.<sup>6</sup> There is also an abundance of reptiles, from forty snake species (including cobras, vipers, coral snakes, and sea snakes), to salt water crocodiles, monitor lizards, turtles, small lizards, and geckos. Wild pigs, megapodes, hornbills, and pigeons are found in the dense forests, and dugongs graze in the off-shore waters, which are also home to hundreds of fish species, crustaceans, and mollusks. There are over a dozen species of indigenous mammals, including shrews, bats, and a palm civet, but no indigenous ungulates, cats, or larger mammals.<sup>7</sup>

5. The outriggers of the Andaman Islanders are flimsy and not used in rough water or for long sea voyages (Cooper 2002:26–29), but are of special interest in the context of Blust’s (1994:60) theory of outrigger monogenesis: “So far as can be determined the outrigger canoe was invented once in human history, by speakers of Austronesian languages either in Taiwan, or in the northern Philippines.” In the same paper, Blust (1994:63) suggests that the Andaman Islanders acquired the outrigger via contact with the Nicobarese, who themselves had contact with Austronesian-speaking people. An Andaman myth reported by Man (1883:165) tells of the time before the great flood, when dug-out canoes were made from pandanus trunks, and had no outriggers.

There is no question that Malays visited the Nicobars and Andamans in historical times, for the purpose of collecting edible birds’ nests and sea cucumbers, as well as to trade with the Nicobarese (Cooper 2002, chap. 2). How far back in time this contact goes is unknown, but it is consistent with PWMP-Ongan look-alikes in section 3.1, Class IV, on page 178, which appear to be loans.

6. Abbi (2006:3) notes that plantations of coconut and banana are a recent phenomenon. However, this should not be interpreted as meaning that coconut trees are not indigenous to these islands. On the contrary, Mathew, Mohandas, and Nair (2004:141) note that “the wild occurrence of certain cultivated species like coconut trees and betel vines in the Andaman and Nicobar Islands, suggests that these islands may perhaps be one of the probable centres of origin of these species.”

7. Whether the Andamanese wild pig originated from a domestic strain is not yet known. Groves (1981) argues that it originated from *Sus vittatus* of Southeast Asia; however, pig bones have been recovered from the lower levels of the Chauldari midden, suggesting it has been in the Andamans for at least 2000 years (Cooper 2002). Dogs were introduced in the late 1800s when the penal settlement was established, the spotted deer (*Axis axis*) was brought to the islands in the 1920s, and elephants have been brought since then from India for work-projects.

Linguistic work on languages of the Andamans up until approximately 1988 is summarized in Zide and Pandya (1989). This annotated bibliography includes references to primary linguistic data, as well as articles on internal and external relationships of the Andamanese languages. Subsequently, a number of important publications have appeared, including Manoharan (1989), Sreenathan (2001), Abbi (2006), and Weber (2006). Manoharan (1989) includes descriptions of Andamanese phonology and morphology, notes on typology, and a discussion of subgrouping as well. Sreenathan (2001) discusses Jarawa language and culture with extensive vocabulary in the areas of flora and fauna, kinship terms, given names, body parts, and material, medicinal, and aesthetic culture. Abbi (2006) provides new data on Great Andaman, Jarawa, and Onge, based on fieldwork carried out in 2001. Weber (2006) consolidates a wealth of information on the Andamanese, their language, culture, history, and physical environment. This on-line resource includes a range of sketches and photographs from early encounters with the Andamanese, and has been extremely useful to me in matching linguistic terms with artifacts, flora, and fauna. Kumar (2005, 2006, in progress) describes Jarawa based on fieldwork started in 2001.

A family relationship between Jarawa and Onge has been assumed or proposed by nearly all researchers working in this area. Radcliffe-Brown (1914) provides a comparative wordlist, Nair (1979) notes additional lexical and structural similarities, as do Manoharan (1989, chap. 4) and Abbi (2006:92–93). Though Abbi refers to “shared innovations in phonology,” none are mentioned explicitly, nor is there any attempt at reconstructing the protolanguage that these shared innovations continue. Section 2 of this paper presents the results of applying the comparative method to Jarawa and Onge data. This supports the conclusion that Jarawa and Onge are indeed related to each other, descending from a common mother tongue.

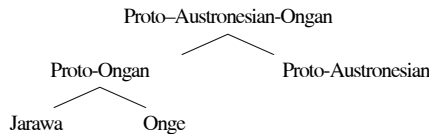
On the other hand, a relationship between Jarawa and Onge and languages of the Great Andaman group is not widely accepted. Radcliffe-Brown (1914:40) found only seven potential cognates between Onge and Bea/Jeru, and noted that the difference between Onge and the Great Andaman languages “is such that it would not be possible from consideration of the vocabulary alone to prove that they belonged to the same language stock. It is only when the grammatical structure of the language is examined that it is possible to show that they are really all derived from one original language.” Abbi (2006:93) is agnostic, stating that “current linguistic analysis does not, with any certainty, indicate any genetic relationship between Great Andamanese and the other two languages.” The only positive evidence offered in support of this relationship is a listing of 17 word pairs as proposed cognates in Manoharan (1989:166–67).<sup>8</sup> There are several problems with Manoharan’s proposal. First, in these sets, Onge and Jarawa words are compared with unidentified Great Andaman languages. Second, although some words are similar in sound and meaning, others are not: consider, for example, Onge *ikalu* ‘hot’, ‘Andaman’ *iximil* ‘hot water’, or Onge *tekala* ‘white cloud’, ‘Andaman’ *ʔɔ:l* ‘white mud’. Third, no argument is presented that similarities reflect direct inheritance as opposed to borrowing or chance. As I show in section 3 below, regular loss of initial \*b /\_u in Proto-

8. Manoharan (1989:166–67) looks like a list of 20 sets, but three sets are repeated: sets (1) and (17) represent the same Onge word for ‘water’, (3) and (18) are the same word for ‘hot’, and (11) and (20) are the same lexeme for ‘drink’.

Ongan leads one to suspect that the shared term *bucu* ‘cooking pot, vessel, aluminum pot’ may be a loan. Manoharan’s (1989:166–67) comparison of Ongan *bucu*, Beja *buj*, and Jeru *pec*, then, may be taken to support not a genetic relationship, but evidence of a term borrowed from Great Andaman into Jarawa-Ongan. Given evidence that shows these languages have been in contact, and the scarcity of data available at present on Great Andaman languages, there remains no persuasive evidence of a family relationship between Jarawa-Ongan and the Great Andaman languages.<sup>9</sup>

The hypothesis argued for in this paper is that Proto-Ongan, mother of Jarawa-Ongan, is a sister of Proto-Austronesian, as depicted in figure 1. This position differs significantly from earlier hypotheses on external relations between Andaman languages and other languages or language families. Three positions on external relations have been taken since study of these languages began. The most widely accepted view, first put forth by A. J. Ellis in his 1882 address to the Philological Society, is that the Andamanese languages are not related to any other known language or language family (see also Bloch 1952). This position may be undermined by the data presented in section 3, which suggest a direct family relationship with Proto-Austronesian. A second position is that Andaman languages (including Jarawa and Ongan) are descendants of “Proto-Negrilo,” a hypothesized language once spoken by other Asian Negrilo groups, itself unrelated to any other known language or language family. Despite Reid’s (1994, 2006) continuing work on identifying unique Philippine Negrilo lexemes, no progress has been made on identifying cognates in Andamanese or elsewhere. I return to aspects of this hypothesis in section 4. The last proposal to consider is Greenberg’s (1971) Indo-Pacific hypothesis, according to which the Andaman languages belong to a linguistic superstock that includes the non-Austronesian (Papuan) languages of New Guinea and surrounding areas, and the indigenous languages of Tasmania. However, from the start, Greenberg (1971:810) is unconvinced of the relation between Great Andaman and Ongan-Jarawa, agreeing with Radcliffe-Brown (1914) that “...there are very few vocabulary resemblances between this language [Ongan—JB] and those of Great Andaman and the only real point of contact is typological. ... A few citations from Ongan have been included in the general Indo-

FIGURE 1. PROTO-AUSTRONESIAN-ONGAN



9. This is not to say that look-alikes in basic vocabulary between Great Andamanese and Proto-Austronesian cannot be found. On the contrary, the Southern Andaman language Bojig-ngiji described by Man (1883) contains several notable look-alikes including *akana* ‘meal, food’ (Man 1883:123) (cf. PAN \*kan-en ‘food’) and *elma* ‘palm of hand’ (cf. PAN \*lima ‘hand’). The same is true of the Andamanese languages described by Portman (1887), where we find, for example, Aka Bia-da *aka-tang-da*, Aka Bojigiab *otong-da* ‘tree’ (cf. cognate set [3.88]), Aka Bojigiab *chemal-da*, Aka Kede *chemil* ‘grass’ (cf. PAN \*Cemel ‘grass’, Blust 1999b). Nevertheless, where regular sound correspondences relate Ongan and Jarawa to PAN, the same cannot be said, at present, for PAN and Great Andamanese.

Pacific vocabulary, but both its special relationship to the languages of the rest of the Andamans and its assignment to Indo-Pacific must be considered highly provisional.” Because the proposal here is that Onge and Jarawa are daughters of a sister language to Proto-Austronesian, they cannot, at the same time, be non-Austronesian languages. If there is indeed an Indo-Pacific stock, Onge and Jarawa do not belong to it.<sup>10</sup>

With this as background, I will briefly present an overview of Jarawa-Onge synchronic and diachronic phonology in section 2, along with a preliminary reconstruction of Proto-Ongan. Evidence supporting a family relationship between Proto-Ongan and Proto-Austronesian is presented in section 3, with a summary discussion in section 4.

**2. ONGAN SYNCHRONIC AND DIACHRONIC PHONOLOGY.** There has been little past work on the synchronic and diachronic phonology of Jarawa and Onge. As a consequence, the present section combines philological work with traditional methods of phonological analysis. On the philological side, a range of works presenting Onge and Jarawa data without phonological analysis have been used. The most important of these are Portman (1887) for Onge, and Sreenathan (2001) for Jarawa. Interpretation of these data sets have been greatly aided by Dasgupta and Sharma’s (1982) work on Onge synchronic phonology, Cogoy’s (2004) on-line Onge dictionary, Abbi’s (2006) preliminary outlines of Jarawa and Onge phonemic systems, and Kumar’s (2005, 2006, in progress) ongoing description of Jarawa. These data sources supplemented, occasionally, by other works on the Onge and Jarawa mentioned in section 1, constitute the primary database for the linguistic analysis of Jarawa and Onge that follows. Though descriptions are limited and preliminary, comparable basic vocabulary (see appendix 1) and grammatical sketches for Jarawa and Onge, along with detailed reconstructions for Proto-Austronesian and its descendants, provide a firm basis for application of the comparative method.

**2.1 SYNCHRONIC PHONOLOGY.** Jarawa and Onge have similar phonological systems. Both languages have oral and nasal stops at bilabial, alveolar/dental, and velar points of articulation. Both languages have palatal affricates, and both lack oral fricatives. In addition, both languages have sonorants *l*, *r*, *w*, *y*. Both languages appear to have a voicing contrast for nonlabial stops and affricates, and Jarawa also appears to have a *p* vs. *b* contrast, though in Onge there is only *b*.<sup>11</sup> Another major difference in the consonant systems of the two languages is that Onge has a labialized velar *k<sup>w</sup>* and no *h*-sounds, while Jarawa lacks *k<sup>w</sup>* but has *h* and *h<sup>w</sup>*. Two aspects of Jarawa and Onge consonant systems are underdetermined by data currently available. First, both languages appear to have surface retroflex sounds (Abbi 2006; Kumar 2005, 2006). However, it is not clear whether these are contrastive or not.<sup>12</sup> Second, a series of aspirated consonants has been

10. Greenberg presents 84 Indo-Pacific cognate sets, 35 with Andamanese words, and only three (36. hair, 77. tooth, and 80. to walk) with Onge forms. No Jarawa data are included. The Onge forms cited are: *de* ‘hair’; *ku*, *kwe* ‘tooth’; *cige* ‘leg’. Based on Onge data from Portman (1887), Dasgupta and Sharma (1982), Cogoy (2004), and Abbi (2006), more accurate phonological transcriptions of these words would be: *-ode-*, *-ore-* ‘hair’; *-akwe-* ‘tooth’, *-icige* ‘leg’.

Greenberg’s Indo-Pacific hypothesis is widely rejected by historical linguists, due to its critical methodological flaws. See Campbell (1998, chap. 13), and Campbell (2003) for discussion of the general issues involved.

described for Jarawa, in addition to the plain voiceless and plain voiced stops for all but the bilabial point of articulation (Abbi 2006; Kumar 2005, 2006). Again, while aspiration is phonetically apparent, it is not clear whether this is an allophonic feature of the voiceless (nonlabial) stop series, or a truly contrastive feature.

The vowel systems of the two languages are also strikingly similar. Both languages have five noncentral vowels *iueoa*, with lax allophones in closed syllables, and both languages have at least one central vowel, which I will write here as *a*, though it is sometimes higher. Abbi (2006) and Kumar (2005) propose an additional high central vowel phoneme, but again, evidence for its contrastive status is lacking.

Jarawa words can be monosyllabic or longer, and all content words (nouns, verbs, adjectives, adverbs) are minimally bimoraic, where vowels and coda consonants count as moras. Jarawa words may begin or end with either vowels or consonants, and the maximal syllable type is CVC. Abbi (2006:58–59) describes four synchronic processes in Jarawa phonology: (i) intervocalic voicing of *c* to [j] in derived environments; (ii) elimination of the first vowel *a* when followed by another vowel across a morpheme boundary; (iii) vowel harmony, whereby *a* becomes [o] when followed by a round vowel in the next syllable; and (iv) elision of whole syllables in fast speech.

Onge words can also be monosyllabic or longer, with no clear word minimum (e.g., *le* ‘honey’). Onge words may begin with either vowels or consonants, and the maximal syllable type is CVC. One major difference between Jarawa and Onge is that, with the exception of imperative verbs ending in glottal stop (e.g., *kaʔ* ‘give’), all Onge words end in vowels. Where a Jarawa word is C-final, one often finds a cognate with final *e* in Onge: Jarawa *ij*, Onge *iʔe* ‘water’; Jarawa *inen*, Onge *inene* ‘foreigner’; Jarawa *dag*, Onge *dage* ‘coconut’. As in Jarawa, mid-vowel harmony is in evidence.<sup>13</sup> For example, the Onge 1st person plural prefix *et-* occurs as [ot-] when the vowel in the next syllable is round *u, o*: *et-ejale* ‘our faces’ but *ot-oticule* ‘our heads’ (Dasgupta and Sharma 1982:11–13). Another phonological process is the delabialization of *k<sup>w</sup>* before round vowels *u, o*: the 3rd person plural prefix *ek<sup>w</sup>-* alternates, as in *ek<sup>w</sup>-ejale* ‘their faces’ but *ok-oticule* ‘their heads’. Although both *d* and *r* are posited as separate phonemes in Onge, *r* is not found word-initially, and *d* alternates with [r] intervocalically. This “flapping” is evident in cases where the dual suffix *-dena* occurs. Compare *inene* ‘foreigner’ and *kue* ‘pig’ with *inendena* ‘two foreigners’ but *kuerena* ‘two pigs’. On the other hand, items that arguably have a phoneme *r* show a differ-

11. Word-initially, the *p* vs. *b* contrast appears to be disappearing in Jarawa as well. In all sources I have consulted, there are words that are transcribed by the same author as having either [p] or [b] initially. Because there are other words that consistently show [b], I assume that words with [p] ~ [b] variants are in the *p* class (from PON \*p), and that a change in progress may be in evidence.

In general, obstruent voicing contrasts are problematic in the Jarawa and Onge data, with much variation in how particular consonants are transcribed. In both languages, this may be due, in part, to conditions on intervocalic voicing. In word-initial position for Jarawa, however, there is still a great deal of variation in transcriptions. These are noted where relevant.

12. Neither Portman (1887) nor Dasgupta and Sharma (1982) describes a contrastive retroflex series for Onge, though retroflexion is transcribed in native Jarawa and Onge words by Abbi (2006) and Kumar (2005, 2006). Retroflexion is audible in both languages, though it seems variable.

While it might be tempting to attribute this feature to Hindi or some other contact language, unlike the Great Andamanese, the majority of Jarawa and Onge are monolingual.

13. Vowel harmony is optional for some speakers, and obligatory for others (Dasgupta and Sharma 1982:11–13).



ent kind of variation, where *r* varies with surface [r], [y], or [l]. The conditions for this are not clear, but it seems especially common after a labial vowel or consonant: compare *koy-age*, *kwarage*, *kwalage* ‘string, rope’ (Cogoy 2004).

Onge has lexical vowel/zero alternations and consonant assimilation processes that apply to affixed words. When two consonants come together across a morpheme boundary, the second triggers place-assimilation with the first, resulting in homorganic sequences, including geminates. For example, Onge *daje* ‘tree; dugout canoe’ loses its final vowel and undergoes place assimilation in the dual and plural forms *dandena* ‘two canoes’ and *danle* ‘canoes’, while *umuge* ‘pigeon’ shows parallel changes: *umuddena* ‘two pigeons’ (where *gd* → *dd*), *umulle* ‘pigeons’ (where *gl* → *dl* → *ll*). In these forms, word-final *e* does not surface under affixation. In fact, these suffixal alternations offer the opportunity for internal reconstruction, as they show a range of assimilatory processes that appear to be reflections of earlier regular sound changes. The relevant assimilations are shown in table 1.

Recall from the preceding discussion that many words ending in consonants in Jarawa end with *e* in Onge. These vowels are historically exrescent, with two pieces of evidence in the synchronic phonology of Onge supporting this historical analysis. First, as illustrated with the examples above, nonetymological final *e* does not surface when nouns are suffixed with number markers. Compare (1a–c) where final *e* does not appear under suffixation, with (1d–e) where it is maintained.<sup>14</sup>

(1)	BASE	DUAL	PLURAL	
a.	<i>inene</i>	<i>inen-dena</i>	<i>inen-le</i>	‘foreigner’
b.	<i>daje</i>	<i>dan-dena</i>	<i>dan-le</i>	‘tree; dugout’
c.	<i>umuge</i>	<i>umud-dena</i>	<i>umul-le</i>	‘pigeon’
d.	<i>kue</i>	<i>kuc-rena</i>	<i>kuc-le</i>	‘pig’
e.	<i>ale</i>	<i>ale-reni</i>	<i>ale-ni</i>	‘child’

A distinct morphological construction reveals a similar contrast between two types of *e*-final words. Nouns in Onge may be suffixed with a definite article that has the two allomorphs *-gi* and *-i*. Words like those in (1a–c) take *-gi*, while all others take *-i*. This suggests a historical lenition of *g* between vowels across the word/clitic boundary. Because words like those in (1a–c) were consonant-final, no such lenition occurred, and the full form *gi* is maintained. Definite forms in (2) illustrate this: words in (2a,b) were historically C-final, while those in (2c,d) were V-final.<sup>15</sup>

(2)	BASE	DEFINITE	
a.	<i>daje</i>	<i>daje-gi</i>	‘tree; dugout’
b.	<i>umuge</i>	<i>umuge-gi</i>	‘pigeon’

14. The dual *-dena*, *-rena* has allomorphs *-deni*, *-reni*, and the same nouns that take the second set also take *-ni* (instead of *-le*) in the plural (Dasgupta and Sharma 1982:13–14).

15. Dasgupta and Sharma (1982:14) remark that “it is not possible to define the distribution of the article *-gi* or *-i* either phonologically or morphologically on the basis of present data.” Comparison with Jarawa illuminates the historical explanation, because cognate forms of nouns taking *-gi* lack final vowels. However, though many Jarawa words are written as consonant final, the same words can be pronounced with light “exrescent” vowels, or final audible release. Whether this is due to the influence of Onge, or is simply a sound change in progress, is difficult to say. Where forms with and without such final vowels are attested, I write only the C-final form; otherwise, I write the vowel-final form. The quality is typically a schwa or copy of the preceding vowel.

c. <i>kuc</i>	<i>kuc-i</i>	'pig'
d. <i>ebe</i>	<i>ebe-i</i>	'thread'

Another process evident in Onge but absent in Jarawa is the reduction of NC clusters to single Cs, as in variants *ijjo-*, *-ijo-* 'to drink' (cf. Jarawa *-ijico*). Onge also shows another type of nasal-consonant alternation: in forms with DVN sequences, where D is a voiced obstruent, and N is a nasal, we find optional DVN > NVN. An example is *bone* (Dasgupta and Sharma 1982:123), *mone* (Portman 1887:65) 'resin, resin torch', which appears to be from \*pone, based on Jarawa *pone* 'resin, resin torch'.

