

The West Papuan Phylum revisited: the genetic status of the Yapen island languages

MARK DONOHUE

National University of Singapore

mark@donohue.cc

The non-Austronesian languages of Yapen island were first assigned to the West Papuan Phylum classificatorily, most closely related to the eastern Bird's Head languages. This decision that was revoked in 1975 as part of the widespread reclassification of Papuan languages, and the languages were grouped with the newly established Geelvink Bay Phylum. This short note seeks to reinstate the non-Austronesian languages of Yapen in the West Papuan Phylum.

1. Introduction

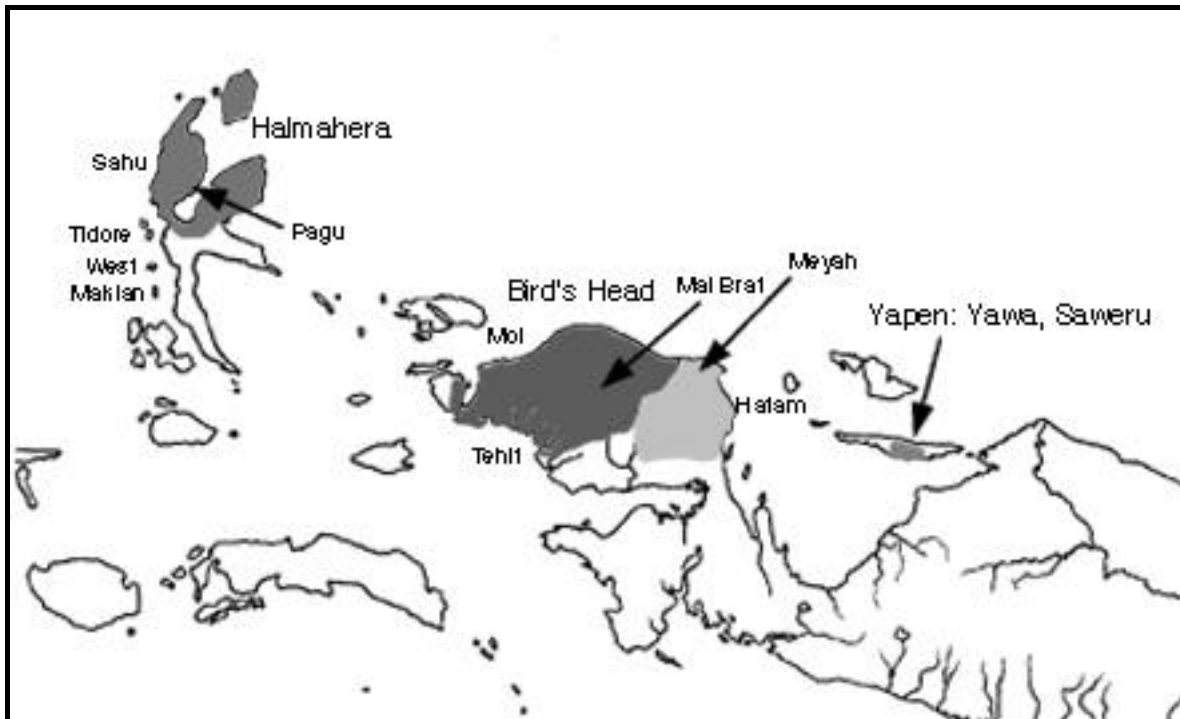
A funny thing happened to me before going to sleep one night, a week or two after returning from a trip to New Guinea. While there, in addition to conducting my own language work, I'd photocopied some materials, including Sarah Wimbish's MA thesis on Pagu (Wimbish 1990). (I'd always been attracted to the North Halmahera languages, with all their prefixes, and Pagu even allows final consonants!) Glancing through it, I came across a table listing the bound pronominal forms, and was immediately convinced that they were related to the Saweru agreement markers that I'd been learning and practising over the last month. The rest of this short article is intended to qualify that conviction in an academically respectable way.

2. The non-Austronesian languages of Yapen

Yapen island, in Cenderawasih Bay on the north-west coast of New Guinea just east of the Bird's Head, is home to ten Austronesian languages (Anceaux 1961, Silzer and Ajamiseba 1980), and two non-Austronesian ones, Yawa (Yava) and Saweru. Saweru was until recently thought to be a divergent dialect of Yawa, though Anceaux did allow that the reported differences might be great enough that 'it will no longer be possible to deny it the status of a separate language.' (Anceaux 1961: 9). I shall refer to Yawa and Saweru collectively as 'the Yapen languages', so as not to misrepresent what my sources have claimed, but also to not create the false impression that Saweru's status has been considered particularly uncertain vis-a-vis the status of Yawa. I am not claiming that the various Austronesian languages of the island are part of this grouping.

Since the major classificatory works were carried out and published in the 1970s, work on the Yapen languages has expanded dramatically (Jones 1986a, Jones 1986b, Larry Jones 1986, etc., Donohue 2001, Ayeri, Karuri, Maniakari and Donohue 2003), as well as more detailed descriptive material becoming available for the languages of both the Bird's Head and of North Halmahera (henceforth referred to simply as 'Halmahera'; just as the label 'Yapen' is used to refer to only the non-Austronesian languages of Yapen, and not to include the Austronesian languages of that island, so too 'Halmahera' is being used here solely to refer to the non-Austronesian languages of the north of the island, as well as Ternate and Tidore to the west). This has made materials available that allow us to reconsider the classifications that have been assumed for the last few decades.

Map: Halmahera, the Bird's Head and Yapen



The map shows the relative locations of the groups under discussion, as well as the rough locations of the individual languages mentioned in the text. From Tidore to the western Bird's Head is approximately 500km, while western Yapen to the eastern Bird's Head is only 150km; from western Halmahera to central Yapen by sea is approximately 1200km in total, well within the range of sailing technology in the region (the Sultan of Tidore traditionally carried out slaving raids on the southern Bird's Head, south and east of Tehit, proving that it is not unreasonable to assume maritime links over such distances). Significantly, in contrast to the abundant harbours along the south coast of Yapen and much of northern Halmahera, the Bird's Head offers little in the way of safe anchorages, a point that will be relevant in section 8.

3. The Yapen languages: classification

Previous comparisons of the Bird's Head and Yapen have focused on lexical correspondences (Cowan 1961, Voorhoeve 1975), with an emphasis on the free pronouns. Problems with this approach are related to the apparent instable status of pronouns in the area: New Guinea generally is the site of a lot of pronominal instability, as evidenced by the behaviour of Malay varieties in the area (Donohue 2002), and a cursory examination of the pronominal forms in the languages concerned.¹ On the other hand, Voorhoeve himself admitted that this method was applied only because of the near-complete absence of any structural information about the languages (Voorhoeve 1975a: 717, 1975b: 872, 874), a dearth that has only really been corrected in the last decade. Cowan (1960) placed the Yapen languages in his expanding West Papuan Phylum, which was centred on the Bird's Head, but extended to North Halmahera,

¹ Diffloth's (1994) statement that 'Southeast Asia is not an area where pronouns are known for their immortality' could equally apply to New Guinea/Melanesia, which fades into Southeast Asia imperceptibly.

Timor, Alor and Pantar, and Yapen. Capell (1975: 678, Diagram A) includes Yapen in his assessment of the West Papuan Phylum, linking it with the languages of what is now known as the East Bird's Head family by virtue of the similarities in first and second person singular subject prefixes (he lists a common pair *t-* for 1SG and *b-* for 2SG). This was later refuted by Voorhoeve (1975b: 874-875) (emphases mine – MD):

What, then was the basis for this reassignment of Yava to a different family? To quote Voorhoeve at length,

The present classification of Yava as a member of the Geelvink Bay phylum is not without problems. ... Yava has phylum-level relationships with a few languages in both the Trans-New Guinea Phylum (TNGP) and the West Papuan Phylum (WPP). ... *With the almost total absence of structural information* on the languages in question ... it is difficult to assess the weight of the lexical evidence; the sporadic cases of phylum-level relationships with languages of the TNGP and WPP seem in themselves not enough to posit either TNGP or WPP membership for Yava, ... *The bulk of the lexical evidence does not support any of these solutions. For this reason, Yava has been united into a separate phylum with the languages of the East Geelvink Bay family.* These at least, in addition to having phylum-level relationships with Yava, share with it the characteristic of not belonging clearly to either the TNGP or WPP.

The lexical data that 'give "phylum-level relationships" between East Geelvink Bay languages and Yapen languages' will be discussed in section 7. For now, we can note that Voorhoeve did state that '[t]he original set-up of this section was to give a survey of the languages of the West Papuan Phylum located on the mainland of New Guinea and on Yapen island...if the West Papuan Phylum was to be interpreted as a phylum in the lexicostatistical sense, several languages had to be dropped from it. These are ... and the Yava language on Yapen' (1975a: 717). This suggests that Voorhoeve had doubts about a complete reliance on lexicostatistics as a classifying tool.

Since the application of the comparative method relies on comparing bound morphology, and not the more ephemeral free lexical items, the available bound morphology, pronominal agreement marking, has been the focus of my examination. Why are only the pronominal agreement markers being compared here? In part, it reflects the short announcement-nature of this article; in part, it reflects the data available. But mostly, it reflects the fact that, especially in the Bird's Head languages, there is not a lot of bound morphology. Descriptions of languages from this area that have emerged since Voorhoeve's 1975 pronouncement show that bound morphology is in the main restricted to the agreement markers (outside the North Halmahera family, which has extensive and cognate verbal morphology). Since this is the material that we have to work with, and since these morphemes are reasonably accurately reported, they have been used.

Equally importantly, we also have a much better idea of the relationships, and non-relationships, of the languages in the area, thanks largely to work carried out in the 1980s and 1990s in Halmahera and the Bird's Head areas. The relationship of the North Halmahera languages and at least some of the Bird's Head languages is now well-established, though the exact nature of that relationship remains in question. At the same time the earlier pan-Bird's Head grouping has been shown not to extend to the eastern end of the peninsula: the languages there form two groups, the East Bird's Head family, and the unresolved Hatam language.

4. Bound morphological correspondences

Firstly, the positive evidence. Table 1 shows ‘subject’ and ‘object’ prefixes in a selection of north Halmahera languages; tentative reconstructions of these affixes for proto-Halmahera are given in table 2. The North Halmahera languages have been chosen to evenly represent the different subgroups of the North Halmahera stock (as described in Voorhoeve 1988); the proposed reconstructions draw on data from other languages as well as those presented here, though the selection given here is reasonably representative.