**2.2 DIACHRONIC PHONOLOGY.** With this much as background, I will move to the goal of this section: to make the family relation between Jarawa and Onge concrete by establishing regular sound correspondences between the two languages. Sound correspondences are shown in tables 2 and 3, based on the comparative data in appendix 2. Regular sound changes are discussed below, with numeric references to sets in appendix 2 (as, e.g., [2.21]).

As can be seen in table 2, the nasals, liquids, glides, and vowels are relatively stable. In the voiceless series, there are several notable changes. Onge has undergone merger of \*p/\*b to *h*. In Jarawa, this merger may be underway (see fn. 11), but seems to be incomplete. Other voiceless stops *t* and *k* undergo intervocalic voicing in Onge, though there may also be prosodic conditions that inhibit this. PON \*k, \*k<sup>w</sup> undergo debuccalization in Jarawa, resulting in *h* and *h<sup>w</sup>*, respectively. Jarawa *k* occurs in words where this sound change was blocked (see below, [3.41]), and in loans. The voiced stop series shows some

TABLE 1. SOME SYNCHRONIC & HISTORICAL CC ASSIMILATIONS IN ONGE

C1	C2	_d	_l
n		nd	nl
m		nd	nl
ŋ		nd	nl
g		dd	ll

TABLE 2. ONGAN CONSONANT CORRESPONDENCES  
(IN NONCLUSTER CONTEXTS)

Proto-Ongan	*p	*b	*t	*d	*k <sup>w</sup>	*k	*g	*y	*w
Jarawa	p, b	b	t	d	h <sup>w</sup> , h	h	g, y	y	w
Onge	b	b	t, d	d, r	k <sup>w</sup> , k	k, g	g, θ	y	w

Proto-Ongan	*c	*j	*m	*n	*ɲ	*l	*r
Jarawa	c	j	m	n	ɲ	l	r
Onge	c, j	j	m	n	ɲ	l, y	r/y/l, θ

TABLE 3. ONGAN VOWEL CORRESPONDENCES  
IN OPEN NONFINAL SYLLABLES

Proto-Ongan	*i	*u	*a	*e	*o	(*ə)
Jarawa	i	u	a	e, ə, o	o	(ə)
Onge	i	u	a	e, ə, o	o	(ə)

lenition in Onge and Jarawa, with \*d going to *r* intervocalically in Onge, and \*g leniting intervocalically to *y* in Jarawa, again with probable prosodic conditioning as well.

The vowels are also stable in open nonfinal syllables. The only exception is that Jarawa *e* is often reduced to *ə* in unstressed syllables. Note that in table 3, Proto-Ongan \*ə is parenthesized. This is because schwa has a limited distribution in PON, occurring only before nasals, in the morphemes \*ən- 'personal prefix' and \*əŋ 'person'.<sup>16</sup> Further work could reveal additional conditions that allow schwa to be viewed as an allophone of \*e before nasals in PON.

Changes not coded in tables 2 and 3 include the consonant assimilations in clusters noted earlier in the context of Onge suffix-induced alternations (table 1) and insertion of excrescent *e* after Onge word-final consonants.

Regular sound changes identified from PON to Onge are listed in (3).

- (3) a. \*d > r / V\_ {V,#} [2.25, 2.77, 2.83, 2.97, 2.103] (precedes and feeds 3b)  
 b. \*{w,r} > Ø / {a,e}\_ # [2.15, 2.69, 2.103]  
 c. \*Ø > e / C\_# [2.3, 2.6, 2.29, 2.30, 2.42, 2.50, 2.53, 2.95]  
 d. \*p > b [2.90, 2.91, 2.92]  
 e. \*aw. > o [3.45]

Onge sound changes in (3) include word-final loss of coda {w, r} after *e* (3b), and the insertion of an excrescent vowel in C-final words (3c), which both result in canonical V-final words. Other regular sound changes are the lenition of \*d (3a), which, based on [2.103], should precede (3b); neutralization of PON \*p, \*b to Onge *b* (3d); and coalescence of \*aw > *o*. Onge shows more intervocalic voicing of obstruents than Jarawa, and in some instances, voiced and voiceless reflexes are found in derivationally related forms, e.g., *g/k* in *-agage* 'breast' vs. *agak-ota* 'nipple', where *-ota* appears to be a reduced variant of *-oto* 'seed, fruit'.

Regular sound changes identified from PON to Jarawa are listed in (4).

- (4) a. \*k, k<sup>w</sup> > h, h<sup>w</sup> [2.4, 2.36, 2.37, 2.47, 2.48, 2.60-2.67, 2.93, 2.96]  
 N.B. blocked by another *h, h<sup>w</sup>* in same word [2.57, 3.60]  
 b. \*Ø > a / h#\_ [3.18b, 3.99, 3.100]  
 c. \*g > y [1.12, 1.33]

The most significant of these sound changes is the debuccalization of the voiceless velar stops (4a), resulting in phonemes *h* and *h<sup>w</sup>*. In addition, there is vowel epenthesis after word-final *h* (4b), and lenition of \*g (4c), which may be in progress (cf. [1.48]). Reduction of the nonround nonhigh vowels \*e and \*a to schwa in unstressed syllables is variable.

Though certain aspects of the Proto-Ongan sound system remain to be discovered, it is possible to establish a solid set of regular sound correspondences between the two languages, and a preliminary reconstruction of Proto-Ongan (appendix 2). With these correspondences and reconstructions in hand, we are in a position to compare Proto-Ongan and Proto-Austronesian. As I show below, this comparison reveals a striking number of cognate sets that seem to defy chance resemblances, and whose semantic and morphological characteristics are difficult to explain in terms of borrowing or contact-induced change alone.

16. These two morphemes could reflect a single morpheme \*en historically, assuming \*n > ŋ word-finally, and unstressed \*e > ə before nasals.

**3. EVIDENCE FOR PROTO-AUSTRONESIAN-ONGAN.** In this section, I suggest that Proto-Ongan and Proto-Austronesian descend from the same mother tongue, as schematized in figure 1. Speakers of a mother language, Proto-Austronesian-Ongan (PAO) split into at least two groups in Austronesian prehistory. In 3.1 and 3.2, I present lexical and grammatical evidence for a genetic relationship between Proto-Ongan and Proto-Austronesian. In 3.3, additional evidence is presented that is neither lexical nor purely grammatical.

**3.1 LEXICAL EVIDENCE.** Preliminary reconstruction of Proto-Ongan allows us to compare this language with Proto-Austronesian. This comparison yields cognates and allows regular sound correspondences to be identified, as shown in tables 4 and 5, with regular sound changes (abbreviated SC below) summarized in table 6.

The best attested regular sound change is the loss of PON \*b word-initially before \*u (SCi.) This sound change results in correspondences of PAN \*bu... with PON \*u.. [3.18–3.23]. Initial \*b-loss also seems to have occurred before \*i, though there are fewer clear examples [3.16, 3.17]. In some words of more than one syllable, PAN initial \*q corresponds to zero in PON (SCiii) [3.2, 3.32, 3.39]. It is unclear whether these forms reflect \*q-loss in PON or \*q-insertion in PAN. Because in some cases there are PAN doublets (e.g., PAN \*qaleb, \*cleb ‘knee’), and because some of these are body-part terms that may

**TABLE 4. SOME PAN-PON CONSONANT CORRESPONDENCES**

Proto-Austronesian	*p	*b	*t	*d	*k	*g	*q	*h
Proto-Ongan	*p	*b, Ø	*t	*d	*k	*j,g	*k	*h,y,Ø
Proto-Austronesian	*m	*n	*ɲ	*ŋ	*N	*ʔ	*ku	*qu
Proto-Ongan	*m,-ŋ	*n,-ŋ	*ɲ,-ŋ	*ŋ	*l,y	*ʔ	*k <sup>w</sup>	*k <sup>w</sup>
Proto-Austronesian	*c,C	*s, S	*j	*z	*l	*r	*R	*w
Proto-Ongan	*c	*c	*j,y	*c	*l	*r	*l,r	*w

**TABLE 5. SOME PAN-PON VOWEL CORRESPONDENCES IN OPEN NONFINAL SYLLABLES**

Proto-Austronesian	*i	*u	*a	*ə (=‘e’)	*ay#
Proto-Ongan	*i	*u, o	*a,c	*e	*e

**TABLE 6. REGULAR SOUND CHANGES IN PON/PAN**

	Proto-Ongan	Proto-Austronesian
SCi.	#bu > #u	#bu
SCii.	#bi > #i	#bi
SCiii.	#V	#qV
SCiv.	q > k	q
SCv.	qw > kw (+ kw < ku)	qw > q, w; kw > k, w
SCvi.	T > Ø / _#*	final Cs maintained (in unstressed syllables??)
SCvii.a.	n > ŋ, Ø / _#	final nasals maintained
SCvii.b.	m, ɲ > ŋ, Ø / _#	final nasals maintained (in progress?)
SCviii.	ay# > e#	ay#

\* The symbol T represents any oral stop.

have been consistently vowel-initial in the protolanguage (see 3.3.1), I reconstruct them as vowel-initial in PAO, assuming \*q-insertion in PAN. Another regular correspondence is that between PAN \*q and PON \*k (ONG *k*, JAR *h*) [3.50–3.53]. Where PON shows a labialized velar \*kw, this continues \*qw or \*kw < \*ku. Interestingly, PAN shows neutralization, either by delabialization to \*q, \*k or by loss of the stop to \*w (SCv) [3.77, 3.89]. In monosyllables, final consonants are maintained [3.14, 3.28, 3.73], but in multisyllabic forms, word-final oral stops are typically lost in final position (SCvi) [3.5, 3.20, 3.32, 3.39]. In some cases, there is evidence for both the C-final and V-final forms in PON [3.23], in which case both reconstructions are shown. Neutralization of final nasals is in evidence, though it appears to be in progress (SCvii). In some forms, final nasals are lost altogether [3.41], while in others, there is loss, or neutralization to the velar nasal [3.34]. Place of articulation appears to be significant, with neutralization complete with dental/alveolars, but incomplete with palatals and labials. Finally, there is evidence of \*ay > \*e in PON (SCviii) as in [3.2, 3.15].

Vowel correspondences are fairly regular, with only minor shifts in the vowel space. In PON, \*o has several sources: unconditioned lowering of \*u [3.47], closed syllable lowering of \*u [3.99, 3.100], rounding assimilation [3.90], and coalescence of \*wa > o [3.18b]. When followed by a palatal, when word-final, or when unstressed, PAN \*a may correspond with PON \*e. In syllables closed by nasals \*e is centralized to schwa when not preceded by a palatal, whereas elsewhere, it is realized as a mid front vowel.<sup>17</sup> Because regular correspondences have yet to be fully understood, where deviations from the correspondences in table 5 occur, they will be noted in the text below each set.

The cognate sets below are organized as follows. Class I includes sets where PAN reconstructions have already been proposed. In a few sets in class I, I propose slightly modified PAN reconstructions based on the earlier PAN/PMP reconstructions of Blust (1995); all of these are marked by double-asterisks, and in cases where only the meaning of the form has been adjusted, double asterisks precede the gloss I have added. Class II includes sets where cognates are found in PMP and Formosan languages, but no PAN reconstruction has been proposed previously. In these cases, I propose new PAN reconstructions, with all of these, again, marked by double asterisks. Sets I and II are the basis of the primary hypothesis in this paper: that PON and PAN are sister languages, as evidenced by regular sound correspondences. Class III shows sets where cognates are found only in Formosan languages. If these sets are valid, they provide further evidence for the central hypothesis in this paper, as there is no evidence that speakers of Formosan languages or their ancestors have been in contact with Andamanese Islanders in historic or prehistoric times. Finally, Class IV includes lexical items in Onge and Jarawa that closely resemble forms in Malayo-Polynesian languages, but where Formosan data are lacking, or where correspondences are incorrect. In this class, PON reconstructions are only proposed if sound correspondences reflect those described above, which are characteristic of classes I/II and III. Otherwise, the sets in class IV stand out as potential borrowings, both in terms of their sound patterns and their semantic fields. If these are borrowings, and the general hypothesis of PAN-PON sisterhood is correct, Ongan is a language family that shows evidence of direct and indirect inheritance relations with Austronesian languages. Directly

17. This centralization appears to be regular before nonlabial nasals, and variable before \*m.

inherited forms from Proto–Austronesian–Ongan show the sound correspondences summarized in tables 4, 5, and 6. Indirect inheritance in the form of loans from Malayo-Polynesian languages spoken to the south and east may involve forms not reconstructible for PAN, and may not show these sound correspondences.

Where relevant, notes on etymologies follow each cognate set. Attempts have been made to use only sets with near-perfect consonantal correspondences. Where irregularities arise, they are noted in the text. Because vowel transcriptions tend to vary much more across sources, and because vowel shifts are less well understood, I have been more lax in allowing seemingly imperfect vowel correspondences.

In sets where only a Jarawa or Onge lexeme is known, the Proto-Ongan reconstructions should be taken as highly preliminary. In some cases, sets are based on PAN CVC roots (Brandstetter 1906, 1918; Blust 1988). Some Jarawa or Onge roots are preceded by initial vowels that have multiple sources, but have been reanalyzed as part of the stem. In all cases where only part of a word is compared, a slash separates the relevant root from the rest of the word. Reconstructions or attested forms on which reconstructions are based are written in bold. Other relevant comparanda are in nonbold italics. PAN reconstructions are from Blust (1995), unless noted otherwise. Language abbreviations, sources, and source abbreviations are listed in footnote 1.

### Austronesian-Ongan Cognate Sets

#### CLASS I. WITH (PREEXISTING) PAN RECONSTRUCTIONS

[3.1] PAO \***aCa** ‘high up’: PON \***-eca-** ‘high up, upper part; face’; \***-ecepto** ‘eye’ (< \***eca-ipo**, \**-ipo* ‘flesh’) (JAR **ece/po** ‘eye’; **ece/kulu** ‘forehead’, *kulu* ‘head’; ONG **-eja/le** ‘face’, **-eja/tore** ‘beard on the face’, *-tore* ‘the hair’; **eje/bo** ‘eye’, **eje/tati** ‘skin of forehead’); PAN \*\***aCa** ‘high up’ (cf. PAN \***aCa/s** ‘high, tall’; PAN \***m/aCa** ‘eye’ [Blust 1999a], PAI **qatsa** ‘big’, **qatsa-qatsa** ‘tall’).

For further discussion of vowel-initial bodypart terms in PON, see 3.2.1. PON \***a** > **e** before palatals may be restricted to unstressed syllables. Though parsing of PAN \***maCa** into earlier \***ma** + **aCa** may seem unmotivated, it is one of the few \***m**-initial PAN content words that is not prefixed. In addition, at least two languages of the Philippines lack reflexes of initial \***ma**-. Binongan Itneg and Kalinga (Guinaang) **?ata** ‘eye’ (Reid 1971). See 3.3.3 for further discussion.

[3.2] PAO \***aCay** ‘liver’: PON \***-ace/ŋ** ‘blood’ (< \***ace-ig** ‘liver-water’) (JAR **-aceŋ**; ONG **-aceŋe**); PAN \***q/aCay** ‘liver’.

For further discussion of vowel-initial bodypart terms in PON, see 3.2.1.

[3.3] PAO \***aku** ‘self, ego; 3rd person pronoun’: PON \***-aku/i** ‘self, ego’, \***-ekw/i** ‘3rd person pronoun’, where \***-i** is PON definite suffix (JAR **əhwi-** ‘3rd person nominative pronoun’; ONG **-ekwi** ‘3rd person definite nominative plural pronoun’; ONG **-akui** ‘self, ego’, added to pronouns to form reflexives, e.g., **m-akui** ‘myself’, **et-akui** ‘ourselves’); PAN \***aku** ‘1st sg. actor; I’.

The PON doublet appears to correlate with stress: where *u* is stressed, it is preserved as a vowel, and the initial *a* is not reduced; where final *i* is stressed, \***ku** > **k<sup>w</sup>**, and \***a** > **e**.

An interesting feature of certain Formosan languages is the obligatory prefixation of *-aku* in the expression of first person singular pronouns, leading Blust (1995) to reconstruct Proto-Atayal \***aku** ‘I’. This feature is more widespread, as shown by: ATY *s-aku* ‘I’, Saisiyat, Thao *y-aku* ‘I, 1 singular nominative’, Amis *k-ako* ‘I’, and

RUK *ko-ako* 'I'. The bound status of this morpheme, then, may be a shared feature of Formosan languages and PON.

- [3.4] PAO \***ala** 'fetch, get, take': PON \***-le** 'verbalizing suffix' (JAR **-le**, e.g., **-ipo/le** 'remove the bark', *ipo* 'flesh, skin'; ONG **-le**, e.g. **-in/le** 'fetch water', *ije* 'water'); PAN \***ala** 'fetch, get, take'.

Final \*a is often reduced to \*e in PON.

- [3.5] PAO \***aNak** 'child, offspring': PON \***ale** 'child' (JAR **-ale** 'child' [Senkuttvan 2000:30]; ONG **ale** 'child [term of address]', **-aye** 'child, step-child'); PAN \***aNak** 'child, offspring'.

Final \*a is often reduced to \*e in PON.

- [3.6] PAO \***aNiC** 'skin': PON \***atiy** (< \***ayit** via metathesis) (JAR **-atiye**; ONG **-ati**); PAN \***q/aNiC** 'skin' (Blust 1999a).

The JAR and ONG forms refer to skin of people and animals, as the PAN term does. For further discussion of vowel-initial bodypart terms in PAO, see 3.2.1. Recall that forms with PAN/PON initial \*q/θ correspondences are assumed to result from \*q-epenthesis in pre-PAN.

If this set is valid, it is notable, as PAN \*qaNiC is replaced by \*kuliit 'skin' in PMP (Blust 1999a).

- [3.7] PAO \***-an** 'locative suffix': PON \***-an** 'locative suffix' (ONG **-a** 'nominal locative suffix', **-an/ka** < \**an-ka* 'locative directional postposition'); PAN \***-an** 'verbal suffix marking locative focus; nominal locative suffix'.

- [3.8] PAO \***apa** 'carry on back': PON \***-apa** (JAR **-apa-**; ONG **-aba-** 'carry on shoulder'); PAN \***apa** (Blust 1999b).

- [3.9] PAO \***aqajaw** 'sun, sunlight, day': PON \***ekeyaw** (JAR **eheya** 'sun, daylight'; ONG **ekuwe** 'day, today'); PAN \***qajaw** 'day'.

JAR appears to preserve a reflex of \*y, with loss of final \*w; while ONG loses \*y and preserves final \*w, with final regular e-epenthesis.

Compare PON \***eke** 'sun' (JAR **ehe**; ONG **eke**), as well as the /aqa/ substring of PAN \***daqaNi** 'noon; day' and PMP \***banagaR** 'radiance, as of rising sun'. PAN loses the initial vowel.

- [3.10] PAO \***aRi** 'come, go (movement toward speaker)': PON \***-ale-** 'come; go (movement toward speaker)' (JAR **-ale/ma** 'come; go'; ONG **ale/ma-** 'return'); PAN \***aRi** 'come!; let's go!; come on!; bring'.

Vowel lowering of \*i > e in PON may be the result of lowering before the imperative \*-ʔ. The formative /ma in JAR and ONG may be a fossilized pronoun \*ma 'to me' (cf. ONG *ma* (< *mi-a*) 'my, me'.) An interesting aspect of this form is the stem-final glottal stop in Onge, because this glottalization is rare and, in some cases, is clearly associated with imperative forms, diachronically or synchronically.

- [3.11] PAO \***ati-a** 'our, us; 1pl possessive pronoun: PON \***eti-a** (ONG **eta** 'our, us' from /**eti-a/**, **eti** 'we'); PAN \***ata** 'I pl. inclusive poss. pronoun: our' (cf. PAN \***ita** 'we [inclusive]).<sup>18</sup>

PON \*a > e may be triggered by i in following syllable.

- [3.12] PAO \***-aya** 'mother': PON \***-aya** 'mother' (JAR **aya** 'mother' [RS:36], **k/aya** 'mother [address form]', **w/aya** 'mother [reference term]'; ONG **k/aye/ri** 'mother [address form]'); PAN \***aya** 'father's sister; \*\*mother'.