Table 1. Halmahera agreement prefixes

	Tidore		Sahu		Pagu		West Makian	
	‘subject’	‘(poss’r)’	‘subject’	‘object’	‘subject’	‘object’	‘subject’	‘(poss’r)’
1SG	to-	ri-	to-	ri-	to-	n-	tV-	ti
2SG	no-	na-	no-	ni-	no-	n-	nV-	ni
3SG.M	wo-	i- ~ wi-	o-	u-	wo-	w-	i-	mV
3SG.F	mo-	mi-	mo-	mi-	mo-	m-	i-	mV
3SG.NH	yo-	ma-	i-	a-	yo- / i-	(Ø)	i-	dV
1PL.EX	mo-	mi-	mi-	muʔu-	mi-	n-	a-	mi
1PL.IN	fo-	ni	wo-	na	wo-	n-	mV-	nV
2PL	no-	ni	ni-	nu	ni-	n-	fV-	fi
3PL	yo-	na- / ma-	ʔ [i/u]-	aʔu-	yo-	k-	dV-	di

Table 2. Tentative proto-North Halmahera agreement prefixes.

	‘subject’	‘object’		‘subject’	‘object’
1SG	*to-	*[r/n]-	1PL.EX	*mi-	*mi-
2SG	*no-	*ni-	1PL.IN	*[f/w/m]o-	*na-
3SG.M	*wo-	*wi-	2PL	*ni- ~ *fV-	*ni-
3SG.F	*mo-	*mi-	3PL	*di- ~ *yo-	? *di-
3SG.NH	*yo-	*Ca-			

Most of the reconstructions in table 2 should be unproblematic. I have assumed a degree of paradigmatic levelling in all languages, particularly West Makian. The plural subject forms, particularly, present a challenge to notions of regularity, but this might in part reflect the material that is being compared: The ‘subject’ prefixes are the main morphological means for indexing the S or A argument of the clause, but also have other extensions and restrictions on their use, varying from language to language. Similar cautions apply to the ‘object’ markers, which are also used as markers of inalienable possession in the Yapen languages, and which are cognate with possessive prefixes (rather than P-marking morphemes) in some North Halmahera languages. Furthermore, in both Halmahera languages and Yapen languages there is a small, but significant class of stative intransitive verbs that inflect by means of the ‘object’ prefixes, rather than the ‘subject’ markers: the languages have a ‘split-intransitive’ or ‘stative-active’ system of verbal indexing.²

² In Yapen there are further complications: Saweru allows a very small class of intransitive predicates to inflect by means of *dative* suffixes, and in both Yawa and Saweru there are several strategies for marking the arguments of transitive predicates (Donohue 2001). These details do not affect the reconstructions presented here.

Table 3 shows the equivalent morphemes in the three Yapen varieties, two dialects of Yawa, one western and one eastern, and Saweru, which is spoken on the island of the same name off the south coast of Yapen.

Table 3. Yapen agreement prefixes

	Yawa (Sarawandori)		Yawa (Yapanani)		Saweru		proto-Yapen	
	‘subject’	‘object’	‘subject’	‘object’	‘subject’	‘object’	‘subject’	‘object’
1SG	ʃo(-)	in-	so-	ny-	o=	ina-	*ʃo	*in
2SG	no(-)	n-	no-	t-	no=	na-	*no	*n
3SG.M	po(-)	Ø-	ɸo-	t-	ɸo=	a-	*ɸo	*a
3SG.F	mo(-)	r-	mo-	t-	mo=	ra-	*mo	*r
1DU	ririmo(-)	ririns-	nimo-	nint-	imo=	isa-	*(r)imo	*ins
2DU	ipo(-)	is-	jo-	y-	i=	ya-	*i(ɸo)	*is
3DU	yo(-)	y-	jo-	y-	yo=	ya-	*yo	*y
1PL.EX	reamo(-)	reans-	nea-	mant-	amo=	asa-	*(r)amo	*(r)ans
1PL.IN	wamo(-)	wans-	nea-	mant-	amo=	asa-	*(w)amo	*(w)ans
2PL	wapo(-)	was-	wao-	mant-	wa=	wa-	*wa(o)	*wans
3PL	wo(-)	m-	wao-	mant-	wo=	ya-	*wo	*m

The correspondences between the bound pronominal morphology in the North Halmahera languages and the bound pronominal morphology in the Yapen languages are close. Amongst the singular forms, the 2SG, 3SG.F and 3SG.NF/3SG.M forms are strikingly alike. The lack of any consonant in the 1SG form in Saweru follows a regular loss of proto-Yapen *ʃ in Saweru, and so its cognacy with Yawa ʃo- is not in doubt, and the adaptation of a *t to a fricative is widely attested in New Guinea and beyond (possibly through a yet earlier form **t[j/i]o, showing the common New Guinean sibilization of before high (front) vowels: *t* [ɟ], [ʒ], [ʃ], [tʃ] / *i* (, u), but this is speculative).

The correspondences in the non-singular forms are closest if we compare the Halmahera plural forms with the Yapen duals; but can this methodology be justified? Examining the uses of the dual forms in Yapen shows that it can. In Saweru, for instance, there is a dual:plural opposition, but unusually it is the duals, not the plurals, that are used for unmarked, generic reference, as in the following example:

Saweru

- (99) *wonggam=o imo=nunu*
board=LNKR ‘1DU’.NOM=sit
‘board that we sit on; chair’
- (99) *Amo=nunu farum iri wonggam=o imo=nunu=ai*
1PL.NOM=sit all LOC [board=LNKR ‘1DU’.NOM=sit]=OBL
‘We’re all sitting on chairs’

Furthermore, when referring to non-humans, we find that the duals, rather than the plurals, are often used in generic cases.

Saweru

- (99) *Raruatan* *wo=no* *Aiwendui=ai.*
 people 3PL.NOM=be.at Saweru=OBL
 ‘There are people on Saweru.’
- (99) *Dian* *yo=no* *rawanan=ai.*
 people ‘3DU.NOM’=be.at sea=OBL
 ‘There are fish in the sea.’

These morphosyntactic quirks make no sense if the duals are compared with respect to their opposition to the contemporary plural; all that is known of the dynamics of markedness suggests that the plural forms, not the dual forms, should be used for generic reference. If, however, the modern duals of Saweru are derived historically from a (paradigmatically unmarked) plural set, then the retention of the unmarked use for generic reference can be understood (though the motivation for a shift to dual reference remains obscure), and the non-dual relic functions of these forms in Saweru is easily explained.

We can now compare some pronominal data from the Bird’s Head languages. Examining table 4, we can first note that there are no striking similarities between the languages of the eastern Bird’s Head (Meyah and Hatam) and those of the centre and west.³ Differences of this sort are part of what convinced Voorhoeve that a reassessment of the original WPP, conceived of as containing all the languages of the Bird’s Head, was required. Meyah is now considered to be part of the small East Bird’s Head family, and the status of Hatam is still unresolved. Comparing the western and central languages Tehit and Mai Brat, and comparing with tables 1 - 3, we can see that, while the Bird’s Head languages form a tightly-knit group, as far as bound morphological correspondences go, they are more divergent as a group from both Yapen and Halmahera than these two groups are from each other, despite being geographically intermediate between the two. Mai Brat, perhaps appropriately for a language in the geographical centre of the Bird’s Head, shows the greatest divergence from the patterns found to the north- and south-west in Moi and Tehit.

Table 4. Bird’s Head agreement prefixes

	Tehit	Moi	Mai Brat		Meyah		Hatam
1SG	t-	t-	t-		di-		di-
2SG	n-	n-	n-		bi-		a-
3SG.M	o- / w-	w-	y-		Ø-		ni-; Ø-
3SG.F	m-	m-	m-				
1DU.EX					ma-		
1DU.IN					na-		
2/3DU					ge-		
1PL.EX	m- / m-	m-			me-		i(g)-
1PL.IN	ɸ- / f-	w-	p-		mi-		ni-
2PL	Ø / n-	n-	n-		i-		ji-
3PL	i- / y-	w- / l- / j-	m-		ri-		si-

³ Other than the fact of prefixal agreement for subject, which is in any event common in a large belt from Southeast Sulawesi to the eastern Torricelli mountains, in which the Halmahera-Bird’s Head-Yapen area is found

Table 5. Tentative proto-(western) Bird's Head agreement prefixes.

1SG	*t-	1PL.EX	*m-
2SG	*n-	1PL.IN	*f-
3SG.M	*w-	2PL	*n-
3SG.F	*m-	3PL	*y-

The irregularities in the pronominal paradigms for the Bird's Head, when compared to the similarities exhibited between the Halmahera and Yapen languages, are just that – irregularities. We can demonstrate this by reference to one of the well-known typological features of the Bird's Head languages, their SVO word order, and a less well-known correlation of the area, that of alignment. The conclusion is that, as is supported elsewhere (Reesink 1996), the Bird's Head has been the site of enormous linguistic areal pressures in the past, resulting in convergence phenomena. This means that the pronominal information from the Bird's Head is not necessarily genetically original.

5. The Bird's Head: order and alignment

I assume that the original proto-West Papuan language had an SOV order; this is in keeping with the dominant order found in the languages of this area, and also reflects the main orders found in both the eastern Yapen languages and most of the western Halmahera languages, meaning that while only one SOV → SVO change would be required if proto-West Papuan was SOV, two SVO → SOV changes would be required if proto-West Papuan was SVO. We have evidence of SOV → SVO in some of the more heavily Austronesian-influenced languages of Halmahera (Ternate, Tidore), and a similar model of change is plausible for the Bird's Head languages. We know that there has been long term contact with the Nusa Tenggara area, evidenced through the extensive integration of the *kain timur* trade into customs throughout the Bird's Head area, and we also know that there has been extensive and intimate contact with the more local Austronesian languages in Cenderawasih Bay. The traders of the *kain timur* were most likely speakers of SVO languages, the order pattern of all the languages between the Bird's Head and Nusa Tenggara, and SVO is the sole word order attested amongst the Cenderawasih Bay Austronesian languages. In short, we have a very plausible mechanism to explain an SVO order change in the Bird's Head languages. If proto-West Papuan was SVO, on the other hand, we need to come up with some explanation of the SVO → SOV changes in both Yapen and Halmahera. Yapen might possibly be explained by appealing to a 'substrate' of SOV-ordered non-Austronesian languages, which have since disappeared; the only plausibility in this line of reasoning lies in the fact that the non-Austronesian languages spoken on and behind the Waropen coast are SOV in order. The almost total lack of relations between these inland groups and the Yapen island languages does not lend much support to these arguments. On the other hand, in Halmahera the same appeal to substratal influences is the only recourse open to explain the SOV order, if we started out with SVO, but here there is no evidence for any other SOV languages in the region. All the evidence suggests that the only non-Austronesian linguistic influences in the area stem from the extant North Halmahera languages themselves.