JAR *kaya* was likely borrowed from ONG, as the expected form is \*\**haya*.

Though Blust (1995) does not reconstruct PAN \*aya as 'mother', his data include the following terms for 'mother': ATY *aia* (address form); Sadong *aya-*; Kembangan *oya* 'mother'; Dampelas *aya* (address form); Arguni *yai*. In this case, if PON and PAN are sisters, PAN \*aya should be reconstructed as meaning 'mother' as well.

18. Kumar (2006) has no number distinction in first person, with *mi* 'nominative' and *ma* 'oblique'. Abbi (2006:60) lists Jarawa *nao* '1 PL inclusive', but shows no other first plural forms.

- [3.13] PAO **\*baqeRuh** ‘new’: PON **\*baro** (ONG **baro**, **-i/baro** ‘good, nice, beautiful’); PAN **\*baqeRuh** ‘new’.

The PON form shows syncope of medial \*e, and subsequent cluster reduction of \*qR > \*R, closed syllable lowering of \*u > \*o, and final \*h-loss.

- [3.14] PAO **\*bel** ‘smoke’: PON **\*bel** ‘smoke’ (ONG **bele/me** ‘smoke [of fire]’); PAN **\*qe/bel** ‘smoke’.

ONG *-me* appears to be a stative suffix (see 3.2.3).

- [3.15] PAO **\*beRay** ‘give’: PON **\*bele** (ONG **bele** ‘give’); PAN **\*beRay** ‘give’.  
 [3.16] PAO **\*biaC** ‘bow; draw a bow’: PON **\*iya** (JAR **eya/ya** ‘make bows and arrows’, **iya** ‘hunt [with bow and arrow]’; ONG **iya** ‘bow’); PAN **\*biaC** ‘draw a bow to its full extent’.  
 [3.17] PAO **\*biran** ‘anger, angry’: PON **\*iran** ‘angry’ (JAR **-ero-** [AA]; ONG **irage/bit** ‘angry’, **-biti** ‘bad’); PAN **\*biran** ‘anger, angry’.

JAR /o/ is unexplained.

- [3.18a] PAO **\*buaq** ‘fruit’: PON **\*wa** ‘fruit’ (JAR **ele/wa** ‘fruit’; ONG **wa** ‘fruit, pl.’ [EC]); PAN **\*buaq** ‘fruit’.

The first formative in JAR *ele/wa* is likely from PON \*ele ‘leaf’ (cf. ONG *gele* ‘leaf’).

This PAO form appears to have given rise to a doublet (see [3.18b]), perhaps under different prosodic conditions. In both forms, word-initial \*bu > \*u in PON; however, in the word for ‘fruit’, the final C is lost, while it is maintained in the word for ‘flower’.

- [3.18b] PAO **\*buaq** ‘fruit’: PON **\*okw** ‘flower’ (JAR **oha** ‘flower’, ONG **okw/ottireye** ‘flower’); PAN **\*buaq** ‘fruit’.

ONG *okw/ottireye* appears to contain two words for ‘flower’. Cf. ONG *otire*, *otikele* ‘flower’, JAR *otikule* ‘petals (of flower)’.

- [3.19] PAO **\*bubu** ‘conical bamboo basket trap for fish’: PON **\*u/bubu**, **\*-ubu** (JAR **ububu** ‘fish-keeping vessel’;<sup>19</sup> ONG **ubu-dage** ‘bucket of giant bamboo’ [P:206-07],<sup>20</sup> where *-dage* is ‘trunk; stick’); PAN **\*bubu** ‘conical bamboo basket trap for fish’.

This set provides internal evidence for word-initial \*b-loss: initial \*b is preserved when a copy vowel precedes, but is lost in ONG when \*bu is word-initial.

- [3.20] PAO **\*buhet** ‘squirrel’: PON **\*uhe** (JAR **uye** ‘squirrel; rat’); PAN **\*buhet** ‘squirrel’.

This Jarawa term ‘squirrel’ likely refers to the palm civet or jungle cat (*Paradoxurus andamensis*), which is common in the Andaman jungles and, according to Man (1883:127), a common food source.

- [3.21] PAO **\*bukeS** ‘hair’: PON **\*ukec/ele** ‘tail of animal’ (ONG **-ukice/le** ‘tail of animal’, *-le* ‘plural suffix’); PAN **\*bukeS** ‘head hair’.

ONG \*e > i may be due to following palatal. The suffix in PON allows final consonant to be preserved.

- [3.22] PAO **\*bukij** ‘forested area, with many trees’: PON **\*ukiy** ‘area with many trees’ (ONG **ukye** ‘grove’ < **\*ukiye**); PAN **\*bukij** ‘mountain; forested inland mountain areas’.

- [3.23] PAO **\*buluq** ‘type of plant with slender stem for use as small poles, arrow shafts, and general construction’: PON **\*uluk<sup>w</sup>**, **\*ulu** (JAR **uluhe** ‘to weave, make net; skeletal cane structure of conical basket’; **ulu/eye** ‘to weave with cane’, *-eye* ‘to take’; ONG **-uluk<sup>w</sup>e** ‘make mat’, **uluk<sup>w</sup>/ene** ‘common cane’ [P:226-27]); PAN **\*buluq** ‘bamboo; *Schizostachyum* spp.’.

The range of glosses for modern reflexes of PAN \*buluq indicate its common use in building, and/or manufacture of common objects. The semantic match then is in the functional features of the plant, as suggested by the PAO gloss.

19. Senkuttuvan (2001:38) writes this as *u:vu’vu*, where *v* represents a voiced labiodental fricative.

20. Portman (1887:206-07) writes this word with initial *k-*, which is likely his writing of the third-singular possessive *g-*. This is not an isolated example; many possessed nouns are written by Portman with their possessive prefixes, e.g., *mobedupe* ‘finger nail’ (pp.52-53), where *m-* is the 1st singular possessive prefix.



- [3.24] PAO \***Cekelu** 'beckon to come back': PON \*-**e/jegulu** (ONG -**ejegulu** 'beckon to come'); PAN \*\***Cekelu** 'beckon to come' (PAN \***Cikel** 'return, go back', PAI **tsikelu** 'come back!', PAZ **ma-/sekela** 'to come across').

Expected ONG form is \*\**e/jegelu*, with initial copy vowel, and intervocalic voicing. Harmony between final and penult may be due to its weak prosodic status.

- [3.25a] PAO \***Cenek** 'hard exterior; sharp exterior': PON \***cenek** (ONG -**uku/cenege, uku/cene** 'heel', -**uku** 'node, protuberance'); PAN \***Cene/k** 'thorn, spine'. (See [3.25b] below for potential stem.)
- [3.25b] PAO \***Cana** 'hard exterior; sharp exterior': PON \***cana** 'shell; hard exterior' (JAR **cana/cew** 'topshell, *Trochus niloticus*', **cew** 'beautiful' [MS:98]; **cana/na** 'snail', **cana/hanap** 'hard-soiled, plain ground'; ONG **cena/gili** 'conch shell', **gili** 'smooth'; **cena/dalu** 'dentalium shell', **dalu** 'big'; -**uku/cenege, uku/cene** 'heel', -**uku** 'node, protuberance'); PAN \***Cene/k** 'thorn, spine' (see [3.25a] above).
- [3.26] PAO \***Cin** 'offal; entrails and internal organs; blood and guts': PON \*-**a/ciŋ** 'kidney' (ONG -**acige**); PAN \***Cin/aqi** 'guts'.

The PON form may be contaminated by \*-a/ce 'blood' in terms of its pre-root vowel.

Blust (1995) suggests PAN \*C-in-aqi, from \*Caqi 'feces, excrement'. An alternative is a PAN formative \*-aqi relating to lower reaches of the body.

- [3.27] PAO \*-**Cu** 'that, there': PON \*-**cu**- (ONG **li/cu/ne** 'here', **la/cu/ne** 'there'); PAN \*-**Cu** 'deixis and spatio-temporal reference: that; there, then'.
- [3.28] PAO \***daŋ** 'bone': PON \*-**daŋ** (JAR -**i/dəŋ** 'bone', -**alata-gi-taŋ** 'collar bone', cf. -**alata** 'shoulder joint'; ONG -**i/dage** 'bone', -**atale/rege** 'collar bone', cf. -**atale** 'shoulders'); PAN \***wa/daŋ** 'collarbone; clavicle' (cf. PAN \***Cuqe/laŋ** 'bone').

It is possible that a doublet \*-**daŋ, -taŋ** 'bone; trunk, log' should be posited for PON. JAR shows voiced and voiceless variants (though see footnote 11). Compare forms above with JAR -**ata** 'shoulder blade', -**ipita** 'chin' (cf. ONG -**ibirae** 'jaw'). ONG has initial *d* or *r*; never *t*, supporting the PON reconstruction above. Whether or not a lenited form of this stem occurs in PAN \***Cuqe/laŋ** 'bone' must await further comparison.

Many words for specific bones in JAR and ONG have been recorded. This is likely a consequence of their cultural significance. After the mourning period is over, bones of the dead are dug up, cleaned, and made into ornaments worn by the living (Man 1883:77–78).

- [3.29] PAO \***daqu** 'ko hardwood tree': PON \***daku** (JAR **dau** 'k.o. tree, *Pterospermum acerifolium*' leaves used for mats; women sit on leaves during menstruation [DTW:62]; ONG **dagu/lee** 'k.o. tree, casuarina', **lee** 'knife' [D&S:117]); PAN \***daqu** 'a tree, *Dracontomelum edule*'.

The expected JAR form is \*\**dahu*.

- [3.30] PAO \***daReq, taReq** 'soil, clay, earth': PON \***dare** 'ko spirit'; \***tale** 'earth, soil' (JAR **to/tale** 'landmass, earth' [MS:23], *to-* definite/relative marker 'the one that'; ONG **to/tan/nape** 'soil', *to-* definite/relative marker [D&S:116], **dare** 'class of spirits who send the south west monsoon [D&S:118]; a class of spirits living below the earth [D&S:112]';<sup>21</sup> PAN \***daReq** 'earth, land, soil'; PMP \***taneq**).

No spirit terms have been recorded for Jarawa. In ONG, *dare-* alternates with *nare-* in words with following nasals, and we find also **nare/cekwe** 'sun', **cekwe** 'hop', which could have the literal meaning 'earth-hopper'.

The PAO doublet is supported by \*/\*d doublets in PON, and by comparison of PAN and PMP with PON.

- [3.31] PAO \***durdur** 'noise of thunder': PON \***durur** (ONG **dururu** 'peal of thunder; cry of varanus [k.o. lizard]'); PAN \***duRduR** 'thunder'.

ONG shows final vowel copy instead of *e*-insertion.

- [3.32] PAO \***eleb** 'knee': PON \***olo** (< \***olob**) 'knee' (JAR -**olo** [RS:40]; ONG -**olo-**); PAN \***eleb** (cf. doublet \***qaleb**).

21. On this second meaning, D&S write *taore*; I assume this is the same word.

For further discussion of vowel-initial bodypart terms in PON, see 3.2.1. PON shows \**eb.* > \**ob.*, followed by rounding harmony (see section 2).

- [3.33] PAO \**esa* ‘one’: PON \**eta* ‘and, also; one; another’, \**-ta* ‘singular suffix’; (ONG *eta* ‘and, also’; *aʔy/eta* ‘again’, with *-aʔy-* ‘more’; *-da* ‘singular suffix’); PAN \**esa*, \**isa* ‘one’.

PON has \**t* instead of expected \**c*.

- [3.34] PAO \**i+dalam* ‘inside’: PON \**i-dalam* ‘inside’ (JAR *itala*, *italaŋ*; ONG *-iram-*, in *iram-angka* ‘inside’ with *-angka* locative ‘at, by’ [D&S:21]); PAN \**i-dalem*.

In ONG, proposed developments are: \**iram* > *irlam* > *irram* > *iram*. In JAR, the final nasal is reduced to *ŋ*, or lost in complex bodyparts, as in *ecepo-tala* ‘pupil’ (lit. ‘eye-inside’). See 3.2.1 for further discussion of complex bodypart terms and this PON postposition.

- [3.35] PAO *ijan* ‘nose’: PON \**ijnjan* (JAR *ijnjan/po*, *ijnjam/po* ‘nose’; ONG *-ijnabo* ‘nose’ [P:53], both with reflexes of \**ipo* ‘flesh’); PAN \**ijuŋ*, \**ujuŋ* (Blust 1995), \**m/ujin* (Blust 1999a) (cf. PWMP \**e(n)juŋ*, \**ijuŋ*, \**ujuŋ*, \**gjuŋ*).

PON \**C* > *NC* / *\_VN* occurs in JAR and ONG as a synchronic process. The shift of final PAO \**n* to \**ŋ* in PAN, and PAN \**u*-vocalism are unexplained. Cf. also PAI *idjiam* ‘point’.

- [3.36] PAO \**isiʔ* ‘flesh, contents, foundation (of humans, animals, fruits, tubers)’: PON \**ici* ‘(JAR *ici/tag* ‘roots of tree, plant’, *tag* ‘stem, trunk’ [MS:21], *-ici/ge* ‘leg; hindleg [of pig]’, *-uge* ‘foot, leg’; ONG *-ici/ge* ‘leg; lower leg’, *ici/bbo* ‘calf muscle’); PAN \**isiʔ* ‘flesh (of humans, animals, fruits, tubers); contents; blade of a knife; inhabitants, residents’ (cf. Tsou *rm/isi* ‘root’; Kalamian Tagbanwa *isi* ‘meat, flesh; tuber’).

- [3.37] PAO \**kwet* ‘tie, bind’: PON \**-a/ikwe-* (JAR *-a/ih\*/le-* ‘tie, fasten’; ONG *-a/ikwe* ‘bind, tie, fasten’); PAN \**Si/ket* ‘tie, bind, attach to by tying’, PAN \*\**ket* (cf. ATY *ket* ‘plait, wind, coil’).

The accreted /*a*/ in PON appears on other verbs with locative or applicative function. Loss of final *C* suggests that this verb was already disyllabic in pre-PON. The Atayal form confirms a PAN/PAO monosyllabic root.

- [3.38] PAO \**iseq* ‘urine’: PON \**ecce/le* ‘urinate’ (JAR *-ecce/le* ‘urinate’; ONG *-eje/le* ‘urinate’); PAN \**iSeq*, \**iSiq* ‘urine, urinate’.

Recall that *-le* in Jarawa and Onge is a suffix meaning ‘to make, to get’, so the verb was historically ‘to make urine’. The shift of \**i...e* > \**e...e* in PON appears to be regular.

- [3.39] PAO \**itaaq*, \*(*h*)*ataaq* ‘eat; chew’: PON \**-ita*, \**-ata* (JAR *-ita*; ONG *-ida-*, *-ada-*; *-ite* ‘bite [by snake]’); PAN \**q/etaaq*, \*(*h*)*ataaq* ‘eat something raw’ (cf. PAZ *-ata* ‘chew’).

A striking aspect of this etymology is the correspondence between the PON and PAN doublet. Lowering of \**i* in PAN \**qetaaq* in unaccounted for.

- [3.40] PAO \**jem* ‘cold, chill’: PON \**jom* (ONG *-jome* ‘be cold’, *-ajomo* ‘cool’); PAN \**da/jem* ‘chills, as of malaria’.

PON shows \**em* > \**om*. It is unclear whether ONG forms with final *-e* vs. vowel copy are variants.

- [3.41] PAO \**kakaj* ‘k.o. crab’: PON \**kaka* (JAR *haka* ‘crab’ [AA:65]; ONG *kaga* ‘crab’ [P: 23,146]); PAN \**kaganj* ‘crab sp.’.

- [3.42] PAO \**kam* ‘mat’: PON \**kam* (JAR *ham* ‘hanging platform inside hut’ [DTW:34]; ONG *kame* ‘mat of split canes used for sleeping’, *kan/le* ‘make mat’); PAN \**Si/kam* ‘mat’.

Sreenathan (2001:47) notes that, though E. H. Man and others provide early reports of well-made Jarawa sleeping mats, “such mats are no longer in use and they mainly sleep on the earth”. This could explain the restricted meaning cited above.

- [3.43] PAO \**kan* ‘eat’: PON \**kan* ‘food’ (ONG *kane* ‘rice’, *kwa/gane* ‘turtle eggs [lit. ‘creek-food’]’); PAN \**kaen* ‘eat’ (Blust 1999a).

Rice is an introduced food in the Andamans, so the etymology assumes a meaning shift FOOD > RICE. Turtle eggs are a prized food, and are also called, in Onge, *tokala* ‘the white things’.

Compare this set with South Andaman *akana* ‘meal, food’ (Man 1883:123).

- [3.44] PAO \***kahiw** ‘forest, wood, stick’: PON \***kahiw**, \***kayiw** (JAR **həwə** ‘stick’ [AA:69], **kəyə**, **kəyəw**, **gəyə**, **kewa** ‘woods, forest’; ONG **kəyi**, **kəye**, **keye**, **ke**, **gə** ‘harpoon; stick’, **gai/borale** ‘forest’, **gai/aŋka** ‘as some height of tree’); PAN \***kayuh**, \***kaSiw** ‘tree, wood, stick’ (cf. PAZ **kahuy** ‘tree; forest’).

Initial voicing of PON \*k in JAR and ONG variants is unexpected. These may be lenited variants of basic forms with initial *k*.

JAR *k*-initial forms suggest earlier medial \*h, with blocking of (4a), as noted earlier. The variant **həwə** ‘stick’, on the other hand, requires a PON form without \*h.

Blust (1999a) has PAN \*kaSiw ‘wood’. If PON reflects PAN \*S as \*c, as suggested by [3.21] and [3.36], then PON is cognate with PAN \*kayuh (or some other form without medial \*S), not \*kaSiw.

- [3.45] PAO \***laŋaw** ‘housefly’: PON \***noŋaw** (JAR **noŋaw** [DS:98], **unaw** [AA] ‘housefly’; ONG **ŋeno** ‘housefly’, **ŋənoi** [MVP:33]); PAN \***laŋaw** ‘fly (insect)’ (Blust 1999a).

Metathesis has occurred in ONG, and possible place assimilation in JAR. In PON, I assume initial \*n < \*l in assimilation with following nasal.

- [3.46] PAO \***ma** ‘stative (of verbs)’: PON \*-**me** ‘stative verbal suffix’ (JAR **-me**; ONG **-me**); PAN \***ma-** ‘stative verbal prefix’.

See 3.2.3 for further discussion of this affix.

- [3.47] PAO \***ŋuŋu** ‘nurse, suckle’: PON \***ŋoŋo** (JAR **ŋoŋo** ‘to feed milk’; ONG **ŋoŋo** ‘suck mother’s breast’); PAN \***ŋuŋu** ‘breast’ [CAD] (cf. \***nunuh** [Blust 1999a]).

Reconstruction of verbal ‘nurse’ rather than nominal ‘breast’ meaning is based on stronger association between reduplication and repeated events (or intensity) in PAN, but is admittedly weak. Though one might claim association between palatal nasals and sucking sounds, the set is significant in this particular context, because the palatal nasal is rare word-initially in the PAN lexicon and in Jarawa and Onge.

- [3.48] PAO \***paCay** ‘die, death’: PON \***peca** ‘die, death’ (JAR **peca** ‘death’ [RS:43], **pecame** ‘die’; ONG **became**, **became** ‘die’); PAN \***paCay** ‘kill; \*\*die; \*\*death’ (cf. PAN \***aCay** ‘death’, \***ma-aCay** ‘die, dead’).

The PAN reconstruction is from Blust (1995), which shows Kananavu *patay* ‘die, dead’, Amis *patay* ‘death, die’, and Kananavu *mia-paca* ‘kill’. For the Austronesian languages of Taiwan, then, there is good evidence of \***paCay** ‘die, death’, with the shift of meaning to ‘kill’ being primarily a PMP innovation.

PON has \*-me suffix corresponding to PAN stative \*ma- prefix. See 3.2.3 for further discussion of this affix.

- [3.49] PAO \***pegu** ‘body excretion; bodily fluid’: PON \***pegu** (JAR **pegu** ‘stool’; ONG **bege** ‘stool’, **beke/lalaje** ‘urine’); PAN \*\***peju** ‘bile, gall, gall bladder’ (PAN \***qa/peju** ‘bile, gall, gall bladder’; Saisiyat **pa?zo?**; Proto-Rukai \***pago**.)

ONG has undergone final vowel reduction. Cf. also JAR *pecu* ‘flatulence; farting’ [MS:31].