In terms of alignment, we find that the Bird's Head languages show a 'standard' (for New Guinea, and indeed for most of the world; statements on 'standardness' of grammatical

features are based on Nichols 1993) nominative(-accusative) alignment, in terms of their agreement markers. While this is found in some Halmahera languages, there are also significant numbers of languages with a split in intransitive marking: a stative/active, or dynamic/nondynamic split.⁴ And this is exactly what is found in the Yapen languages as well: the agreement marking on the verb shows separate prefixal/proclitic forms marking the A and the P of bivalent clauses, and these same agreement markers both being used to index different single arguments of a monovalent clauses. The dominant alignment type of the Cenderawasih Bay Austronesian languages mirrors that of the Bird's Head languages; with clearly nominative prefixes, and restricted accusative suffixes.⁵

6. Significance of the pronominal evidence

Testing this closed data set (the seven members each of the 'subject' and 'object' sets, not counting the difference between inclusive and exclusive forms (only the inclusive have been used), or between dual and plural forms (only the Yapen duals have been used) for its status as a genetic marker (Nichols 1996), we find that it more than meets the criteria that she sets up. Of the fourteen possible consonantal matches, there are six clear matches. There are other correspondences as well: the 1SG.SUBJ *to-:ʃo-*, 3SG.NF.SUBJ *wo-:ʔo-*, 1NSG.SUBJ *mi-:(r)imo* and 2NSG.SUBJ *fV-:i(ʔo)* are highly likely to be cognates, as discussed earlier, and the 3PL.OBJ *k-:ya-* is also likely. They have not been considered in the discussion immediately following in order to minimise the chance of my hypothesis of relatedness being proven. The following correspondences are unproblematic:

Table 6. Clear matches in form and function

		Halmahera	Yapen
'subject'	2SG	<i>no-</i>	* <i>no</i>
	3SG.F	<i>mo-</i>	* <i>mo</i>
	1NSG	<i>mio-</i>	* <i>(r)imo</i>
	3NSG	<i>yo-</i>	* <i>yo</i>
'object'	1SG	<i>n-</i>	* <i>in(a)</i>
	2SG	<i>n-</i>	* <i>n(a)</i>

Ignoring the vowel correspondences, the chances of these six consonants all appearing in the closed pronominal paradigms examined, all with the functions ascribed to each member, can be calculated, assuming (following Nichols 1996: 48-56) a one in five chance of any consonant appearing in a particular function. I shall assume no value to the vowel correspondences. Multiplying these probabilities together yields the following chance of random occurrence.

$$\begin{array}{cccccc} 2SG & 3SG.F & 1NSG & 3PL & 1SG & 2SG \\ 0.2 & \times & 0.2 & \times & 0.2 & \times & 0.2 & \times & 0.2 & = & 1/15,625 \end{array}$$

⁴ This is explicit in Galela (Shelden 1991, 1998), and can be inferred in Pagu from comments in Wimbish (1990: 33: 'the object markers may also be used alone (without a preceding subject marker) when the subject is an experiencer or patient, rather than an agent').

⁵ The Cenderawasih Bay Austronesian languages typically have accusative suffixes only for third persons, and then only optionally. Bird's Head languages vary as to whether they have accusative suffixes or not, but it is rare.

While this is not within the realm of individual identifying evidence (which requires a probability of approximately 1/100,000 or less),⁶ it is less than the 1/10,000 level that is ‘at least interesting’ (Nichols 1996: 49). If we factor in the other likely correspondences (assuming roughly double the chance for the unproblematic correspondences, that is, a 0.4 probability), the figure drops to 1/3,800,000, which would represent a clear match. Even if we ignore these extra correspondences, there is considerable supporting evidence that links the two families together:

- in both cases we see an *n-* in the 2SG.SUBJ, but an *n-* in both 1SG and 2SG for object;
- that the languages concerned make, and mark, a gender distinction in the same way;
- the ‘object’/(undergoer) paradigm is also used to mark possession.
- there is a split in the first person nonsingular into an inclusive and an exclusive category.⁷
- these bound morphemes are in both cases prefixal/proclitic forms;
- the languages mark a split-intransitive (= ‘stative-active’) system, and not the typologically more frequent nominative-accusative or ergative-absolutive system;

Not all of these factors can be judged to be of equal weight; prefixation for subject agreement is the common pattern in a wide band of languages, stretching from southern Indonesia west of Timor past the area in question here to the eastern Torricelli ranges in Papua New Guinea. Showing a more restricted range, split-intransitive alignments are common in the languages from southern Indonesia to Yapen; in both cases, these features cross genetic boundaries, indicating that areal spread has occurred. The prefixal agreement markers are also used to indicate possession, something that is not typical of the area described, but which is widely found amongst the non-Austronesian languages of the New Guinea region, and so which possibly represents an earlier, pre-Austronesian areal feature.

The remaining morphosyntactic features, however, mark the two groups as highly likely to be related. Note that comparing these features with the Bird’s Head languages yields results that are not nearly as impressive. The pronominal evidence can only be based on seven forms, since most Bird’s Head languages do not have a separate object marking paradigm. In the person marking, then, there are five matches with the Halmahera languages (1SG: *t-*, 2SG: *n-*, 3SG.M: *w-*, 3SG.NON-M: *m-*, 1PL: *m-*), yielding a 1/3,125 chance of random occurrence, which is insignificant. (The figures for Bird’s Head:Yapen are similarly non-significant. The 1SG form cannot count, and the 3SG.M/3SG.NF is only a partial match, but the 3PL *i-/y-* is a match: it works out to 1/1,250). Furthermore, there is no syncretism in the object marking paradigm to call on, nor the typologically rare split-intransitive alignment system. We shall return to these questions in section 8 below.

⁶ The 1/100,000 level is based on a standard 1/20 or 0.05 chance of error, multiplied by 6,000 (the approximate number of languages known to be extant in the world).

⁷ This is also true of most of the Bird’s Head languages, but not Mai Brat, the most inland of the group. It might be that Mai Brat represents the Bird’s Head language least influenced from without.