- [3.50] PAO \***qala** ‘shining, shimmering, dazzling’: PON \***kala** ‘white’ (JAR **hala**; ONG **kala**); PAN \***qala-**, \***qali-** ‘spirit, spirit world’ (Blust 2001); PAI **qal’aqal’** ‘glittering, dazzling’.

See Blust (2001) for extensive evidence that PAN \***qala-**, \***qali-** is associated with the spirit world, including Austronesian spirit names beginning with reflexes of this prefix. He does not highlight connections with white or luminescent creatures, but this is evident in his comparanda where \*qala-, \*qali- is found in words for iridescent beetles, fireflies, luminous millipedes, and termites.

- [3.51] PAO \***qeCuR** ‘sexual arousal, sexual excitement’: PON \***ketol** (ONG **ketolo** ‘copulate’); PAN \***qeCuR** ‘sexual arousal, sexual excitement’ (cf. Amis **?etol** ‘sexual arousal, sexual passion’).

Expected PON form is \*\*kecol, with medial palatal.

- [3.52] PAO \***qidum** ‘dark, black’: PON \***kiru** ‘black’ (JAR **hiru**; ONG **kiu**); PAN \***qudem** ‘dark’ (with metathesis); PMP \***qitem** ‘black’.

ONG **giru** ‘brown’ may be related.

- [3.53] PAO \***quai** ‘rattan’: PON \***kwai**, \***kwe** ‘cane’ (JAR **da/ai/g** ‘cane basket’; ONG **kwege** ‘cane; rope’, **da/kwege** ‘k.o. cane (slender stem)’); PAN \***quay** ‘rattan’.

PON \*kw is lost in JAR. Ongan final -g may be a reduced form of the definite suffix \*-gi; *da/ta-* is ‘the one that...’. Cf. [2.28].

- [3.54] PAO \***sakaya** ‘walk, go (imperative?)’: PON \***cakaya** ‘walk’ (JAR **cawaya** ‘walk’ [PK]); PAN \***sakay** ‘walk, go’; PAZ **zakayi** ‘walk!’.

JAR shows irregular lenition of PON \*k > (g) > w. A final vowel must be posited for PAO, else final \*ay would surface as PON \*e.

If this set is valid, it is notable, because PAN \*sakay is replaced by \*lakaw/panaw ‘go, walk’ in PMP (Blust 1999a).

- [3.55] PAO \***Sada** ‘house’: PON \***cada** ‘house’ (JAR **cada** ‘house’); PAN \***Sad/iRi** ‘house post’.

I have no suggestions for the possible meaning of PAN \*iRi. The only PAN reconstruction with this string is \*wiRi ‘left side or direction’.

- [3.56] PAO \***Sulij** ‘lie or sleep next to’: PON \***tule** ‘lie down, sleep’ (JAR **tule** ‘sleep, lie down’ [PK]; ONG **tule** ‘lie down’ [D&S:115]); PAN \***Sulij** ‘lie or sleep next to’.

Lowering of PAO \*i in closed syllable. PAN \*S corresponds to PON \*c, except before \*u.

- [3.57] PAO \***tuk** ‘strike (as flint for a fire)’, \***tuki** ‘kindle fire; fire’: PON \***tuke** ‘fire’ (JAR **tuhe**, **ittuge** ‘torch’, ONG **tuke**); PAN \***tuk** ‘knock, pound, beat, \*\*strike’; PAN \*\***tuki** ‘kindle fire’ (ATY **atuk** ‘chipping knife’, **ta/tuki/an** ‘match’ [CAD]; PAZ **pa/tuk** ‘to strike against’, **pa/tuki** ‘flint’ [used to kindle fire]); Proto-Agta \***t/əm/uk** ‘burn’, \***ti/duk** ‘burn’, **duku** ‘fire’ [Reid 1994]). See 4.3 for further discussion.

The PON final vowel must be present, else we expect JAR \*tuk > tok > toh > \*\*toha. PAN \*tuk follows Blust (1988), with addition of primary meaning ‘strike’, while \*\*tuki ‘kindle fire’ is new.

PAI *pa-tseki* ‘make fire with flint and metal’, *pa-p-edjek* ‘start fire’, *edjek* ‘flame’, if cognate, appear to be borrowed, because the expected reflex of PAN \*t is PAI *tj*. Blust (1995) has \*tuqtuq ‘hammer, pound, crush’, based on PAI *tjuqtjuq* ‘blacksmith’s hammer’, and forms derived from this, but Formosan forms above suggest PAN \*tuk.

Understanding the etymology of this word could reveal a great deal about the prehistory of PAN and PON speakers. If these are cognate, speakers of PON knew how to make fire with flints, or had contact with people who did. This contrasts with common fire-making by friction in Oceania (Ross, Pawley, and Osmond 1998:144–45), terms for which PAN initial \*s seems likely: cf. PAI *patut* ‘drill for firemaking’, *ph/atut* ‘make fire by friction’; PMP \*usuq ‘rub, wipe’; POC \*suka, \*suka-i ‘make fire with fire plough’.

- [3.58] PAO \***uNay** ‘pierced by thorn, get splinter’: PON \*-**ulay-** (ONG **-ull/uke** ‘be pierced by thorn in the foot’, *-uke*, *-uge* ‘foot’); PAN \***uNay** ‘sliver, splinter, foreign body embedded in the flesh’.

Geminate *ll* in Onge is attributed to syncope and assimilation: \*uleyuke > *ulyuke* > *ulluke*.

- [3.59] PAO \***uma** ‘other; more; again’: PON \*-**uma** ‘more, big’ (ONG **-uma/me** ‘grow up, become big’); PAN \***uma/n** ‘add; repeat; again’; PAI **-uma-** in **z-uma** ‘other (persons, things)’, **m-uma-ly** ‘to repeat, do again; to change, renew’; **ma-uma-ly** ‘to be (come) changed’.

- [3.60] PAO \***uqay**, \***quqay** ‘leg, foot’: PON \***uke** ‘foot, feet’ (JAR **-uge**; ONG **-uge**, **-uke**); PON \***kuke** ‘wing, fin’ (JAR **huke** ‘wings [of bird]; lateral fins [of fish]’); PAN \***qaqay**, \***waqay** ‘leg, foot’.

For further discussion of vowel-initial bodypart terms in PAO, see 3.2.1. The two PAO forms appear to have split in PON, with unusual consistent voicing in JAR *-uge* ‘foot’, perhaps in contrast to the term for ‘wing’ or ‘fin’. In JAR, where PON \*k, \*k<sup>w</sup> >

*h, hʷ*, there appears to have been a constraint that only one *h* can occur per word, blocking multiple instances of the sound change.

- [3.61a] PAO \***uRaC-u** ‘blood organ; liver’: PON \*-**orac** ‘liver’ (ONG **-o-acu**); cf. PAN \***uRaC** ‘artery, blood vessel, vein; muscle; nerve; sinew; tendon’ in [3.61b] below.

ONG \**r*-loss is found elsewhere, e.g., ONG *kiu* ‘black’ < \**kiru*; the vowel cluster provides evidence of the former intervening consonant.

- [3.61b] PAO \*(**q**)**uRaC** ‘artery, blood vessel, vein’: PON \***kuat/o** ‘blood vessel, vein, nerve’ (JAR **hʷeto**; ONG **kʷeto**); PAN \***uRaC** ‘artery, blood vessel, vein; muscle; nerve; sinew; tendon’.

This is one of the few bodypart terms where PON shows an initial \**q* absent in PAN. As in [3.61a], there is medial \**r*-loss, but here it can be traced back to PON. Reflex of \**C* is irregular PON \**t*, instead of expected \*\**c*.

- [3.62] PAO \***izaN** ‘rain’: PON \***ucen** ‘rain’ (ONG **ujeŋe, uŋe** ‘rain’); PAN \***q/izaN**.

This form is irregular. PAN \**N* should match PON \**l* (or be lost finally). This irregularity might justify putting it into class IV.B. However, as this is the only form with final \**N*, the PON nasal could be a conditioned reflex.

- [3.63] PAO \***wakwak** ‘to caw; crow’: PON \***wawa** ‘to caw; crow’ (ONG **wawa** ‘crow; cry of crow’); PAN \***wakwak** ‘bird sp.; to caw; crow’.

Although bird names of this sort are poor comparanda as they may be onomatopoeic, this set is notable, because \**w*-initial reconstructions are rare in PAN, and *w*-initial words appear to be rare in ONG as well.

- [3.64] PAO \***zalan** ‘path, trail’: PON \***i/cala** (JAR **-i/cʰala** ‘forest path, trail’; ONG **-i/cele** ‘path, way’); PAN \***zalan** ‘path, road’ (Blust 1999b).

Vowel raising in ONG is unexplained, possibly conditioned by initial palatal.

## CLASS II. PMP AND FORMOSAN COGNATES, WITH SOME NEW PAN FORMS

- [3.65] PAO \***alaŋ** ‘mouth’: PON \***alaŋ** (ONG **-alaje** ‘mouth’, **-alaj/ete** ‘yawn’, **-alan/daje** ‘tongue’; JAR **-ala** ‘lips’ [AA:67]); PAN \*\***aŋaN** (with consonant metathesis) (PAI **aŋan, aŋalʷ** ‘mouth’, **m-aŋalʷ** ‘to open one’s mouth’, **ru-aŋal-an** ‘chatterbox’; PMP \***r/aŋa** ‘open the mouth wide’).

Cf. also PAZ *rahal, rahan* ‘mouth’. PAN \**ŋ* continues as PAZ *ŋ*, so this form, if cognate, shows irregularities. Cf. also [3.6] ‘skin’, where metathesis also occurs.

- [3.66] PAO \***alew** ‘honey, sweet food’: PON \***lew** (JAR **lew** ‘honey [from white honeycomb]’; ONG **le** ‘honey’); PAN \*\***alu** (PAI **alʷu** ‘honey, candy, sweets’; PAZ **w/alu** ‘honeybee’; PMP \***w/ani** ‘honeybee’).

- [3.67] PAO \***betiew** ‘taut, springy; bowstring’: PON \***betiew** ‘bowstring’ (JAR **weito**; ONG **becuwe** ‘bowstring’); PAN \*\***betiu-** (PAI **vetjius** ‘to pull forcefully on something [as string or wire], Western PAI **vetjun** ‘springy [trap]; catapult’ [cf. PAN \***beties** ‘calf of the leg’]; PMP \***bentuk** ‘curve’ [cf. Tiruray **betuk** ‘bulging; to set a bowstring’]).

- [3.68] PAO \***beR** ‘mud, murky water’: PON \***bela** ‘mud, dirt’ (JAR **bela** ‘mud, soil, land’;<sup>22</sup> ONG **bela** ‘sand; shore’); PAN \*\***beR** ‘turbid, murky, muddy’ (cf. PAN \***buR**, possible later form \***beR**, ‘turbid, murky’ [Blust 1988:89]; PAI **ber/eq** ‘watery mud’; PEO \***bʷela** ‘mud, swamp, taro swamp’).

- [3.69] PAO \***bukul** ‘lump, bump, hump; back (of body); shell of turtle’ (cf. \***buku**); PON \***ukel, \*bukel** ‘turtle’ (JAR **ukela** ‘turtle’; ONG **gi/bukece** ‘catch turtle’); PAN \*\***bukul** ‘lump, bump; back (of body)’ (PAZ **bukul, bukun** ‘back [of body]; shell of turtle’, **ti/bukul** ‘hunchback’); PMP \***bu(ŋ)kul** ‘hump (including on back of some animals),

22. Kumar (2005) writes Jarawa *bela, pela* ‘land, soil’, while Senkuttuvan (2001) has initial /b/, with meanings including ‘mud’ and ‘ball made of mud’. Compare Abbi (2006) *bela* ‘dig’. There seems to be a different *p*-initial lexeme, with closely related meaning: *pila, pile* ‘sand, mud, soil’, *pile-he* ‘sandy shore’ (Sreenathan 2001:24). If one reconstructs PON \**pela* instead of \**bela*, the PON/PAN initial \**p/b* correspondence is irregular.

lump, bump, swelling'; PWMP **\*bukel** 'lump, bump, swelling' (cf. also Kalinga Limos *bukol* 'round', ISN *bokál* 'round', *bokálo* 'round gastropods that cleave to stones').

The semantic shift here is from 'turtle shell' > 'turtle'. It is unclear whether Kalinga Limos *bukol* 'round', ISN *bokál* 'round', *bokálo* 'round gastropods that cleave to stones' belong in this set.

Turtle hunting was, and continues to be, an important part of Andamanese life, with turtle hunting a common activity for coast-dwellers, turtle meat and eggs a favorite food, and turtle meat taboos during initiation (Man 1887; Sreenathan 2001).

In JAR, the expected reflex of \*k is /h/, marking this as a potential loan. In ONG, *gi-* is 3rd singular prefix. The expected ONG verb however is **\*\*gi/bukele**, **\*\*gi/bukelle**, not attested *gi/bukece*. Compare South Andaman *-bukara* in *yadi-bukura* 'the larger of two or more large green turtles' (*yadi* 'green turtle'). *tau-bukura* 'the larger of two or more large Hawk's bill turtles' (*tau* Hawk's bill turtle) (Man 1923:141).

- [3.70] PAO **\*buqu** 'bamboo node, used as container/conduit for liquids; container for liquids; knot in wood or rope; joint, knuckle': PON **\*uku** (JAR **uhu** 'wooden basket, made from a soft-wood tree, used for honey collection'; ONG **uku** 'wooden bucket'; **-uku-**, in **-uku-tamandage** 'ankle', **uku-cene** 'heel'); PAN **\*\*buqu** 'node of bamboo (used for collecting or drawing liquid)' (PAZ **bukubuk** 'bamboo pipe for drawing water; jar, bottle'; PAI **buqu**, **buqubuqu** 'a protuberance; ankle, knuckles'; PMP **\*buku** 'node [as in bamboo, or sugarcane]; joint; knuckle; knot in wood or rope').

This set is striking for the range of semantics continued in daughter languages: the functional status of the object (container for liquids), along with the association with joints, including the ankle.

- [3.71] PAO **\*buliq** 'cowrie shell; shells, stones, or other weights used as sinkers for fishnets': PON **\*uli** 'stone' (JAR **uli** 'stone', ONG **uli** 'stone'); PAN **\*\*buliq** 'cowrie shell; shells, stones, or other weights used as sinkers for fishnets' (ATY **baliq** 'iron, bullet'; PAI **vuliq** 'a black or white stripe around neck of animal of opposite color', **vuli-vuliq-an** 'a black animal with white neck; white animal with black neck'; PMP **\*buliq** 'cowrie shell').

The PAN reconstruction is based on Paiwan forms that appear to refer to the dark/light patterns of cowrie shells, and ATY **baliq** 'iron, bullet'.

Blust (1995) reconstructs PMP **\*buliq** 'cowrie shell', without the meaning 'weight, sinker', though some languages show only this meaning, e.g., ISN, Ilokano **buli** 'lead; lead sinker of a fishing net', Bontok **boli** 'lead (metal)', while others maintain both meanings, e.g., Roti **fuli** 'kind of shell; shell or bits of lead used as sinkers for a fishnet'. Also supporting the 'stone (as weight)' meaning is ISN **buli/nawa** 'a very hard black pebble, the usual anvil of blacksmiths' (*anawa* 'fishing rights').

Man (1883:146) remarks on the use of large fishing nets by the Andamanese, "one side sunk by means of stones, and the other... kept up by floating sticks".

- [3.72] PAO **\*cilaq** 'light emanating from heavenly bodies': PON **\*cila** (JAR **cila/he**, **cila/be** 'star; moon', ONG **cile/me** 'moon; candle; white beeswax'); PAN **\*\*cilaq** (PAI **tselalaq** 'lightning bolt'; PMP **\*sila(q)** 'outpouring of light', PMP **\*sila(k)**, **\*cilak** 'to shine [of heavenly bodies]').

JAR *-he* refers to places: *pile* 'sand', *pilehe* 'sandy shore' (Sreenathan 2001:24). ONG *-me* appears to be a stative suffix.

- [3.73] PAO **\*dag** 'coconut, coconut shell container': PON **\*dag** 'coconut' (JAR **dag** 'coconut'; ONG **dage** 'coconut' [P:23]); PAN **\*\*dag** (POC **\*d[r]ag[i]** 'k.o. container' [?] [LPO1:75] [Yapese **ða:Y** 'coconut-shell container' (Ross 1996); PNNG **\*dag** 'wooden bowl' (Ross 1996)).

Cf. also Proto-North Agta **\*tug** 'ladle of coconut shell' (Reid 1994), as well as PAN **\*-Duk** 'ladle, spoon' (Blust 1988), PPH **\*sa/duk** 'coconut shell ladle' (Reid 1994).

The Jarawa and Onge did not use coconuts as a food source prior to recent contact with Indian officials, but apparently did use parts of the (wild) plant for various imple-

ments, including cups and perhaps brooms. Portman (1887:19) transcribes Onge 'broom' as *dage*, the same word he writes for 'coconut'. It is common in Jarawa and Onge for implements made from a particular plant to be called by the plant name (Sreenathan 2001).

- [3.74] PAO \***ebej** 'wind around, tie up, wrap up; thing used to wind, tie, wrap': PON \***ebe** 'thread' (ONG **ebe** 'thread'); PAN \*\***ebej**, \***bej** (Blust 1988), \***bejbej** 'wrap around repeatedly, bind by wrapping'; PWMP \***embej** 'bind the waist'.

The WMP forms on which \***embej** is based include Sasak *embet* 'black cord around the middle of newborn or sick person', and Balinese *embed* 'band or ring which firmly closes something'. Note also the nominal reflex of the reduplicated \***bejbej** in PAI *vetvet* 'fancy waist-sash (for dancing)'.

- [3.75] PAO \***gil** 'shiny, smooth': PON \***gili** (ONG **giligilie** 'smooth; action of smoothing and polishing', **cena/gili** 'conch shell'; **wa/gili** 'iron'); PAN \*\***gin**, \*\***giNak** 'shiny', \*\***gil** 'smooth' (PAI **g/al/iF/gil'** 'shining', **ma-gilgil** 'be rubbed against something'; PAZ **lilak** 'shining'; PMP \***gil/ak** 'shine').

Perhaps related to set [3.72].

- [3.76] PAO \***ijncu** 'drink'. PON \***ijncu**- (JAR **-ijncu**-; ONG **-ijncu**-, **-ijncu**-); PAN \*\***ijnu** (Tsou **-imo**; RUK **wa-ogolo**; PMP \***inum**).

PMP \***inum** may be reformed on the basis of \***dalum** 'water' (< PAN \***daNum**). Given the reconstruction of PAO \***ijncu** 'water' it is possible to view PAO \***ijncu** as morphologically complex.

- [3.77] PAO \***kwaray** 'rope, string': PON \***kware** (ONG **k<sup>w</sup>ara/ge**, **k<sup>w</sup>ala/ge**, **k<sup>w</sup>e/ge**, **koya/ge** 'rope, string'); PAN \*\***kara**, \*\***waray** (Proto-Atayalic \***waray** 'thread' [Rau 2004]; PAZ **waray** 'thread, hemp yarn for weaving'; PAI **alay**, **kala** [Southern dialect] 'thread'; PMP \***waRej** 'rope, twine; vine, creeper').

If initial PAO \***kw** is reconstructed for PAN, the Paiwan doublet is partially explained.

- [3.78] PAO \***kuai** 'game, wild animal': PON \***kui** 'game, wild animal' (JAR **hwəuə** 'wild pig'; ONG **kue**, **kui** 'wild pig; animal [in compounds]'); PAN \*\***kuai** (Tsou **zua** 'deer'; ITB **kuyi/s**, **koyi/s** 'pig, boar'; Tolai **vəva/guəi** 'domestic animal', **kuə/bər** 'wild animal, wild pig'; Lau **kui** 'animal; dog').

- [3.79] PAO \***lalaq** 'trochus shell, and ornamental necklaces and bracelets made from it': PON \***lele** (JAR **lelele** 'oliva shell; shell necklace'); PAN \*\***lalaq** (RUK **saki/ʔ/ə/θan** 'long necklace'; PMP \***lalak** 'trochus shell'; POC \***lala/k(o)** 'trochus shell; bracelet made from it' [LPO1:103–04]).

- [3.80] PAO \***li** 'proximal, near': PON \***li** (JAR **li**; ONG **li**); PAN \*\***li** (PAN \***za/Ni** 'near'; PAZ **a/lih** 'near', **di/ni** 'here', **mi/ni** 'this'; **uhu/ni** 'now', **li/a** 'already'; PAI **Fia-dut** 'near, nearby', **mal'i-tsa** 'here it is!', **za-zi-tsu** 'this'. **Fia** 'dawn, daybreak').

See 3.2.2 for further discussion.

Of note is the semantic category of proximal/distal time, as opposed to space, in Pazeh and Paiwan. Proximal time, or proximate temporal distance, relates 'now; already' in Pazeh with 'dawn' in Paiwan.

- [3.81] PAO \***lu** 'distal, far': PON \***lu** (JAR **lu**; ONG **lu/a**); PAN \*\***lu** (PMP \***da/iuq** 'far', \***Za/uq + ma-** 'far'; PAZ **a/nu-** 'far', **du/a** 'that, yonder; there'; PAI **ma-lua** 'far apart', **mu/dulu** 'send someone on an errand', **i-zua** 'there, over there', **tsuay** 'long time').

See 3.2.2 for further discussion.

- [3.82] PAO \***nam** 'firewood': PON \***nam** (JAR **nam**, **name** 'firewood, pieces of dry wood'; ONG **name** 'firewood; make fire'); PAN \*\***namaR** (Kanakanavu **namar-** 'to burn [of brush]' [Tsuchida 2003:127, 133]; Central Amis **namar** 'fire' [ABVD]; ITB **namay** 'fire, house fire'; Tetum **namot/tuk** 'burn' [ABVD]; Manobo **namu** 'a smoky torch used to smoke and subdue nest of bees' [Elkins 1968:126]).

- [3.83] PAO \***nape** ‘fish’: PON \***napo** (JAR **napo** ‘fish’); PAN \*\***qunap** ‘scales (of fish)’; PAI **qa/napi** ‘k.o. fish, *Flussfisch*’, **qu/rip** ‘scales of fish’; RUK **ka|a/labə** ‘fish scale’; PMP \***qu/nap** ‘scales (of fish)’ (Dempwolff 1938) (POC \***qu/nap** ‘fish scales; turtle shell’ [LPO1:166]; POC \***qunap-i** ‘scale fish, vb’ [LPO1:166]).
- [3.84] PAO \***-otab** ‘face, head, top’: PON \***-otab** ‘face, head’ (JAR **-ota**, **-otab**, **-otap** ‘head’, **tape** ‘moon’; ONG **-otabe** ‘head’); PAN \*\***atab** ‘top’ (cf. PAN \***aCab** ‘cover’; Agta **atubag** ‘front’ [Healey 1960:28]).

This set is discussed further in section 3.2.1. The only real change I have made is adding the meaning ‘top’ to PAN \***aCab** ‘cover’. The PON/PAN \***o/a** correspondence in the first syllable is not regular.

- [3.85] PAO \***pelpel** ‘butterfly’: PON \***pelpel** (JAR **pe|pel**; ONG **bebele** [MVP:19]); PAN \*\***pelpel** (Western PAI **qu|ipepe** ‘butterfly sp.’; POC \***bebe**; Roviana **pepele**; Botolan Sambal **palo-palo**; Tagalog **ali/paro** ‘butterfly sp.’).

See Blust (2001) on the difficulty of reconstructing a single term for ‘butterfly’, and its apparent role in the Austronesian spirit world.<sup>23</sup>

- [3.86] PAO \***qwaHiR** ‘water, stream, creek’: PON \***k<sup>w</sup>ai** ‘fresh water’ (JAR **h<sup>w</sup>əw** ‘creek, stream’, **wə/wə** ‘rain’; ONG **k<sup>w</sup>a**, **koa**, **k<sup>w</sup>ai** ‘stream, creek, river, body of water’); PAN \*\***quSiR** (ATY **qušiya?** ‘water’ [CAD]); PMP \***wahiR** ‘fresh water; stream, river’.

If Proto-Atayal is reconstructed with an initial \***qu**, \***qw**, and evidence is found for PAN \***wahiR**, then this set would be further evidence for a \***qw** vs. \***q** contrast, maintained in Ongan languages, but neutralized in all PAN daughters.

- [3.87] PAO \***talū** ‘long, big’: PON \***talū** (JAR **talū** ‘long’; ONG **dalū** ‘big’); PAN \*\***talū** ‘big’, \*\***dalūq** ‘big, long’ (PAZ **ma-taru-** ‘big, large’; RUK **maa/ɟaw** ‘big’; PAI **ku/ɟal** ‘big’; Bunun **ma/dulo?** ‘long’ [ABVD]; PAI **laɟuq** ‘long’ [with metathesis]; PAN \***adaduq** ‘long [of objects]’).

Regular correspondences are between PON \***talū** and PAN \*\***talū** ‘big’. PAN \***adaduq** appears to be an irregular reflex of \*\***dalūq** ‘big, long’.

- [3.88] PAO \***taŋ** ‘tree trunk’: PON \***taŋ** (JAR **taŋ** ‘tree; canoe; tree branches; vertically arranged sticks on roof’, **deya/taŋ** ‘wooden handle of machete’; ONG **daŋe** ‘tree; dug-out canoe’); PAN \*\***taŋ** ‘tree trunk’ (RUK **taŋ/adana**, Yami **atŋə na** ‘tree stump’; PMP \***ba/taŋ** ‘tree trunk; log’ [Blust 2004:403], \***a/taŋ** ‘cross-beam’; PWMP \***daŋ/anaŋ** ‘handle of knife or sword’; ISN **aru/taŋ** ‘tree trunk’; Sundanese **taŋ/kal** ‘tree trunk’; ITB **o/taŋ** ‘wooden handle of axe’; ITB **taŋ/kal** ‘wood sled drawn by carabao’).

This stem occurs in many JAR and ONG plant names, as well as for handles of implements. By metaphorical extension with dugout canoe, it is also used for other boats, cars, and means of transportation.

### CLASS III. FORMOSAN ONLY, WITH SOME NEW PAN FORMS

- [3.89] PAO \***-aquid** ‘canine tooth, fang’: PON \***-ak<sup>w</sup>ed** ‘tooth’ (JAR **-ahod** ‘tooth; incisor’; ONG **-ak<sup>w</sup>e-** ‘tooth’); PAN \*\***w/aqit** (Proto-Atayalic \***waqit** ‘fang’ [Rau 2002:63]; ATY **waqit** ‘canine tooth, fang, tusk; tooth of a tool or instrument’).
- [3.90] PAO \***batu** ‘seed’: PON \***otu** ‘seed, fruit’ (\***utu** < \***butu**) (JAR **otu**; ONG **oto**); PAN \***batu** ‘stone’ (Blust 1999a), \*\***madu** ‘fruit, seed’ (PAZ **madu** ‘fruit, seed’; RUK **maɟo** ‘fruit’).

Vowel lowering in JAR and ONG is unaccounted for. Formosan data, if cognate, support meaning assigned to PAO lexeme.

- [3.91] PON \***bebe** ‘leaf, leaves’: (JAR, ONG **bebe** ‘leaf, leaves’); PAN \*\***bebe** ‘new growth, fresh shoots’ (PAI **veve** ‘fresh shoots of tree’, **v/n/eve** ‘to sprout’).

23. Though butterflies (apart from some exotic clicking species) do not make sounds, their wing fluttering appears to result in similar sounding words around the world. Compare, for example, Welsh *pili-pila*, Nahuatl *pa:pa:lo:tl*, and Hebrew *parpar* with the forms in this potential cognate set. I include this set for the sake of completeness, but it is possible that convergence, as opposed to direct inheritance, is at play here. Thanks to Yoram Meroz (pers. com. 2006) for reminding me of these facts.



- [3.92] PAO \***ig** ‘water’: PON \***ig** (JAR **ig**; ONG **ige**);<sup>24</sup> PAN \*\***-uŋ** (ATY **ya/uŋ** ‘stream, tributary, small river’, **lilɪ/uŋ** ‘river’; PAZ **rax/uŋ** ‘river, stream’; PAI **Iuŋ** ‘pool, puddle’, **va/uŋ** ‘stagnant water’, **ma-lilʷ/uŋ** ‘water gathers’).

Vowel clusters in ATY and PAI point to a formative **-uŋ** ‘water’.

The \*i/\*u vowel correspondence is unexplained, making this set tenuous, at best; however, the vocalism could be based on that of PAN \*daNum ‘water’.

- [3.93] PAO **kuhiŋ** ‘louse’: PON \***kuhi** ‘louse’ (JAR **kuhi/le** ‘bedbug, louse’, **-le** ‘plural’ [AA:64]; ONG **torag/gile** ‘bedbug, louse’); PAN \*\***kuhiŋ** ‘louse’ (Proto-Atayal \***kuhiŋ** ‘head louse’ [Rau 2004]).

The ONG reflex has undergone reduction \*kuhile > *kuile* > *kile*, with regular post-nasal voicing to *-gile*. ONG *t-ora* is a definite form of *-ora* ‘skin, outer surface’. In JAR, the medial \*h blocks debuccalization of initial \*k > h. Given PAN \*kuCu ‘head louse’, perhaps this was an ancient term for body lice.

#### CLASS IV. PMP ONLY.

##### A. UNLIKELY LOANS, PON RECONSTRUCTIONS, AND REGULAR CORRESPONDENCES

- [3.94] PON \***cikwed** ‘net’ (ONG **cikwe** ‘fishing net’; **cikol/le** ‘make fishing net’); PWMP \***siqud** ‘kind of fish net or trap’.

The PWMP reconstruction is based on two forms, Hiligaynon *siʔud* and Brunei

Malay *siut*. If borrowing occurred, one would expect the source language to have a medial *k* or *kw*, *ku* cluster. Borrowing from PWMP itself can not be ruled out.

- [3.95] PON \***-i/bale** ‘scar’ (ONG **-i/bare**, **-i/bale** ‘scar, wound’ [Portman 1887: 67, 91]); PMP \***balaj** ‘scar’.

- [3.96] PON \***-i/bubu** ‘pour’ (ONG **-i/bube** ‘pour [D&S:21], spill [Portman 1887: 75]’); PMP \***buqbuq** ‘pour, pour out’.

- [3.97] PON \***-i/but** ‘thing’ (ONG **ən/i/bute** ‘thing’); PMP \***buhat** ‘do, make, create; to work; deed; thing, creation’.

- [3.98] PON \***-ilip** ‘to see, to know’ (JAR **-iliye/ma** ‘see, look’; ONG **-ilije**, **-ilijte**, [D&S:69; AA:85]); PWMP \***ilij** ‘see, examine, look closely at’.

Compare ONG *-iaba* ‘see’.

- [3.99] PON \***nok** < \***nuk** ‘bird; hen, rooster’ (JAR **noha** ‘bird; hen, rooster’); PAN \*\***anuk** (Sarangani Blaan **anuk**; ISN **anu?**, **an?anu?** ‘bird’; Timugon Murut **anak**; Toba Batak **anak**; Indonesian **anak ayam**); PMP \***ma/nuk** ‘chicken’ (Blust 2002).

Sound changes are regular: PON closed syllable vowel lowering, followed by JAR \*k > h, and final *a*-insertion.

Along with regular sound correspondences, JAR direct inheritance is supported by absence of the initial syllable and lack of reduplication for the meaning ‘bird’. Blust (2002) does not discuss Malayo-Polynesian forms without initial /m/.

- [3.100] PON \***-omok** < \***-omuk** ‘sleep’ (JAR **-omoha**, **-omohə** ‘sleep’; ONG **-omoka** ‘sleep’); PWMP \***amak** ‘mat, especially one for sleeping’ (cf. Singhi *omuk* ‘mat’).

Note the JAR expression *umoha-ka-cada* ‘mat’ (lit. ‘sleeping place in the house’ [AA:67]).

- [3.101] PON \***umuke** ‘pigeon, dove’ (ONG **umuge** ‘pigeon, dove’); PMP \***muken** ‘omen dove’.

##### B. POTENTIAL LOANS (SOME WITH IRREGULAR CORRESPONDENCES)

- [3.102] ONG **aba** ‘an exclamatory sound denoting disgust or surprise’; PWMP \***aba?** ‘exclamation: ah!, oh!, alas! (said in context of surprise, reproach)’.

- [3.103] JAR **bedai** ‘house’ (Manoharan 1989:165, from Colebrooke 1790); ONG **berai**, **bedai**, **bera** ‘hut, communal hut; mat covering roof of hut’; PMP \***balay** ‘public building, open-sided building, community house, house.’

24. Abbi (2006) transcribes the words for ‘water’ in both Jarawa and Onge with a final palatal nasal /ɲye/. All other authors show a velar nasal.

Note Southern Andamanese *barai* ‘village’ (Man 1883:196)

- [3.104] ONG **bercay**, **berēa** ‘canoe’; PWMP \***baraŋay** ‘kind of large boat, large canoe’.  
 [3.105] ONG **beruge** ‘outrigger (of canoe)’; PWMP \***bilug** ‘type of boat (dugout with out-rigger?)’.  
 [3.106] PON \*-**bule-**, \***ule** ‘evil spirit associated with illness; illness, pain, sickness’ (JAR **ule** ‘pain, fever’, **ule/da** ‘sickness; uneasiness caused by fever, cough, cold, injuries, or any pain’, **-obule** ‘vomit’;<sup>25</sup> ONG **-obule-be** ‘to be sick; to vomit’, **e/bule/be** ‘vomit (n.)’, **ule/te** ‘pain’, **ga/bule/mbe** ‘a class of harmful spirits living above the sky’, **ulu/ge** ‘God’ [P:35]); PPH \***bula-**, \***bulV-** ‘fossilized affix associated with the spirit world; evil spirit associated with illness and protruding tongue’ (cf. PPH \***bula/lakaw** ‘shooting star, meteor; spirit of the shooting star’; Batad Ifugao **bul/la:yaw** ‘fireball with a tail; comet—tradition is that it eats and drinks the blood of a person at night who is not protected by a fire’; ISN **bul/lileway** ‘a spirit who frightens children: his tongue protrudes from his mouth at least six inches. Shamans pronounce his name when they attend children who have a fever’; Batad Ifugao **bula** ‘for someone to push out foreign matter [as food, spent betel quid] from the mouth with the tongue’).  
 Note that \*b is lost word-initially in PON, but not medially, where protected by an initial vowel, as shown by the Onge verb ‘to be sick’. Blust’s (2001) study of the range of PAN fossilized prefixes with spirit associations includes \*buli- and \*buri-, but not PAN \*bula-. However, PPH \***bula-**, with spirit associations, is suggested in Blust (1995).  
 [3.107] PON \***buluga** ‘powerful spirit; creator spirit’ (ONG **uluge** ‘God; creator spirit’ [P:35], [EHM]). Compare ISN **bulúgan** ‘the spirit who talks through the mouth of the shaman, when she is in her trances.’

Southern Andamanese *puluga* ‘the creator spirit’ (Man 1883:31ff) and Aka Bia-da *puluga-da* ‘God’ show the initial labial lost in ONG *uluga* ‘God’ [P:34–35]. Though one might imagine on these grounds that Southern Andamanese is related to Ongan, another possibility is that this lexeme, like [3.100], is a relatively old borrowing, borrowed into both PON and Andamanese prior to initial /b/-loss in Ongan.

- [3.108] PON \***erug** ‘thunder; monitor lizard’ (JAR **erug** ‘monitor lizard, crocodile’; ONG **əyuge**, **eyuge** [EC], **əluge** [P:97] ‘monitor lizard; thunder’); PMP \***duRduR**, \***enduR** ‘thunder’ (cf. ISN **addug** ‘thunder’; ITB **raag/orog** ‘thunder’, **rag/orog** ‘person with low voice’).

Synchronic variation in Onge supports [y], [r], and [l] as variants of the same phoneme.

Although it is possible to reconstruct a PON form, there are several reasons to view this as a borrowing from some PMP language. First, there is a near doublet for ‘thunder’ [3.31], with expected reflexes. Second, there is no PAN cognate with final \*g, but PMP \***enduR** ‘thunder’ yields vowel-initial g-final words in some languages, as illustrated above.

- [3.109] JAR **paraw** ‘raft’; PMP \***para[qh]u** ‘boat’; PWMP \***padaw** ‘kind of sailboat’ [LPO1:183], \***paraSu** ‘boat’ [LPO1:183–84].

**3.2 GRAMMATICAL EVIDENCE.** In addition to the 88 cognate sets in classes I and II above, comparison of the Ongan languages with Proto-Austronesian reveals certain shared grammatical and distributional properties that also support direct inheritance of linguistic subsystems from a mother tongue.

**3.2.1 Bodypart terms with incorporated adpositions.** There are two types of nouns in Jarawa and Onge. Independent or alienable nouns can occur without pronominal prefixes, and may begin with vowels or consonants. Dependent or inalienable nouns can only occur with pronominal prefixes, and are vowel-initial. Dependent nouns include bodypart and kinship terms, and part-whole relations. Some examples of independent vs.

25. Senkuttuvan (2000:41) writes *o:p’pul’la* ‘vomit’.

dependent nouns are shown in (5) for Jarawa and Onge, along with PON reconstructions. In some cases, as in the word for 'bone', comparison of PON forms with cognates [3.28] allows us to analyze the initial vowel as historically epenthetic, or a reanalyzed prefix-final vowel. In other cases, as in the word for 'nose', PAN shows an initial vowel [3.35], suggesting that the PON vowel is not epenthetic. Further discussion of these vowels can be found in 3.2.2.

(5)	INDEPENDENT	JARAWA	ONGE	PROTO-ONGAN	
	'water'	iŋ	iŋe	*iŋ	
	'log; canoe'	taŋ	daŋe	*taŋ	
	DEPENDENT				
	'bone'	m-idəŋ	m-idaŋe	*m-i-dəŋ	'my bone'
	'nose'	m-iŋjanpo	m-iŋabo	*m-iŋjanpo	'my nose'
	'son/daughter'	m-ayi	m-aye	*m-ayi	'my child'

The possessive prefixes that cooccur with dependent nouns are shown in (6), along with proposed PON reconstructions. Note that Jarawa does not make a singular/plural distinction, while Onge does. Both languages have a fourth person prefix that is indefinite and human ('possessed by someone, someone's').

(6)	JARAWA	ONGE	PROTO-ONGAN
1SG	m-	m-	*m-
2SG	ŋ-, n-	ŋ-	*ŋ-
3SG	w-	g-	*gw-
4SG	ən-	ən- (on-)	*ən-
IPL		et- (ot-)	*et-
2PL		n-	*n-
3PL		ek <sup>w</sup> - (ek-, ok-)	*ek <sup>w</sup> -

A salient feature of certain Jarawa and Onge bodypart terms is that they are compositional. Some clearly compositional terms from Jarawa are shown in (7).

(7)	JARAWA	ONGE	
a.	-iŋapo	-iŋabo	'nose'
b.	-iŋapo-talaŋ	-odanaŋe	'nostril'
c.	-ikh <sup>w</sup> a	-ik <sup>w</sup> age	'ear'
d.	-ikh <sup>w</sup> ag-dalaŋ		'ear canal, inner ear'
e.	-cɛɛpo	-cɛɛbo	'eye'
f.	-cɛɛpo-talaŋ		'pupil'

While it may appear that Onge is lacking in this construction type, the forms in (8), all relating to the eye, show that this is not true.

(8)	ONGE	JARAWA	
a.	-cɛɛbo	-cɛɛpo	'eye'
b.	-cɛɛbo-wataykala		'white of eyeball'
c.	-cɛɛbo-tati		'eyelid'
d.	-cɛɛbo-tikiu		'iris'
e.	-cɛɛbo-tugetore		'eyelashes'

Returning to (7), the semantics of the Jarawa terms makes it clear that Jarawa *-talanj* means the ‘inside’ or ‘inner part’ of the bodypart term that it modifies. For those familiar with PAN reconstructions, a cognate is evident. The set [3.34] is repeated below:

[3.34] PAO \***i+dalam** ‘inside’: PON \***i-dalam** ‘inside’ (JAR **itala**, **italanj**; ONG **-iram-**, in **iram-aŋka** ‘inside’ with **-aŋka** locative ‘at, by’ [D&S:21]); PAN \***i-dalem**.

The sound correspondences are regular. The final consonant is lost in an unstressed syllable, but preserved intact in ONG when nonfinal and followed by a vowel. In JAR, the loss of final nasals is incomplete, so that some bodyparts show a final velar nasal or zero.