7. The Geelvink Bay hypothesis reviewed

To counter this bound morphological data, we should consider the alternative hypothesis, that the Yapen languages are in fact related to the languages of the Waropen coast hinterland, the East Geelvink Bay family, in Voorhoeve's Geelvink Bay phylum. This classification was based solely on lexical comparisons. What are these lexical data that give 'phylum-level relationships' between East Geelvink Bay languages and Yapen languages? Voorhoeve (1975b: 875-876) lists the following forms (Baropasi and Bauri [= Bauzi] were also compared, but only Tarunggare forms were given for comparison). This is presented in table 7.

Table 7. Voorhoeve's Yapen: Geelvink Bay correspondences

	Yava	Tarunggare		Yava	Tarunggare
'come'	<i>nde, re</i>	<i>nere</i>	'tail'	<i>ateva</i>	<i>otapara</i>
'eat'	<i>rai(s)</i>	<i>ghayo</i>	'water'	<i>karu</i> (moisture)	<i>waro</i>
'fly'	<i>bariri</i>	<i>bunana</i>	'wind'	<i>obar</i>	<i>bwa</i>
'foot'	<i>najo</i>	<i>nal</i>	'you (sg)'	<i>uein</i>	<i>ei</i>
'give'	<i>ra</i>	<i>nore</i>	'you (pl)'	<i>wea</i>	<i>wi</i>
'I'	<i>nei, rei</i>	<i>ei</i>			

Examining these lexical items, the following cautions apply: Yapen *rai(s)* is morphologically *ra-* '3SG.FEM.GEN/ACC' plus *i* 'eat (transitive)' (I cannot explain the bracketed '(s)', except to note that it is not heard in Yawa or Saweru these days). The probable cognate is then reduce to an *i:ghayo* 'correspondence'. Yapen *bariri*: Tarunggare *bunana* evidences only one out of six phonemes as cognate, unless an *r:n* correspondence can be established, and so remains tentative. Yapen *najo* is morphologically *na-* '2SG.GEN' plus *yo* 'foot, leg', yielding no correspondences at all with Tarunggare, and only a single vowel with Baropasi *naro*. A single consonant defines Yawa *-ra* and Tarunggare *nore* 'give', but Saweru *awe* 'give' means that *-ra* cannot be reconstructed for proto-Yapen in any case. Yapen *inei*: Tarunggare *ei* relies solely on vocalic correspondences, as does 'you (singular)' and 'you (plural)'. Yawa *obar*: Tarunggare *bwa* is possible, and would confirm the *b:b* seen in 'fly'. The set Yapen *re*: Tarunggare *nere* 'come', and *ateva*: Tarunggare *otapara*, in which *a-* is the 3SG.NFEM.GEN prefix for the Yapen forms, seem plausible, as is Yawa *karu*: Tarunggare *waro*, but these are then only three possible cognates, not enough to base a genetic relationship on.

We can examine the pronominal data from the one East Geelvink Bay family language for which we have adequate data, Bauzi (Briley 1997). The base pronominal forms are shown in table 8. There is no 3PL pronominal form in Bauzi; all the pronominal bases in this table must be suffixed for case in order to be used in grammatical clauses; the choices are *-ho* ergative, *-m* absolutive, and *-ba* dative, as well as some pragmatically, and not syntactically, determined choices. Notice that the alignment marked in Bauzi matches neither the nominative-accusative pattern of the Bird's Head, nor the stative-active systems of Halmahera and Yapen. Significantly, there is no verbal agreement, another point of structural disagreement; in terms of word order, the East Geelvink Bay family languages all appear to be basically SOV languages, the sole point of similarity with Yapen, and one shared with the vast majority of all New Guinea area languages.

Table 8. Bauzi pronominal forms

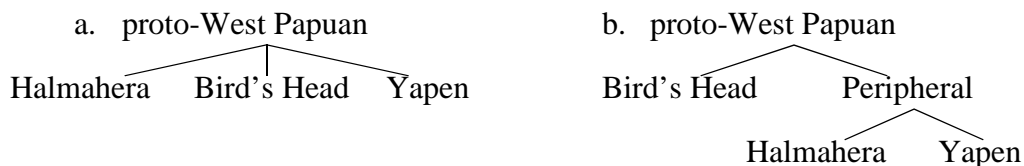
1SG	<i>e-</i>	1PL	<i>i-</i>
2SG	<i>o-</i>	2PL	<i>u-</i>
3SG	<i>a-</i>		

Establishing any sort of correspondence between these purely vocalic *free* pronominal forms and the bound morphology discussed in section 4 is highly unlikely. In short, there do not appear to be any compelling reasons to accept the Geelvink Bay phylum hypothesis.

8. The position of Yapen in a ‘West Papuan Phylum’

The genetic model proposed assumes one of the following arrangements of the West Papuan languages:

Figure 1. Higher-order relationships of the West Papuan languages: two models



I have only indicated the possibilities for macro-level relatedness. Details of the relationships between the Halmahera languages can be found in Voorhoeve 1988; a more recent assessment of the relationships existing (or not existing) between the Bird's Head languages can be found in Reesink 1996. At the language level the Yapen languages show a two-way split between Yawa and Saweru; details of Yawa dialects can be found in Larry Jones 1986.