The significance of set [3.34], however, goes beyond a good lexical cognate set. PON postpositions were clearly productive in bodypart word-formation, as (7) illustrates for Jarawa. In addition, as is clear from (7b), over time, shortening of common terms has occurred so that instead of Onge ‘nostril’ being the expected \*\**-iŋabo-danaŋe* ‘nose-inside’, it is *-odanaŋe*, ‘inside’ in modern Onge.

Forms like those in (7), then, provide additional evidence that PAO and PAN are related languages. The sound/meaning correspondence of the entire CVCVC string in [3.34], along with regular sound correspondences, is an unlikely result of chance, while the grammatical status of PON \***i-dalam** ‘inside’ makes borrowing implausible.<sup>26</sup> In addition, truncation of ‘nose’ in Onge suggests a long internal history of this word-formation process.

**3.2.2 Proximal and distal morphemes.** Jarawa and Onge both have cognate proximal and distal markers, allowing a system to be reconstructed for Proto-Ongan. The Onge system, described in Dasgupta and Sharma (1982:19), shows three distinctive categories for ‘near’, ‘not far away’, and ‘far away’, which I label ‘proximal’, ‘medial’, and ‘distal’, respectively, in (9). These morphemes can occur on their own, as demonstrative pronouns, or can be combined with elements like *-kai* ‘person, thing’, or *-oya* ‘place’, as illustrated in (9).

(9)		+ <i>kai</i>		+ <i>oya</i>	
		‘person, thing’		‘place’	
	PROXIMAL	<i>li</i>	<i>likai</i>	<i>lioya</i>	‘this/this one/here (near)’
	MEDIAL	<i>lə</i>	<i>ləkai</i>	<i>ləoya</i>	‘that/that one/there (not far)’
	DISTAL	<i>lua</i>	<i>luakai</i>	<i>luaoya</i>	‘that/that one/there (far)’

Unlike definite articles, these modifiers precede the nouns that they modify, as in Onge *licelege* ‘through this path’ from *li icele-gi-e* where *li* is the proximate marker, *icele* ‘path’, *-gi* is the definite marker, and *-e* is a postposition meaning something like ‘making use of’.

Although fewer details of the Jarawa system are known, in her section on deixis, Abbi (2006:60) lists Jarawa *li* ‘proximate’ and *lu* ‘distal’, and on this basis, we can reconstruct the same for PON: \***li** ‘proximal’, \***lu** ‘distal’. These forms appear to be cognates of Pazeh and Paiwan deictics, and, coupled with earlier PAN reconstructions, support the PAO and new PAN reconstructions in [3.80], [3.81], repeated below.

[3.80] PAO \***li** ‘proximal, near’: PON \***li** (JAR **li**; ONG **li**); PAN \*\***-li** (PAN \***za/Ni** ‘near’; PAZ **a/lih** ‘near’, **di/ni** ‘here’, **mi/ni** ‘this’, **uhu/ni** ‘now’, **li/a** ‘already’; PAI **Fia-dut** ‘near, nearby’, **mal’i-tsa** ‘here it is!’, **za-zi-tsu** ‘this’. **Fia** ‘dawn, daybreak’).

26. As far as I am aware, there is no language in the world where a preposition or postposition is borrowed exclusively for the purpose of bodypart word-formation.

- [3.81] PAO \***lu** ‘distal, far’: PON \***lu** (JAR **lu**; ONG **lu/a**); PAN \*\***-lu** (PMP \***da/luq** ‘far’, \***Za/luq + ma-** ‘far’; PAZ **a/nu-** ‘far’, **du/a** ‘that, yonder; there’; PAI **ma-lua** ‘far apart’, **mu/dulu** ‘send someone on an errand’, **i-zua** ‘there, over there’, **tsuay** ‘long time’).

**3.2.3 Other grammatical morphemes.** Proto-Austronesian is reconstructed with numerous grammatical prefixes, infixes, and suffixes, and Ca- reduplication as well (Blust 2003). Because most of these affixes are single vowels or CV- or -VC- sequences, and have multiple meanings, good sound-meaning correspondences between PAN and PON may be attributed to chance. Nevertheless, it seems worthwhile to highlight a few potential cognates that may be significant.

Apart from the prefixes shown in (6), the proximate/distal markers in (9), and seemingly sporadic reduplication, Onge and Jarawa appear to be wholly suffixing. A rough template for Ongan verb structure is shown in (10), based on Dasgupta and Sharma (1982:22) and Kumar (2006). In both Jarawa and Onge, tense/aspect suffixes follow the verb, and these suffixes can be preceded by one or more derivational suffixes, plus directional and number markers in Onge. In both languages, mood markers follow tense/aspect markers.

(10) (PRONOMINAL PREFIX) – VERB STEM – (DERIVATIONAL) – (INFLECTION)

Because the majority of bound CV- morphemes are prefixes in PAN and suffixes in PON, it must be assumed that these elements were clitics in PAO, with procliticization dominant in PAN and encliticization dominant in PON. With this as background, I tentatively suggest reconstruction of a PAO clitic, \***ma** ‘stative’ [3.46], repeated below. Unlike other sets given, the minimal phonological form of this particle means that chance could easily be responsible for sound/meaning correspondences, given the small vowel and consonant inventories involved.

- [3.46] PAO \***ma** ‘stative (of verbs)’: PON \***-me** ‘stative verbal suffix’ (JAR **-me**; ONG **-me**); PAN \***ma-** ‘stative verbal prefix’.

Onge and Jarawa both show evidence of a stative verb suffix \*-ma, though it is not clear whether this morpheme is productive in either language. The cognate set for ‘die’ [3.48], which includes this suffix, is repeated below.

- [3.48] PAO \***paCay** ‘die, death’: PON \***peca** ‘die, death’ (JAR **peca** ‘death’ [RS:43], **pecame** ‘die’; ONG **became**, **bencame** ‘die’); PAN \***paCay** ‘kill; \*\*die; \*\*death’ (cf. PAN \***aCay** ‘death’, \***ma-aCay** ‘die, dead’).

In this set, PON has \*-me suffix corresponding to PAN \*ma- prefix marking stative verbs.<sup>27</sup> Other Onge stative verbs that appear to have this suffix are: *-ak<sup>w</sup>arame* ‘feel tired’, *balame* ‘be big surf’, *beleme* ‘smoke (of fire)’ and *-ireame* ‘stout (body)’.

Though both Ongan languages have productive morphological imperatives (Onge *-nene*, Jarawa unsuffixed stems), irregular imperatives with final laryngeals are notable, because (i) final contrastive laryngeals are otherwise not attested in verbs, and (ii) PAN \*-aʔ is reconstructed as an imperative suffix. Both Jarawa and Onge have *koʔ* ‘take it!; hold it!; catch it!’, and for Onge, *caʔ* ‘take’ and *kaʔ* ‘give’ are also attested. These verbs may be relics of a once-productive imperative in PAO, similar to that reconstructed for PAN.

27. An additional PAN \*ma- ‘active verb prefix’ is listed in Blust (2003:471).

**3.3 OTHER EVIDENCE.** In addition to the shared grammatical cognate morphemes noted above, certain distributional properties also support direct inheritance of certain phonological, morphological, and grammatical features from a mother tongue.

**3.3.1 A distributional argument: Explaining a restriction on the PAN lexicon.**

Having introduced the dependent noun system of Onge, Jarawa, and PON, it is worth repeating a fact mentioned in passing above: all dependent nouns in PON are vowel-initial, with pronominal prefixes ending in consonants, producing a canonical C-V pattern at the prefix boundary. This much said, a peculiar phonosemantic feature of the reconstructed PAN lexicon may now be explained.

As illustrated in (11), all reconstructed PAN common nouns beginning with vowels are either body parts or kinship terms, or items that are typically possessed (homes, domestic dogs, and splinters that are embedded in the body).<sup>28</sup>

- |                                 |                                |
|---------------------------------|--------------------------------|
| (11) i. KINSHIP TERMS           | ii. BODYPART TERMS, EXCRETIONS |
| *aki ‘grandfather’              | *ujuj, *ijuj ‘nose’            |
| *ama ‘father, father’s brother’ | *ikuR ‘tail’                   |
| *aNak ‘child, offspring’        | *iSeq ‘urine’                  |
| *apu ‘grandparent/grandchild’   | *uRat ‘vein, sinew’            |
| *aya ‘father’s sister’          | *utaq ‘vomit’                  |
| *ina ‘mother, mother’s sister’  |                                |
| iii. OTHER                      |                                |
| *asu/*wasu ‘dog’                |                                |
| *aCab ‘a cover’                 |                                |
| *ian ‘place of residence’       |                                |
| *uNay ‘sliver, splinter’        |                                |

One clear exception to this generalization is \*aCab ‘a cover’, to which I return below.<sup>29</sup> I suggest that this seemingly arbitrary restriction is a relic of the type of dependent nominal system described for the Ongan languages in 3.2.1. I hypothesize that PAO had dependent nouns that were primarily bodypart items and kinship terms. These nouns were all vowel-initial, preceded by C-final prefixes. In PAN, this system was lost, but remnants of it remain in the form of the phonosemantic peculiarity noted: vowel-initial common nouns are limited to the semantic field of typical “inalienables,” as they were all once preceded by C-final possessive prefixes.

In this context, PAN \*wasu/asu ‘dog’ (Blust 2002) is of interest. Some languages, like Pazeh, reflect \*wasu, while others, like many languages of the Philippines (Reid 1971), reflect \*asu. While there are many PAN doublets, loss of \*w- in daughters of PAN has not only occurred in this word, but in several others as well. As noted by Blust (1999b), initial \*w is lost in Pazeh in \*wanaN > *anan* ‘right side’ and in \*wiRi > *ixi* ‘left side’, “presumably as a result of some now-observed morphological process.”

28. This list includes all PAN V-initial nouns from Blust (1995), and has been checked against more recent publications. PAN \*aNay ‘termite’, suggested in Blust (1999b), is revised to PAN \*SaNay in Blust (2002:124) (where it is misprinted as \*SayaN).

29. This exception was brought to my attention by Bob Blust (pers. comm. 2006). PAN \*aCab ‘a cover’, is based on Puyuma (Tamalakaw) *aʔav* ‘a cover (in general)’, Tagalog *ataʔb* ‘covering, roofing, or siding of palm leaves’, Iban *antap* ‘cover, patch, lining, backing’.

Given comparative data from PON, these facts fall together with the \*wasu/asu facts. As illustrated in (12), certain \*w-initial forms in PAN may reflect earlier PAO third person possessed forms (cf. PON \*kw- ‘3rd person singular’: Onge *g-*, Jarawa *w-* ‘3rd person’).

(12)	PAN		PAO	
	*wasu ‘dog’	<	*kw-asu	(3SG possessed)
	*wiRi ‘left side’	<	*kw-iRi	(3SG possessed)

In this case, there are two potential explanations for the apparent doublet in PAN. One is that, with the domestication of dogs, and their use in the hunt, PAO \*asu was a dependent noun. Another possibility, related to this, is that there was no PAN \*asu, only \*wasu, a lexicalization of an earlier 3 singular possessed form, as proposed in (12). In either case, the mysterious morphological process Blust (1999b) notes for Pazeh is illuminated: initial \*w- was lost in Pazeh and other immediate daughters of PAN, because during this early period, the morphological structure of these terms was still recoverable, with the ancient possessive prefix “subtracted” from the base.

The case of PAN \*-aCab ‘cover’, however, is vexing. Where does this vowel-initial form come from? Based on JAR *-ota*, *-otab*, *-otap* ‘head’ and ONG *-otabe* ‘head’, we may reconstruct a PON form \**-otab* ‘head’. As suggested in [3.84], PAN \*-aCab ‘a cover’ may continue an earlier PAO form meaning ‘head, face, top’, accounting for its appearance in (11).

**3.3.2 PAN monosyllabic roots.** Blust (1988) provides a comprehensive review of the theory, first advocated by Brandstetter (1916), that many disyllabic words in PAN are composed of smaller monosyllabic “roots.” These roots, as Blust (1988) convincingly demonstrates, are recurrent partials of PAN lexemes that, in many cases, are sound symbolic. As Blust shows, however, the sound symbolism is not defined by any intrinsic property but, rather, appears to be defined by the analytic operation of recurrent association. In addition to sound symbolic roots, there are also a fair number of roots that do not appear to be sound symbolic. These recurrent CVC partials are identified as roots on the basis of recurrent sound/meaning associations, and include \**baw* ‘high, top’, \**Duk* ‘ladle, spoon’, and \**pil* ‘attach, join’ (Blust 1988).

Of the 1,804 lexemes from living languages used by Blust (1988) to reconstruct 231 identifiable CVC roots, only 17 lexemes are instances of the monosyllable itself. Blust (to appear a) provides the best explanation of this conundrum in the form of a canonical disyllabic word template for the Austronesian family as a whole. If there were CVC roots in the prehistory of PAN, most of these were eliminated by affixation, reduplication, and vowel epenthesis.

Finding a language that instantiates unaffixed cognates of PAN CVC roots could be viewed as stronger evidence for a family relationship than finding a language with exact cognates. If exact cognates are found, borrowing (before any relevant sound changes) is still a possibility. However, if bare CVC roots are discovered, they arguably reflect an older stage of the language, prior to lexicalization of the disyllabic forms found in so many Austronesian languages. In this regard, several PAO/PON cognate sets in section 3.1 are of special interest, because they include monosyllabic lexemes. Some of these are repeated in (13).

(13) PAO	PON	PAN
*bel ‘smoke’ [3.14]	*bel ‘smoke’	*qe/bel ‘smoke’
*kam ‘mat’ [3.42]	*kam ‘mat’	*Si/kam ‘mat’
*taŋ ‘tree trunk’ [3.88]	*taŋ ‘tree trunk’	*ba/taŋ ‘tree trunk; log’ (PMP)

These cognates provide further evidence of PON and PAN as related languages, and shed new light on PAN word structure, pointing to new monosyllabic roots that have not yet been extensively identified and catalogued. For example, if [3.42] is a valid set, there is further evidence for PAN \*Si- as a recurrent partial, where PAN \*-kam is to \*Sikam as, for example, \*-kaen ‘cat’ is to \*Sikan ‘fish’.

### 3.3.3 Another distributional argument? \*m-initial words in PAN/PON.

While there is no question that the bilabial nasal \*m should be reconstructed for the consonant systems of PAN and PON, both languages have interesting restrictions on words beginning with /m/. In PAN, apart from \*maCa ‘eye’, there are no \*m-initial common nouns reconstructed that do not have vowel-initial doublets, and few, if any, nonderived verbs or adjectives. In contrast, \*m occurs medially and finally in all word classes, seemingly without restriction. In PON, the same is true, except that the word for ‘eye’ is vowel-initial. These distributional facts are quite remarkable in the context of the world’s languages: *m* is relatively stable across time; and there are few languages where *m* does not occur root-initially.

A working hypothesis is that PAO lacked \*m-initial roots, but contained a range of \*m-initial prefixes. In PAN, it is primarily \*m-initial derivational prefixes that survived (see, e.g., Blust 1999a), while in PON, \*m-initial first person singular inflectional prefixes came to dominate. If this is true, both languages have a peculiar phonological restriction on common nouns that one may attribute to shared inheritance.<sup>30</sup>

## 4. DISCUSSION

**4.1 ESTABLISHING LONG-RANGE RELATIONSHIPS.** Most historical linguists believe that the only reliable way of demonstrating common origins of languages is to eliminate chance and borrowing as plausible explanations for observed similarities. This is usually accomplished by means of lexical comparisons because, as Blust (1996:117) emphasizes, “the normal arbitrary association of sound and meaning in the lexicon can be seen as a mass of *independent* historical events not likely to be duplicated by convergent development, and, in the case of basic vocabulary, not likely to be acquired by contact.” The lexical comparisons in section 3, classes I and II, include basic vocabulary items, which are typically stable and resistant to borrowing (EAT, BONE, EGO, BLOOD, FRUIT, HAIR, GIVE, CHILD, INSIDE, URINE, DIE, SKIN, BLACK, RAIN, etc.). Furthermore, they show multiple regular sound correspondences, including the loss of word-initial \*b before high vowels \*i and \*u, loss of final consonants in unstressed syllables, and PAN \*q corresponding with PON \*k. In addition, certain semantic nuances of words (e.g., [3.70] PAO \*buqu ‘BAMBOO JOINT, USED AS CONTAINER; NODE; JOINT’) appear to be preserved over time.

30. With respect to PAN \*maCa ‘eye’, set [3.1] provides a model by which it may be historically derived from earlier \*ma-aCa. If this is so, there may once have been a doublet \*maCa, \*aCa in the same way that we find PAN \*mujin, \*ijun, \*ujun ‘nose’.



While researchers have very different criteria for accepting lexical comparisons, at least two distinct criteria are met by the cognate sets in classes I–III above. First, they satisfy Blust’s (to appear b, chap. 10) quantitative measure of sound correspondences: “a body of, say, 50 cognates would convincingly establish the genetic relationship of two languages or language groups, provided that each proposed sound correspondence is exemplified in at least two, and preferably three forms.” If we take the correspondence between PON \*u and PAN \*bu, we have at least five cases from class I [3.18–3.23], and the correspondence between PON \*k, \*kw and PAN \*q, \*qu is exemplified at least six times in class I [3.9, 3.23, 3.29, 3.50, 3.51, 3.52]. Of note are sets like [3.23] where both initial \*b-loss and \*q > \*k are in evidence. From this class of 64 cognate sets, which nearly all contain well-established PAN reconstructions, Blust’s quantitative measures appear to be met.

Another criterion that solidifies lexical comparisons is the percentage of basic, stable vocabulary involved. Pawley (to appear) notes the difficulties involved in using lexical criteria to establish distant genetic relationships, but stresses that across Indo-European, Austronesian, and Trans New Guinea, there is close agreement in the most stable word meanings across time. For Austronesian, only the first 10 meanings show cognation rates above 50 percent. These meanings are: FIVE, TWO, EYE, WE, LOUSE, FATHER, TO DIE, TO EAT, MOTHER, FOUR. Of this set, firm PAN reconstructions exist for: TWO, EYE, WE, LOUSE, FATHER, TO DIE, TO EAT, MOTHER.<sup>31</sup> And of this set, all but ‘two’ and ‘father’ are included in class I above. By Pawley’s estimates, under the favorable conditions where a protolanguage has left many surviving daughter languages belonging to several different high-order subgroups, and the cognate lexicon has not suffered too much phonological change, on average, about 8 percent of the words in the 200-word list should remain in a single language after 10,000 years. Because, in this case, the hypothesized PAO has only two high-order subgroups, and only two (poorly attested) languages in one of them, the first favorable condition is not met. But if correspondences in tables 4–6 are accurate, the phonology of these languages is extremely conservative, so that at least one favorable condition is met. Under these circumstances, one might demand that at least 16 words from a basic 200-word list be clearly cognate. Considering only class I above, we find: EGO; FETCH/GET/TAKE; CHILD; SKIN; CARRY ON BACK; SUN/DAY; COME/GO; OUR/US; MOTHER; SMOKE; GIVE; FRUIT; BONE; EARTH/SOIL; ONE; INSIDE; NOSE; URINE; EAT/FOOD; FOREST/TREE; DIE/DEATH; BLACK/DARK; WALK/GO; LIE/SLEEP; PATH/TRAIL. Even with nonideal conditions, then, 25 meanings from a basic 200-word list seem relatively stable, meeting the criteria for cognation in basic vocabulary.

31. See Sagart (2004) on PAN numerals. The most important hypothesis he suggests relating to this study is that the numbers 5–10 were innovations in PAN. This is consistent with the little we know about Jarawa and Onge numerals, where only meanings for ONE, TWO, THREE, and MANY appear:

	JARAWA	ONGE	ONGE NOMINAL SUFFIXES
‘1’	<i>wəyə</i>	<i>-i-uaiya</i>	<i>-da, -ra</i> (singular)
‘2’	<i>nayə</i>	<i>-i-naga</i>	<i>-dena, -rena</i> (dual)
‘3’	<i>kaŋ itəjile</i>	<i>-i-rejida</i>	
‘many’	<i>mala-(wə)</i>	<i>-ilake</i> ‘many’	<i>-le</i> (plural)

Onge *eta* means ‘and, also’, so the word for ‘three’ in Onge could be a compound ‘one-and’, meaning ‘one more’.

In short, though lexical material for Jarawa and Onge is extremely limited (no more than 600 words per language, not including unidentified plant names), I have nevertheless been able to identify regular sound correspondences, begin reconstruction of PON, and demonstrate a solid core of lexical cognate sets with regular sound correspondences that appear to defy chance resemblances. Some of these cognate sets result in hypotheses about new PAN forms (classes II and III), many of which have been confirmed by a close examination of additional Austronesian data.

In addition to lexical comparisons with regular sound correspondences, morphological and grammatical evidence is also important in establishing language families. The morphological evidence presented above is of several different kinds. In 3.2.1 I showed that bodypart terms in Ongan languages incorporate reflexes of a cognate of PAN \*i + \*dalem 'in(side)'. Because no known PAN language has bodypart terms of this sort, and because the base of the term is also not found in any Austronesian language, one cannot claim that these terms are borrowed, as there would be no language to borrow them from. Another morphological feature relating PON and PAN is the expression of proximal vs. distal reference. The prefixes or proclitics reconstructed for PON are claimed to be cognate with those for PAN, but again, borrowing does not seem possible, because these particles are independent in PON, but in PAN, they are fused with other elements and do not occur on their own. In fact, it is only with the PON reconstructions in hand that a clear reconstruction of PAN morphemes becomes possible.

Morphological irregularities, which are often key factors in establishing long-distance relationships, have also been identified. In set [3,24], a PAN doublet \*qetaq, \*(h)ataq 'eat something raw' corresponds to PON \*-ita, \*-ata 'eat'. A further irregularity noted in 3.2.3 is the Jarawa and Onge imperative verb *ko?* 'take it!; hold it!; catch it!', which appears to be a relic of a once-productive imperative formation corresponding to that reconstructed for PAN. In the case of PAN/PON doublets, the languages being compared have a "shared aberrancy" of the type highly valued in establishing distant genetic relationships. In the second case, an irregular feature of the Ongan languages is understood as an archaism of a regular morphological process evidenced in Austronesian languages.

Finally, in 3.3 I have suggested distributional evidence that argues for a distant relationship between Onge-Jarawa and the Austronesian language family. If we assume that these language families descend from the same mother tongue, Proto-Austronesian-Ongan, then we can understand why Austronesian vowel-initial common nouns are limited to body parts, kinship terms, and other typical "inalienables," and perhaps even understand the one exception to this generalization. My hypothesis is that PAO resembled PON in having a dependent noun system, where dependent nouns were vowel-initial and nondependent nouns consonant-initial. When this system was lost, dependent nouns were left bare, giving rise to the peculiar distribution noted for PAN. I also suggest that the occurrence of bare monosyllabic roots in PAO/PON, cognate with bound PAN roots, is a strong argument for a family relationship. Given the structure of PAN, we assume that CVC roots were once independent formatives. However, the PAN imperative for disyllabic lexemes (Blust, to appear a), forced these monosyllables into bigger words, via vowel epenthesis, reduplication, and affixation. Some ancestor of PAN, then, should have

monosyllabic free roots, cognate with PAN bound roots. Ongan languages provide evidence for pre-PAN monosyllabic free roots, and support this general hypothesis.

Though borrowing can safely be ruled out for most of the lexemes in classes I, II, and III, this is not true for items in class IV, which have several interesting properties. First, some fail to show expected sound correspondences. For example, in [3.104], where we find ONG *bereay*, *bereə* ‘canoe’, we expect PMP final \*-ay to correspond to PON \*-a or \*-ay (ONG *-aye*). In addition to showing unusual phonological correspondences, some items on the list may be associated with cultural borrowing. Consider, for example, the Onge term *beruge* ‘outrigger’ [3.105]. If Blust (1994) is correct, the outrigger was invented only once in human history. Because everything known about the Andaman Islanders, and particularly the Onge and Jarawa, suggests the most rudimentary technologies, the flimsy outriggers mentioned in section I are likely imitations of those constructed by visitors to the islands. If this is the case, it is not surprising that a word for these constructions was borrowed along with them. Though recent genetic studies suggest long-term isolation of the Andamanese, long-term continued contact between Malay seafarers in search of edible birds’ nests, sea cucumbers, and slaves is evidenced by their knowledge of the geography of the Andamans, including local place names (Cogoy 2002, chap. 2).<sup>32</sup> Under these circumstances, it would be surprising if borrowings were not in evidence.

In the remaining sections, I address two issues that merit further discussion. The first is the general proposal that PON and PAN are sisters. A final section briefly discusses potential relationships between PON and Negrito languages of the Philippines.

**4.2 PAN–PON SISTERHOOD.** The tree in figure 1, repeated below, shows PAN and PON as sisters, daughters of a hypothesized Proto–Austronesian–Ongan. This tree employs a default strategy, namely that every language be regarded as a primary branch of the family to which it belongs. For most readers of this journal, it will be clear that PON is neither a mother nor daughter of PAN. PON is not a mother language of PAN because contrasts that exist in PAN are neutralized in PON, for example, the contrast between \*bu- and \*u-initial words.

If PON is not the mother of PAN, could PAN be a mother of PON? Here also, there is at least one phonological argument, as PON appears to show an inherited contrast between labialized and plain velars and uvulars that was neutralized in PAN. However, the strongest arguments against PON as a daughter of PAN are morphological ones. Most importantly, if PON is a daughter of PAN, we can not explain the existence of bare CVC roots in PON. Recall also from 3.2 that the free-standing proximal and distal markers in Jarawa and Onge allow reconstruction of PON \*li ‘proximal’ and \*lu ‘distal’, while 3.3 highlights the significance of reconstructed PON monosyllabic lexemes. In both cases, isolating morphemes are reconstructed for PON, though in (nearly) all Austronesian languages where cognates occur, the same morphemes do not occur as independent words. Given general patterns of grammaticization where free particles or clitics become parts of bigger words, and given more specific details of PAN phonological drift demonstrating a clear preference for disyllabic lexemes (Blust, to appear a), the obvious conclusion to draw is that PON reflects an earlier stage of PAN, where morphemes like the proximal/

32. Recent genetic studies suggesting a long period of isolation for the Andamanese are Thangaraj et al. (2002), Endicott et al. (2003), and Thangaraj et al. (2005).

distal markers and monosyllabic roots were free forms. But if this is the case, then PON cannot be a daughter of PAN, for PAN shows a stage where these elements are no longer free, with proximal/distal markers undergoing morphophonemic processes that make protoforms difficult to recover.

A final argument for the two families as sisters involves the PON dependent noun system outlined in 3.2 and 3.3. If PON is a mother of PAN, then there is a problem, because all Austronesian languages appear to have lost the dependent noun system. If, on the other hand, PAN is a mother of PON, then we have no explanation for the very peculiar distribution of vowel-initial common nouns in PAN, and it is simply a coincidence that dependent nouns in PON evolved to be vowel-initial. On the sisterhood account, properties of both PON and PAN are explained by inheritance from PAO. I suggest that PAO had a dependent noun system, and that this was inherited in PON. On the other hand, a major innovation in PAN was the loss of this system, with reconstructible remnants in the form of vowel-initial bodypart and kinship terms, and rare \*w- subtraction of identifiable prefixes of dependent nouns from early PAN daughters.

**4.3 PON AND PHILIPPINE NEGRITO LANGUAGES.** As noted in section 1, Negrito populations are found in other parts of the Austronesian-speaking world, most notably in the Philippines. Reid (1994) identifies possible non-Austronesian lexical elements in these languages. With preliminary PAO reconstructions, it is worth reviewing these lexical elements and comparing them to what is known of Ongan and Proto-Ongan. Reid (1994) lists 158 unique forms found in one or more Philippine Negrito languages, but not cognate with any known Austronesian forms. Reid originally assumed that these unique terms were Austronesian in origin, and that they were either lost in most Austronesian languages (or subject to radical semantic or phonological change to make them unrecognizable as cognates), or the result of Negrito coinages, once the Negritos had adopted local Austronesian languages. Reid's (1994) hypothesis, however, is that these unique Negrito lexemes may reflect a non-Austronesian language spoken by Negritos prior to their adoption of Austronesian languages via contact in the Philippines.

Although equivalent terms for a great number of these unique Philippine Negrito lexemes are not known for Ongan languages, the comparisons in (14) suggests a continuing possibility that Reid's original hypothesis was correct: unique lexemes in Philippine Negrito languages may reflect an ancient Austronesian, or rather an Austronesian-Ongan, substrate. Philippine Negrito language data in (14) are from Reid (1994). The number in parentheses is the number of Philippine Negrito language groups reflecting these unique forms.

(14)		Negrito	Ongan
1.	'rattan'	*lati (4)	ONG tati (P:63)
15.	'fire'	*duku (2)	PON *tuke [3.57]
16.	'hair, feather'	*pulug (2)	JAR puliyu 'gill of fish'
58.	'sweat'	*asub (1)	JAR -otahila/ahu 'sweat'
84.	'burn'	*tʷəm/uk (1)	JAR tuhəb-, PON *tuke [3.57]
103.	'burn'	*ti/duk (1)	JAR tuhəb-, PON *tuke [3.57]

However, more striking than these lexical resemblances are two general phonotactics of both language groups. First, Philippine Negrito languages and PON both tolerate mono-

syllabic lexemes. Reid (1994:49) notes that “the Negrito languages prior to Austronesian contact were far more tolerant of monosyllabic morphemes than Austronesian languages typically are.” Though Reid (1994:49–50) assumes that monosyllables, as well as other words, are the result of shortening or abbreviation, this may not be the case. He compares Central Agta *aba* ‘shoulder’ with Proto-Philippine \**abaRa*, but this Central Agta form may be cognate with PON \*-*apa* (cf. ONG *-aba-* ‘carry on shoulder’), and more generally with [3.8] (note that Central Agta is a geographic grouping of possibly related Negrito languages).

A second property that Pramod Kumar (pers. comm. 2006) notes for Jarawa is a strong tendency for final obstruents to be voiced (rather than voiceless). Reid’s (1994) list of Philippine Negrito reconstructions shows 38 lexemes with final voiced stops, as opposed to 51 with final voiceless stops. While final voiceless stops still predominate, as they did in PAN, the ratio of voiced to voiceless stops is much higher than one would expect, given PAN phonotactics.

The most remarkable similarity in the Philippine Negrito reconstructions and PON forms are the retained sound patterns with relatively large semantic shifts. For example, Reid reconstructs \**urat* ‘river, water’. The phonology suggests that this form belongs squarely in [3.61b] with ‘vein, artery’, yet this sort of semantic shift is not accepted by all historical linguists as a sound/meaning correspondence. The same could be said for some of the shifts in meaning I have proposed above, and could be due to long time depth, the striking cultural differences between PAN-speakers and ancestors of the modern Negrito hunter-gatherer populations, or both.

**5. CONCLUDING REMARKS.** In this study, the comparative method has been applied to two related languages of the southern Andaman Islands, Jarawa and Onge. This comparison has allowed reconstruction of a protolanguage termed “Proto-Ongan,” and the identification of regular sound changes in each daughter language as discussed in section 2, and summarized in (3) and (4). The same method was used in section 3 to demonstrate that Proto-Ongan is related to Proto-Austronesian. Lexical and grammatical evidence supports the view that Proto-Ongan and Proto-Austronesian are sisters, daughters of Proto-Austronesian-Ongan. Lexical evidence is provided in section 3.1, and includes dozens of cognate sets demonstrating regular sound correspondences, including basic vocabulary items. Though some of the items in class IV look like potential loans, the majority of the hundred-plus sets appear to be due to direct inheritance. Grammatical evidence is presented in 3.2, and includes an incorporated postposition, the proximal and distal deictic particles, inherited doublets matching those in PAN, and fossilized irregular imperatives in PON inherited from PAO. Additional distributional evidence is presented in 3.3, and is specific to the reflexes of a proposed PAO dependent noun system, and the PAN root system, as reconstructed by Blust (1988).

A striking feature of PON and PAN is their conservative historical phonologies. Both are recognizable as sisters, and have changed little from PAO. I suggest that this conservatism is a consequence of two independent features of Proto-Austronesian-Ongan: first, its relatively simple (C)V(C) syllable structure; and second, its dearth of inflectional morphology. Inflectional morphology gives rise to multimember paradigms, and when para-

digms exist, analogical changes (e.g., leveling, extension) are likely to follow. PAO did not have such a system and, as a consequence, analogical change has not muddied the clear waters of regular sound change. As a consequence of relatively simple sound patterns and isolating morphology, then, it appears that PON, PAN, and many of their daughters have changed little over thousands of years.<sup>33</sup>

Although comparative study of Austronesian and Ongan languages has just begun, and a number of details of the analysis presented are certain to need revision, I hope to have demonstrated that these two language families are very likely related. The implications of this discovery are wide ranging, from new solutions to old problems in PAN grammar, to potential PAO homelands and revised migration routes. In these areas, and many more, Andamanese studies may illuminate our understanding of Austronesian pre-history, and provide new general perspectives on long-distance linguistic comparisons.

## REFERENCES

- Abbi, Anvita. 2004. The great experience: A linguistic fieldtrip to the Andaman Islands, 2001–2002. Centre of Linguistics and English, Jawaharlal Nehru University, New Delhi, India. Available at: [www.andaman.org/BOOK/originals/Abbi/art-abbi.htm#visit](http://www.andaman.org/BOOK/originals/Abbi/art-abbi.htm#visit).
- . 2006. *Endangered languages of the Andaman Islands*. Munich: Lincom Europa.
- Austronesian basic vocabulary database. University of Auckland. <http://language.psy.auckland.ac.nz/austronesian/>
- Bloch, D. N. 1952. L'Andaman. In *Les langues du monde*, ed. by A. Meillet and M. Cohen, 511–21. Paris: Éditions du CNRS.
- Blust, Robert A. 1988. *Austronesian root theory*. Amsterdam: John Benjamins.
- . 1994. The Austronesian settlement of mainland Southeast Asia. In *Papers from the Second Annual Meeting of the Southeast Asian Linguistics Society*, ed. by Karen L. Adams and Thomas John Hudak, 25–83. Tempe: Program for Southeast Asian Studies, Arizona State University.
- . 1995. Austronesian comparative dictionary, computer file, University of Hawai'i, Honolulu.
- . 1996. Beyond the Austronesian homeland: The Austric hypothesis and its implications for archeology. *Transactions of the American Philosophical Society* 85 (5), *Prehistoric settlement of the Pacific*, 117–58.
- . 1999a. 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. Symposium Series of the Institute of Linguistics No. 1. Taipei: Academica Sinica.
- . 1999b. Notes on Pazeh phonology and morphology. *Oceanic Linguistics* 38:321–65.
- . 2001. Historical morphology and the spirit world: The \*qali/kali- prefixes in Austronesian languages. In *Issues in Austronesian morphology: A focusschrift for Byron W. Bender*, ed. by Joel Bradshaw and Kenneth L. Rehg, 15–73. Canberra: Pacific Linguistics.
- . 2002. The history of faunal terms in Austronesian languages. *Oceanic Linguistics* 41:89–139.

33. This is the likely reason that identification of PAN cognates in Jarawa and Onge was possible in the first place, even with limited lexical data of variable quality.

- . 2003. Three notes on early Austronesian morphology. *Oceanic Linguistics* 42:438–78.
- . To appear a. Disyllabic attractors and anti-antigemination in Austronesian sound change. *Phonology* 24.
- . To appear b. *The Austronesian languages*. Cambridge: Cambridge University Press.
- Brandstetter, R. 1906. *Ein Prodromus zu einem vergleichenden Wörterbuch der malaio-polynesischen Sprachen für Sprachforscher und Ethnographen*. Luzern.
- . 1918. *Root and word in the Indonesian languages. An introduction to Indonesian linguistics*. (C. O. Blagden, trans.). London: The Royal Asiatic Society.
- Campbell, Lyle. 1998. *Historical linguistics: An introduction*. Edinburgh: Edinburgh University Press.
- . 2003. How to show languages are related: Methods for distant genetic relationship. *Handbook of historical linguistics*, ed. by Brian D. Joseph and Richard D. Janda, 262–82. Oxford: Blackwell.
- Cogoy, Esteveao. 2004. *Onge dictionary*. [www.freelang.net/dictionary/ong.html](http://www.freelang.net/dictionary/ong.html).
- Colebrooke, R. H. 1795. On the Andaman Islands. *Asiatic Researches* 4:385–95.
- Cooper, Zarine. 2002. *Archaeology and history: Early settlements in the Andaman Islands*. Oxford: Oxford University Press.
- Dasgupta, D., and S. R. Sharma. 1982. *A Handbook of the Önge language*. Calcutta: Anthropological Survey of India.
- Dempwolff, Otto. 1938. *Austronesisches Wörterverzeichnis, Supplement 19*. Berlin: Dietrich Reimer.
- Directorate of Tribal Welfare, Andaman and Nicobar Administration. 2004. Resources in the Jarawa Territory. Ms. posted at [www.andaman.nic.in](http://www.andaman.nic.in).
- Egerod, Søren. 1980. *Atayal-English dictionary*. London: Curzon.
- Egli, Hans. 2002. *Paiwan Wörterbuch*. Wiesbaden: Harrassowitz Verlag.
- Elkins, Richard E. 1968. *Manobo-English dictionary*. Honolulu: University of Hawai'i Press.
- Ellis, A. J. 1885. The president's annual address for 1882: Report on researches into the language of the South Andaman Island. *Transactions of the Philological Society* 1882–84:44–73.
- Endicott, Phillip, M. Thomas, P. Gilbert, Ch. Stringer, C. Lalueza-Fox, E. Willerslev, A. J. Hansen, and A. Cooper. 2003. The genetic origins of the Andaman Islanders. *American Journal of Human Genetics* 72(1):1590.
- Ferrell, Raleigh. 1982. *Paiwan dictionary*. Canberra: Pacific Linguistics.
- Greenberg, Joseph. 1971. The Indo-Pacific hypothesis. In *Current trends in linguistics* vol. 8, ed. by Thomas A. Sebeok, 807–71. The Hague: Mouton.
- Groves, C. 1981. Ancestors for the pigs: Taxonomy and phylogeny of the genus 'Sus'. *Technical Bulletin no.3, Department of Prehistory, Research School of Pacific Studies*. Canberra: Australian National University.
- Healey, Phyllis M. 1960. *An Agta grammar*. Manila: Bureau of Printing.
- Kumar, Pramod. 2005. Phonological structure of Jarawa. Handout from work-in-progress talk, January 18, Max Planck Institute for Evolutionary Anthropology, Leipzig.
- . 2006. Jarawa morphology. Handout from work-in-progress talk, November 7, Max Planck Institute for Evolutionary Anthropology, Leipzig.
- . In progress. A descriptive grammar of the Jarawa language. PhD diss., Max Planck Institute for Evolutionary Anthropology, Leipzig.
- Li, Paul Jen-kuei, and Shigeru Tsuchida. 2001. *Pazih dictionary*. Taipei: Institute of Linguistics, Academia Sinica.
- Lynch, John. 2001. *The linguistic history of southern Vanuatu*. Canberra: Pacific Linguistics.