We need to add some cautions before accepting either of the family trees in figure 1. Reesink (1996) notes that there is very little, if any, compelling lexical evidence to suggest the coherence of the Bird's Head language, and suggests that the whole peninsula might in fact represent a linguistic area, in which common pronominal morphology has been grafted on to a previously diverse group of languages. It is accepted that the languages of the eastern Bird's Head such as Meyah and Hatam are unrelated to the central- and western languages (such as Mai Brat, Tehit and Moi), despite their typological similarities. If Reesink's speculations are proven to be correct, then another possibility emerges to explain the 'irregular' features of the Bird's Head languages within the West Papuan Phylum: they are not originally members of this family. Under this scenario only the Halmahera and Yapen languages represent the 'original' West Papuan grouping, with superstratal influence from a considerable time depth affecting the (northern) Bird's Head area, which display some of the characteristics of the Halmahera-Yapen languages, but not in a regular fashion. Whether the homeland of this non-Bird's Head grouping was in Yapen or in Halmahera cannot yet be determined: the fact that the southern half of Halmahera, and both the eastern and western thirds of Yapen, are now occupied by speakers of Austronesian languages means that it is likely that much of the linguistic prehistory has been lost to us. The fact that a (relatively) large part of the lexicon of Yapen island Austronesian languages consists of forms that have no cognates in Austronesian languages outside the Yapen area, including a lack of cognates in Wandamen, a close relative of these languages, points to extensive early contact between the

Austronesian and non-Austronesian peoples of the island. It is noteworthy that the shape of prefixation in the Yapen languages, both Austronesian and non-Austronesian, varies between a single syllable and three syllables of length; in the Halmahera languages prefixes are typically a single uniform syllable in length. While syllabic agreement marking is not in itself remarkable, it is unusual in this part of the world given the widespread areal preference for S,A prefixal agreement on the verb to be marked by means of monoconsonantal prefixes, a tendency that is found between (roughly) eastern Flores and the eastern Torricelli languages, throughout Maluku and along the north coast of New Guinea, in both Austronesian and non-Austronesian languages.

The subgrouping of the two Yapen languages together as a new branch, rather than trying to fit them into one of the other branches of West Papuan, arises from what are apparently regular correspondences between forms in Yapen and Halmahera, which indicate common innovations on the part of the Yapen languages. We have seen that the Yapen 1SG.SUBJ prefixes are plausibly related to the Halmahera and Bird's Head ones, but require a $*t > **ʃ$ change (and subsequent $*ʃ > \emptyset$ in Saweru, as regularly attested).

9. Conclusions

In sum, we have seen that, rather than being classified together with the languages of the Waropen hinterland in a 'Geelvink Bay Phylum', the languages of Yapen show every sign of being genetically related to the languages of northern Halmahera. This classification is in line with Cowan's early assumptions, which were based on very little data, but it contradicts Voorhoeve's (1975) reclassification, which was primarily based on apparent correspondences in a few lexical items. The main reasons for the discrepancy between Voorhoeve's account and this account lies in the kind of pronominal data employed: while Voorhoeve was forced to rely on reports of free pronominal forms, which are unreliable in this area, the current study reports on bound morphological paradigms. Access to the morphological structure of many of the lexical items that Voorhoeve compared with Geelvink Bay family languages, such as Tarunggare and Baropasi, shows that apparent cognates are not convincing cognates, once non-cognate morphology is stripped away. Examining more recent data on the bound pronominal morphology of Bauzi shows no correspondences between the Bauzi agreement markers and those in the Yapen languages.

On the other hand, the cognacy between bound morphological forms in the Yapen languages and the Halmahera languages is striking, and, taking the membership of the basic pronouns (first, second and third person, singular and non-singular, feminine and non-feminine for the third person singular) as a set of seven, and allowing both subject and object pronouns, we find convincing matches in six of the fourteen forms, and a likely correspondence ($wo:\textcolor{blue}{f}o$) in the 3SG.M.SUBJ form. The matches with the Bird's Head languages are less convincing, probably because of the status of the Bird's Head languages, which might or might not be genetically related to the Halmahera languages (for a critique, see Reesink 1996).

This study has shown the importance of careful quantifiable investigation when trying to unravel the conflicting evidence of what are rare (at least in this area) structural features, embedded in a slew of areal features. The methodology of Nichols (1996) has proven to be robust, when the methodology is applied rigorously: a closed set of forms was examined, and all of the members of the set were examined as a group. While the results of this comparison

were not overwhelmingly convincing, they are suggestive of a possible relationship. Combining these factors with the grammatical features mentioned in section 6, which are rare features cross-linguistically (Nichols 1993), we find a strong indication of genetic relatedness. All that remains is a detailed lexical comparison, and that must wait on the lexical materials from the different groups becoming available first. But this short squib has perhaps pointed the way to a direction for fruitful comparative work.