- Man, E. H. 1883. On the aboriginal inhabitants of the Andaman Islands. *Journal of the Anthropological Institute* 12:69–116, 327–434.
- . 1923. *Dictionary of the South Andaman language*. Bombay: British India Press.
- Manoharan, S. 1989. *A descriptive and comparative study of Andamanese language*. Calcutta: Anthropological Survey of India.
- . 1997. Pronominal prefixes and formative affixes in Andamanese language. In *Languages of tribal and indigenous peoples of India: The ethnic space*, ed. by Anvita Abbi, 457–73. Delhi: Motilal Banarsidass.
- Mathew, Sam P., A. Mohandas, and G. M. Nair. 2004. *Piper sarmentosum* Roxb.: An addition to the flora of Andaman Islands. *Current Science* 87:141–42.
- Nair, V. S. 1979. A note on the language of Jarawa. *Bulletin of the Anthropological Survey of India* 28:3–4, 17–38.
- Pawley, Andrew. To appear. Recent research on the historical relationships of the Papuan languages, or, What does linguistics say about the prehistory of Melanesia? In *Population genetics, linguistics, and culture history in the Southwest Pacific: A synthesis*, ed. by Jonathan Friedlander. New York: Oxford University Press.
- Pejros, I. 1994. Some problems of Austronesian accent and \*T ~ \*C (Notes of an outsider). *Oceanic Linguistics* 33:105–28.
- Portman, M. V. 1887. *Manual of the Andamanese languages*. Calcutta: Office of the Superintendent of Government Printing. [Reprinted in 1992 by Manas Publications, Delhi.]
- . 1899. *A history of our relations with the Andamanese*. 2 vols. Calcutta: Superintendent of Government Printing.
- Radcliffe-Brown, A. R. 1914. Notes on the languages of the Andaman Islands. *Anthropos* 9:36–52.
- . 1922. *The Andaman Islanders: A study in social anthropology*. Cambridge: Cambridge University Press.
- Rau, Der-Hwa Victoria. 2004. Lexical similarity, sound change, and intelligibility of Atayalic dialects. In *Papers in Austronesian subgrouping and dialectology*, ed. by John Bowden and Nikolaus Himmelmann, 37–95. Canberra: Pacific Linguistics.
- Reid, Lawrence A. 1971. *Philippine minor languages: Word lists and phonologies*. Oceanic Linguistics Special Publications No. 8. Honolulu: University of Hawai'i Press.
- . 1994. Possible non-Austronesian lexical elements in Philippine Negrito languages. *Oceanic Linguistics* 33:37–72.
- . 2006. Historical linguistics and Philippine hunter-gatherers. Paper presented at the workshop “Historical linguistics and hunter-gatherer populations in global perspective,” Leipzig, August 10–12. Ms., University of Hawai'i.
- Ross, Malcolm. 1996. Is Yapese Oceanic? In *Reconstruction, classification, description: Festschrift in honor of Isidore Dyen*, ed. by Bernd Nothofer, 121–60. Hamburg: Verlag Meyer & Co.
- Ross, Malcolm, Andrew Pawley, and Meredith Osmond, eds. 1998. *The lexicon of Proto Oceanic, 1: Material culture*. Canberra: Pacific Linguistics.
- . 2003. *The lexicon of Proto Oceanic, 2: The physical environment*. Canberra: Pacific Linguistics.
- Sagart, Laurent. 2004. The higher phylogeny of Austronesian and the position of Tai-Kadai. *Oceanic Linguistics* 43:411–44.
- Senkuttuvan, R. 2000. *The language of the Jarawa: Phonology*. Anthropological Survey of India. Calcutta: Government of India, Ministry of Culture, Youth Affairs, and Sports, Dept. of Culture.
- Sreenathan, M. 2001. *The Jarawas: Language and culture*. Anthropological Survey of India. Calcutta: Government of India.



- Temple, Richard C. 1902. *A grammar of the Andamanese languages, being Chapter IV of Part I of the Census Report on the Andaman and Nicobar Islands*. Port Blair: Superintendent's Printing Press.
- Thangaraj, Kumarasamy, Lalji Singh, Alla G. Reddy, V. Raghavendra Rao, Subhash C. Sehgal, Peter A. Underhill, Melanie Pierson, Ian G. Frame, and Erika Hagelberg. 2003. Genetic affinities of the Andaman Islanders, a vanishing human population. *Current Biology* 13:86.
- Thangaraj, Kumarasamy, Gyaneshwer Chaubey, Toomas Kivisild, Alla G. Reddy, Vijay Kumar Singh, Avinash A. Rasalkar, and Lalji Singh. 2005. Reconstructing the origin of Andaman Islanders. *Science* 13 (308/5724):996.
- Tryon, Darell T. 1995. *Comparative Austronesian dictionary*. Berlin: Mouton.
- Tsuchida, Shigeru. 2003. *Kanakanavu texts (Austronesian Formosan)*. Kyoto: Endangered languages of the Pacific Rim.
- Vanoverbergh, Morice. 1972. *Isneg-English vocabulary*. Oceanic Linguistics Special Publication No. 11. Honolulu: University of Hawai'i Press.
- Weber, George. 2006. *The Andamanese*. www.andaman.org/BOOK/text.htm.
- Yadaav, Yogendra. 1985. Great Andamanese, a preliminary study. Series A-67. Canberra: Pacific Linguistics.
- Yamada, Yukihiko. 2002. *Ibayat-English dictionary*. Kyoto: Endangered languages of the Pacific Rim.
- Zide, Norman, and Vishvajit Pandya. 1989. A bibliographical introduction to Andamanese linguistics. *Journal of the American Oriental Society* 109:639-51.

### Appendix 1. Basic vocabulary of Jarawa, Onge, (and Proto-Ongan) \*

	JARAWA	ONGE	GLOSS	PROTO-ONGAN
I.1	mi	mi	'I'	*mi
I.2	ŋi	ŋi'	'you sg'	*ŋi
I.3	mi	eti	'we'	*eti
I.4	li	li	'this'	*li
I.5	lu	lua	'that'	*lu
I.6	one	ŋakai	'who'	*one-akai 'which one'
I.7	onaha	one-nare	'what'	*one 'what, which'
I.8	nadem	narema	'no, not'	*nadema
I.9	mala-awde	-lotalota	'all'	
I.10	mala; tapo	kutu; waje; -ilake	'many, much'	
I.11	wəya	iwaiya	'one'	*waya
I.12	naya	inaga	'two'	*naga
I.13	hutu, uma-me [MS:86]	dalū, ue, ume	'big'	*uma 'big'
I.14	talū	jagai, ue, uye	'long'	*talū 'long, big'
I.15	boiya	baya, bayai	'small'	*baya
I.16	tayoha	əŋa-cebe, əŋ-ale	'woman'	
I.17	taho (A)	ʉŋagile	'man'	
I.18	əŋ, eŋ	əŋe	'person, native'	*eŋ, *əŋ (< *en)
I.19	noha	tuge	'bird'	*nuk
I.20	wəm(e)	weme	'dog'	*weme, *wem?
I.21	ləhubə (A) kuhile	torəŋi torəŋ/gile	'louse'	*kuhi
I.22	taŋ	dəŋe	'tree, log'	*taŋ
I.23	otu	-oto	'seed; fruit'	*otu
I.24	bebe	bebe	'leaf'	*bebe
I.25	icitəŋ	icindəŋe	'root, buttress'	*ici-təŋ 'tree root'
I.26	nodəhe (A)	-əŋwi 'skin, bark'	'bark (of tree)'	*-əŋwi

**Appendix 1. Basic vocabulary of Jarawa, Onge, (and Proto-Ongan) (Continued)\***

JARAWA	ONGE	GLOSS	PROTO-ONGAN
I.27 -atiye	-ati	'skin'	*-ati
I.28 -ipo, -epo	-ibo	'flesh'	*-ipo
I.29 ceŋ, -aceŋ	-aceŋe	'blood'	*-aceŋ, *ceŋ
I.30 -idəŋ	-idəŋe, -ireŋe	'bone'	*dəŋ, *-idəŋ
I.31 -epotaŋ (A)	-eboteŋe	'tears'	*-epo-t-iŋ 'water of eye'
I.32 tuhə	tuke	'fire'	*tuke
I.33 yane	gane	'egg'	*gane
I.34 wiru		'horns'	*-iru, *wiru (?)
I.35 awug 'fish tail' yawugu 'bird tail' yaw 'tail (of dog)' [RS:56]	ukicele	'tail'	*-ukicele
I.36 -ode	-ode, -ore	'feather'	*-ode 'hair, feather'
I.37 napo	coge	'fish'	*napo
I.38 -ode	-ode, -ore	'hair'	*-ode 'hair, feather'
I.39 -otap	-otabe	'head'	*-otab 'head, face'
I.40 -ik <sup>w</sup> a, -ik <sup>w</sup> agu	-ik <sup>w</sup> age	'ear'	*-ikwag
I.41 -ecepə	-ejebo	'eye'	*-ecepə (< *cca-ipo)
I.42. -opetaŋ	-oboraŋe, -obeduŋe (P)	'fingernail'	*-(o,e)pe-taŋ (*taŋ, -daŋ 'bone; nail)
I.43 -ijnapo	-ijabo, ijaibo	'nose'	*-ijjan-ipo
I.44 -olu, alu	-alaŋe	'mouth'	*-alaŋ
I.45 -ahod	-ak <sup>w</sup> e	'tooth'	*-akwed
I.46 -adal(e)	-alaŋe	'tongue'	*-adalaŋ
I.47 -ug(e)	-uge	'foot'	*uge (< *-uke)
I.48 -olo	-ola, -olo	'knee'	*-olo
I.49 -om(e)	-ome	'hand'	*-ome
I.50 -ore, -oh <sup>w</sup> e, -ih <sup>w</sup> ede, -ih <sup>w</sup> edbo heyonu	-ioŋe	'stomach'	
I.51 -anjito	-anjito	'fish abdomen'	
I.51 -anjito	-anjito	'neck'	*-anjito
I.52 -akag	-agage, -agake	'breast'	*-akak
I.53	-ukukotabe	'heart'	
I.54	-oacu	'liver'	*-oracu
I.55 -ijŋco	-ijŋjo, -ijjo-	'drink (vb)'	*-ijŋco
I.56 -ita-	-ida, -ada	'eat (vb)'	*-ita, *-ata (dbl.)
I.57	-ibaga-	'bite (vb)'	
I.58 -yewəba, -iliyema	-atiyebe, -ilijŋe, -ilijŋe	'see, look (vb)'	*-ilijŋe 'see, know'
I.59	-koʔlak <sup>w</sup> e, -yeyeʔ	'hear, listen (vb)'	
I.60 -iniyələ	-ilijŋe, -ilijŋe	'know (vb)'	*-ilijŋe 'see, know'
I.61 omohə	omoka	'sleep (vb)'	*omok < *omuk
I.62 pecame	became, bencame	'die (vb)'	*peca-me
I.63 aik <sup>w</sup> a	otak <sup>w</sup> e	'hit w. arrow, kill'	
I.64 h <sup>w</sup> ara	k <sup>w</sup> a(n)e	'swim (vb)'	*kwara
I.65 ugikə-	-ugiʔkci	'fly (vb)'	*ugike-
I.66 cayway(a)	bujeye	'walk (vb)'	
I.67 alema	alema, -ayo-	'come (vb)'	*alema
I.68 tule	tule, nelaŋa (A)	'lie down (vb)'	*tule
I.69 -ətəhə	-antəkə	'sit (vb)'	*antek
I.70 dokəte-	doka-	'stand (vb)'	*doka
I.71 -iya	bele; -gabe-	'give (vb)'	*bele

**Appendix 1. Basic vocabulary of Jarawa, Onge, (and Proto-Ongan) (Continued)\***

	JARAWA	ONGE	GLOSS	PROTO-ONGAN
1.72	-atiba	bini 'speak;drop;throw' cenene- 'talk'	'say, speak (vb)'	
1.73	che	eke	'sun, daylight'	*eke
1.74	tape, tabe	cileme	'moon'	
1.75	cilape, ciləhe	kəyəkəye	'star'	
1.76	ij	iŋe	'water'	*ij
1.77	wəwə	ujəŋe, ujəŋe	'rain'	
1.78	uli	uli	'stone'	*uli
1.79	pile, bilə	belai; agəŋe, agəŋe	'sand'	
1.80	pela, bela	gea; tutano	'earth; land'	*bela 'mud, dirt, earth'
1.81	paŋŋaŋ, paŋŋan-	beŋŋoŋe, bennaŋe	'sky'	*(p,b)aŋŋan 'flat (surface)'
1.82	pənel, bənel	beleme	'smoke'	*bel 'smoke'
1.83	h <sup>w</sup> iyə, omoha (A)	toŋkute	'ashes'	
1.84	tuhəb-le-yame		'burn (vb)'	
1.85	icala	icele	'path, trail'	*icala
1.86	wed	we	'clay'	*wed
1.87	horgidu	gero, geru	'red'	*g <sup>i</sup> idu
1.88	ele, elu	-i/ela	'blue'	*ele
1.89	alam	alame	'red ochre'	*alam
1.90	hala	kala	'white'	*kala
1.91	hiru	i-kiu	'black'	*kiru
1.92	kitale	toŋki, boykwe	'night'	
1.93	hulu	kulu	'hot'	*kulu
1.94	əmeyn (A)	-ajomo, -jome; -ugite-	'cold'	*jom
1.95		-ejebute	'be full'	
1.96	ŋoŋo	ŋoŋo	'nurse, suckle'	*ŋoŋo
1.97	cew	bado, baro	'good, nice, 'beautiful'	*baro
1.98	biti, piti	-ibiti	'bad'	*(p,b)iti
1.99	h <sup>w</sup> aw	k <sup>w</sup> ai	'stream, creek'	*k <sup>w</sup> ai
1.100	atiba	-atibe	'name'	*-atiba

\* Words marked (A) are only attested in Abbi (2006). Other Jarawa words are from Sreenathan (2001) or Kumar (2006, in progress). Onge words marked (P) are from Portman (1887); others are from Dasgupta and Sharma (1982) and Cogoy (2004).

**Appendix 2. Some Jarawa-Onge cognates with Proto-Ongan reconstructions**

	PROTO-ONGAN	JARAWA	ONGE	
2.1	*-aceŋ	-aceŋ	-aceŋe	'blood'
2.2	*-agi	-agi	-agi/cebe	'husband'
2.3	*-akak	-akag	-agage, -agake	'breast'
2.4	*-akwed	-ahod	-ak <sup>w</sup> e	'tooth'
2.5	*-ala	-ala	-al/larŋe	'shoulder'
2.6	*alam	alam	alame	'red ochre'
2.7	*ale	-ale	ale, -aye	'child'
2.8	*alema	alema	-alema-	'come (toward speaker)'
2.9	*-aŋa	-aŋab	-aŋa/be 'make s.o. ones wife; take a wife (of a man)'	'wife'
2.10	*-aŋito	-aŋito	-aŋito	'neck, throat'

## Appendix 2. Some Jarawa-Onge cognates with Proto-Ongan reconstructions (Continued)

PROTO-ONGAN	JARAWA	ONGE	
2.11 *-apa	-apa	-aba	'carry on back'
2.12 *-ataŋ	-ataŋ	-ata	'shoulder blade'
2.13 *atiba	atiba	-atibe	'name'
2.14 *-atiy	-atiye	-ati	'skin'
2.15 *aw	aw	aʔ	'bow'
2.16 *-aya	-aya	k/aye/ri	'mother (address term)'
2.17 *baŋnan	baŋnan, baŋnaŋ	bennaŋe	'sky, clouds; flat'
2.18 *baya	baiya	baya	'small'
2.19 *bebe	bebe	bebe	'leaf'
2.20 *bei	bai-te-	bei	'go, come'
2.21 *bela, *pela?	bela	bela	'mud, dirt, earth'
2.22 *biti, *piti?	biti	-ibiti	'bad, not good'
2.23 *-bule	ule, -obule	ule-, -obule	'illness, pain, sickness'
2.24 *cana-	cana-	cena-	'shell; thorn, spine, pointy protrusion'
2.25 *cataŋ	cataŋ 'chisel'	cerage 'k.o. arrow; spade'	'sharp instrument'
2.26 *cile-	ciləhe, cilape 'star'	cileme 'moon'	'shine (of heavenly body)'
2.27 *dag	dag	dage	'coconut tree, coconut'
2.28 *dakweg	daaig	dakwage	'cane; cane basket'
2.29 *daŋ	-idaŋ	-idaŋe	'bone'
2.30 *day	daw, daa	daye	'boy; son'
2.31 *doy	dow	doyi	'adze'
2.32 *-eca	-ece-	-eja-, -cje-	'face, forehead'
2.33 *ecele	əcaʎa, icele	eecele	'child' (cf. *ale)
2.34 *-ecepo	-ecepo	-ejebo	'eye'
2.35 *-ecele	-ecele	-ejele	'urinate'
2.36 *eke	ehe	eke	'sun, day'
2.37 *ekw-	eh-, əh-	ekw-	'3PL subject prefix'
2.38 *-ekwal	-ehwal	-ekwale	'navel'
2.39 *-ele	-ele-, -elə	-i/ela	'blue, green'
2.40 *-eni	-enia	-epi	'laugh (vb)'
2.41 *-eŋjapo	-eŋjapo	-enabo	'buttocks'
2.42 *əŋ	əŋ	əŋe	'person; member of group; Ongan'
2.43 *ən-	ən-, on-	ən-	'4th person prefix'
2.44 *icala	-icala	-icele	'path, trail'
2.45 *-icale	icale	-ijele	'sharp edge (e.g., of knife)'
2.46 *-ici	-ici-	-ici-	'base, root (of tree), calf (of leg)'
2.47 *-ikola	iholə	ikola	'side, lateral side (of body)'
2.48 *-ikwag	-ik <sup>w</sup> a, -ik <sup>w</sup> agu	-ik <sup>w</sup> age	'ear'
2.49 *-ilije	-iliye-	-ilije, -ilije	'see, know'
2.50 *inen	inen	inene	'foreigner, stranger, not Ongan'
2.51 *-iŋanpo	iŋanpo, -iŋanpo	iŋaibo, -iŋabo,	'nose'
2.52 *-iŋco	-iŋco	-iŋjo	'drink (vb)'
2.53 *iŋ	iŋ	iŋe	'water'
2.54 *-ipil	-ipil	-ibile	'hand, arm, branch, appendage'
2.55 *-ipo	-epo, ipo	-ebo, -ibo	'flesh, cheek, fleshy part'
2.56 *-ita	-ita	-ida, -ite	'eat, bite (vb)'
2.57 *kaka	haka	kaga	'crab'

## Appendix 2. Some Jarawa-Onge cognates with Proto-Ongan reconstructions (Continued)

	PROTO-ONGAN	JARAWA	ONGE	
2.58	*kala	hala	kala	'white'
2.59	*kam	ham	kame	'cane mat or platform'
2.60	*kayiw	həwə	kəyə, kəye, gai	'forest; stick'
2.61	*kiru	hiru	-ikiu	'black'
2.62	*kiya	hiya	ka? (imp.)	'give'
2.63	*koriko	horko	korige	'owl'
2.64	*kulu	hulu	kulu	'hot'
2.65	*kui	h <sup>w</sup> əuwə, hoi	kui	'pig; animal'
2.66	*kwai	h <sup>w</sup> aw	k <sup>w</sup> ai	'creek, stream'
2.67	*kwara	h <sup>w</sup> ara	k <sup>w</sup> ae	'swim (vb)'
2.68	*lee	-le, leo	leə, lee	'knife, dagger, iron'
2.69	*lew	ləw	le	'honey'
2.70	*lətə	yətə	lətə	'bat'
2.71	*li	li	li	'this'
2.72	*lu	lu	lua	'that'
2.73	*-me	-me	-me	'stative verbal suffix'
2.74	*mi	mi	mi	'I (1SG subject pronoun)'
2.75	*m-	m-	m-	'1 SG possessive/subject prefix'
2.76	*na	na	na	'2 PL possessive pronoun; yours PL'
2.77	*nadema	nadem	narema	'no, not; verbal negator'
2.78	*naga	naya	ninaga, inaga	'two'
2.79	*nam	nam(e)	name, yame	'firewood; ignite fire, burn firewood'
2.80	*ŋoŋo	ŋoŋo	ŋoŋo	'nurse, suckle'
2.81	*ŋi	ŋi	ŋi	'2SG subject pronoun; you SG'
2.82	*ŋ-	ŋ-	ŋ-	'2SG possessive prefix; your SG'
2.83	*-ode	-ode, -od	-ode, -ora	'hair, feather'
2.84	*okw	oha	ok <sup>w</sup> -	'flower'
2.85	*-olo	-olo	-ola, -olo	'knee'
2.86	*-ome	-om(e)	-ome	'hand'
2.87	*-omuk	-omohə	-omoka	'sleep (vb)'
2.88	*-otab	-otab, -otap	-otabe	'head, face'
2.89	*otu	otu	-oto	'seed; fruit'
2.90	*partu	parto, partu	barto	'arrow'
2.91	*pecame	became	became	'die (vb)'
2.92	*pone	pone	bone, mone	'resin, gum'
2.93	*takwe	tahē	takwe	'pipe made of crab claw'
2.94	*talū, *dalū?	dalū	dalū 'big'	'long'
2.95	*taŋ, *daŋ?	taŋ, daŋ	daŋe, taŋe	'log, tree, trunk, stem; canoe'
2.96	*tuke	tuhə	tuke	'fire'
2.97	*udu	udu	uru-ge	'cough'
2.98	*uge (<*-uke)	-ug(e)	-uge	'foot'
2.99	*uku	uhu	uku	'wooden bucket'
2.100	*uli	uli	uli	'stone'
2.101	*wana	wana	wana	'weep, cry (vb)'
2.102	*waya	wəyə	i-uaiya, i-woya	'one'
2.103	*wed	wed	we	'clay, earth'
2.104	*wem	wəm, wə-wəm	wəmə, weme	'dog'
2.105	*-yaba	yo-yəba	-iaba	'see (vb)'
2.106	*-ya	-yə	-a, -ə	'nonpast, noncompletive'