References

- ANCEAUX, J. C. 1961. *The linguistic situation in the islands of Yapen, Kurudu, Nau and Miosnum*. Verhandelingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde, deel 35. Martinus Nijhoff.
- AYERI, ALFONS, ALFRED KARURI, NOAH MAFIA and MARK DONOHUE. 2003. *Ayawo Yafan: pengalaman dari Pulau Saweru*. [stories from Saweru]. National University of Singapore.
- BRILEY, DAVID. 1997. Four grammatical marking systems in Bauzi. In Karl Franklin, ed., *Papers in Papuan linguistics No. 2*: 1-131. Canberra: Pacific Linguistics A-85.
- BROWN, JANE. 1999. Mai Brat. MS, Summer Institute of Linguistics, Irian Jaya branch, Indonesia.
- BROWN, WILLIAM U. 1990. Maybrat nominal phrases. *NUSA* 32: 43-61.
1991. A quantitative phonology of Mai Brat. In Tom Dutton, ed., *Papers in Papuan Linguistics No. 1*: 1-27. Pacific Linguistics A-73.
- CAPELL, A. 1975. The "West Papuan Phylum": general, and Timor and areas further west. In *New Guinea area languages and language study Vol. 1: Papuan languages and the New Guinea linguistic scene*: 667-716. Canberra: Pacific Linguistics C-38.
- COWAN, H.K.J. 1958. A large Papuan phylum in West New Guinea. *Oceania* 28: 159-166.
1960. Nadere gegevens betreffende de verbreiding der West-Papoease taal-groep (Vogelkop; Nieuw Guinea). *Bijdragen tot de taal-, land- en volkenkunde* 116: 350-365.
- DIFFLOTH, GÉRARD. 1994. The lexical evidence for Austric, so far. *Oceanic Linguistics* 33 (2): 309-321.
- DOL, PHILOMENA. 1999. *A grammar of Maybrat: a language of the Bird's Head, Irian Jaya, Indonesia*. Leiden: Grafisch Bedrijf, Universiteit Leiden
- DONOHUE, MARK. 2001. Split intransitivity and Saweru. *Oceanic Linguistics* 40 (2): 321-336.
- DONOHUE, MARK. 2002. *Papuan Malay*. To appear in David Gil, ed., *NUSA: studies in Indonesian and miscellaneous languages of Indonesia*. Jakarta: Universitas Katolik Atma Jaya.
- Available online at <http://courses.nus.edu.sg/course/ellmd/PM/PapuanMalay.pdf>.
- FLASSY, DON AUGUSTHINUS LAMAECH. 1991. *Grammar sketch of Tehit, a Toror language, the West Doberai peninsula, New Guinea (Irian Jaya)*. Thesis, Faculteit der Letteren, Vakgroep Zuid-oost Azië en Oceanië, Rijksuniversiteit te Leiden.
- GRAVELLE, GILLES. *Meyah morphology and syntax*. MA thesis, The Northern Territory University.
- JONES, LARRY B. 1986. The dialects of Yawa. *Papers in New Guinea Linguistics No. 25*: 31-68. Canberra: Pacific Linguistics A-74.
- JONES, LINDA K. 1986a. The question of ergativity in Yawa, a Papuan language. *Australian Journal of Linguistics* 6: 37-55.
- 1986b. Yawa phonology. *Papers in New Guinea Linguistics No. 25*: 1-30. Canberra: Pacific Linguistics A-74.
- nd. *Notes on the phonology of the Rosbori dialect of Yawa*. MS, Summer Institute of Linguistics, Irian Jaya branch.
- HESSE, RONALD GERHARD. 1983. *Imyan Tehit phonology*. MA thesis, University of North Dakota.
- MIEDEMA, J. and F.I. WELLING. 1985. Field notes on languages and dialects in the Kebar district, Bird's Head, Irian Jaya. *Papers in New Guinea Linguistics No. 22*: 29-52. Canberra: Pacific Linguistics.

- NICHOLS, JOHANNA. 1993. *Linguistic diversity in space and time*. University of Chicago Press.
1996. The comparative method as heuristic. In Mark Durie and Malcolm Ross, eds., *The comparative method reviewed: regularity and irregularity in language change*: 39-71. Oxford University Press.
- REESINK, GER P. 1996. Morpho-syntactic features of the Bird's Head languages. *NUSA* 40: 1-20.
1999. *A grammar of Hatam, Bird's Head Peninsula, Irian Jaya*. Canberra: Pacific Linguistics C-146
- SHELDEN, HOWARD. 1991. Galela pronominal verb prefixes. In *Papers in Papuan Linguistics No. 1*: 161-175. Canberra: Pacific Linguistics A-73.
1998. *Transitivity and Galela pronominal reference*. SIL Electronic Working Papers No. 5. Available at: <http://www.sil.org/silewp/1998/005/SILEWP1998-005.html>.
- VAN STADEN, MIRIAM. 2000. *Tidore: a linguistic description of a language of the North Moluccas*. PhD thesis, Leiden University.
- STOKHOF, W.A.L. and DON A.L. FLASSY. 1985. A recently discovered M(o)oi vocabulary in the national museum (Jakarta). *Papers in New Guinea Linguistics* 22: 53-130. Canberra: Pacific Linguistics A-63.
- VOORHOEVE, C.L. 1975a. West Papuan Phylum languages on the mainland of New Guinea: Bird's Head (Vogelkop) peninsula. In S.A. Wurm, ed., *New Guinea area languages and language study Vol. 1: Papuan languages and the New Guinea linguistic scene*: 717-728. Canberra: Pacific Linguistics C-38.
- 1975b. East Bird's Head, Geelvink Bay phyla. In S.A. Wurm, ed., *New Guinea area languages and language study Vol. 1: Papuan languages and the New Guinea linguistic scene*: 867-878. Canberra: Pacific Linguistics C-38.
1982. The West Makian language, North Moluccas, Indonesia: a fieldwork report. In C.L. Voorhoeve, ed., *The Makian languages and their neighbours*: 1-74. Canberra: Pacific Linguistics D-46.
1988. The languages of the North Halmahera stock. In *Papers in New Guinea linguistics*: 181-209. Canberra: Pacific Linguistics A-76.
- VISSER, L. E. and C.L. VOORHOEVE. 1987. *Sahu-Indonesian-English dictionary and Sahu grammar sketch*. Dordrecht: Foris Publications
- WIMBISH, SANDRA GAY. 1991. *An introduction to Pagu through the analysis of narrative discourse*. MA thesis, University of Texas at Arlington.