**Chapter 40. Kinship semantics in Papuan languages**

Nicholas Evans, Bruno Olsson & Wolfgang Barth

**1. Introduction**

Kinship terminology is a crossroad of language and culture. As such, systems of kinship terminology have been studied both by linguists and by anthropologists, with the result that we have a relatively large set of data to draw on – over 100 Papuan systems – compared to most other linguistic domains. [Once we’ve assembled our data, add sentence or two of comments on whether the diversity of kinship systems is as great as that found in other typological domains.] Kinship terminology is typically also the most recursive and systematically structured part of a language’s vocabulary, exhibiting a number of cross-cutting semantic features, such as generation, sex of referent, consanguineal vs affinal kin, lineal vs collateral relatives. The juridically reckoned nature of kinship in many societies results in the transmission of genealogies going back many generations, so that in principle someone like your father’s father’s father’s father’s brother’s son’s son’s son’s son is your kin. The recursiveness of such reckoning (here chaining together multiple instances of the ‘father of’ relationship) generates huge numbers of *kin types* (logically distinguishable points in the genealogy) running to the thousands once we get out to e.g. great-great-great-great-great-grandparents and seventh cousins. No language in the world has distinct *kin terms* for all *kin types*. Thus ‘mother’s mother’ (MM) and ‘father’s mother’ (FM) are distinct *kin types*, but a single English *kin term* lumps them together: *grandmother*. Likewise, the kin types ‘mother’s sister’ (MZ) and ‘father’s sister’ (FZ) are grouped together under the kin term *aunt*. The lumpings together of multiple kin types under a single kin term are generally known as *equivalences* or *syncretisms*; we will generally use the latter term here. Much of the typological variation between kinship system can be understood in terms of different patternings of syncretism. Usually syncretisms are treated as absolute – i.e. there is a binary opposition between terms being identical (*aunt* ‘MZ’ = *aunt* ‘FZ’) or different, regardless of whether the difference is total (Russian *otec* ‘F’, *djadja* ‘FB, MB’) or partial (Spanish *hermano* ‘brother’, *hermana* ‘sister’).[[1]](#footnote-1)

Systems of kinship terminology are coupled with other aspects of social structure, though how tightly is a matter for debate. The three most important of these are:

(a) *marriage rules*. For example in a system where you marry your ‘cross-cousin’ (e.g. your father’s sister’s child)[[2]](#footnote-2) then your paternal aunt is also your mother-in-law, often resulting in terminological equivalence (FZ=WM[[3]](#footnote-3)); likewise it is common for systems favouring cross-cousin marriage to lump parallel cousins (e.g. FBC, non-marriageable) as siblings (e.g. FBD=Z, female parallel cousin called ‘sister’), stressing their non-marriageability, while employing a distinct term for cross-cousins (e.g. FZC), who are typically marriageable. In §2.1 we will see an interesting case of a particular marriage-rule – symmetric sister-exchange – correlating with a set of kinship terms, used only by the children born to such exchanges, for the special types of cousins and nuncles[[4]](#footnote-4) this marriage-rule produces.

(b) *descent/filiation rules*, determining who belongs to such key social units as clans. In matrifiliative clan systems, for example, syncretisms are commonly found between father and father’s sister’s son – both males belonging to the father’s matrilineage, whose membership he cannot transmit himself, but which is transmitted through his sister to her children. This particular syncretism is generally called a Crow skewing[[5]](#footnote-5) rule by anthropologists of kinship. Skewing rules of this type tend to correlate with the logic of clan filiation: matriclan systems correlate with Crow skewing rules, and patriclan systems with so-called Omaha skewing rules in which MBC=MB: both share the same patriclan (see Nen example in §2.1).

(c) *residence rules*, primarily whether a couple live with the man’s or the woman’s clan/household after marriage, often impact on where elaboration is found in the kinship system. Co-residents know more about each other’s genealogies and are bound by stricter interactional norms.

(a)-(c) have long made kinship an area of language used for arguments about social structure, current or reconstructed, though the arguments rest on the assumption that the relationship of vocabulary to society is tightly coupled – an assumption that is not always uncontroversial (insert 2-3 refs). However, full considerations of these links would take us too far afield in this chapter, so we only make passing references to them where they seem particularly salient.

Aside from their anthropological import, kinship terms display many special linguistic features (Dahl & Koptjevskaja-Tamm 2001).

As two-place predicates (logically one cannot *be a father* without *being a father to someone*) they are realised as transitive verbs in many languages (Evans 2000), although this phenomenon does not appear in any Papuan language to our knowledge. However, another manifestation of their logical status as two-place predicates is that in many languages kin terms are obligatorily possessed, thereby ensuring that both arguments are brought in: the possessor, expressed by the possessor affix or NP, and the referent of the expression. As a result, kinship terms often have special possessive morphology, and it is not uncommon for this to result in suppletion according to the person of the possessor, e.g. Kaluli *na:la:* ‘my daughter’, *ga:la:* ‘your daughter’ versus *ida:* ‘her/his daughter’ (Grosh & Grosh 2004) – see Baerman (2014) for a survey.

The frequency with which humans are referred to using kinship terms in many societies has led to the grammaticalisation of various specialised structures, such as kinship dyad terms for referring to pairs of kin, e.g. father and son, mother and daughter – see Ch. 49, §5). There are also a few languages with special pronouns reflecting specific kinship relationships. Consider these Nagovisi pronouns (Nash 1974), which encode the kinship relations between members of the group: *nɛ* ‘we’, *nɛnabora* ‘we [FZ/BC; HM/SW]’, *nɛnamisira* ‘we[BW/HZ]’, *nii* ‘we [WM/DH]’, *ninga* ‘we [relationship uncertain, strangers], or these pronouns from Menya (Whitehead 2004): *yä-mät-qiye* ‘we two [F/S], *qe-mät-aŋgi* ‘you two [F/S]’; in each case these are just some forms from a large paradigmatic set. Significantly, in Nagovisi the dyads encoded in these pronouns dictate behavioural norms (Nash 1974: 48-48), e.g. friendly, informal relationship between mother’s brother and sister’s daughter, joint economic pursuits between sisters, or restraint in regard to sexual matters between brothers and sisters. These dyad-based pronouns thus evoke particular normative behaviours within the named pair, in addition to simply designating kin relationships, in line with Strathern’s (1988) influential notion of Melanesian personhood as ‘partible’, composed out of a number of ‘dividuals’, each reflecting particular patterns of relationship.

Implicit in everything that has been said so far is that the notion of ‘kinship’ is self-evident, as a system of social reckoning based on the biological facts of reproduction and descent. This neat assumption is not always unproblematic. On the one hand, many languages extend the formal properties of kinship terms to other types of social relationship, such as the Marind term KKK for ‘person born on the day as someone else’ [Bruno? Can you insert?] – should these be treated as ‘emic kin’ in such a language? More fundamentally, the notion that kinship is primarily rooted in biological reproduction has been challenged by anthropologists such as Bamford (2009).[[6]](#footnote-6) We bracket these concerns aside here, and concentrate on the large quantum of material where meaningful comparisons of lexical semantics can be made, across a very large number of languages. It is these structural patternings of conceptual equivalence that will occupy us for the rest of the chapter.

One final remark is needed. There tends to be a relationship between individual syncretisms (e.g. F=FB [same term for father and his brother]) on the one hand, and, on the other, whole kinship systems which link multiple syncretisms in a sort of ‘harmonic logic’ by which, for example, syncretisms at the parent/nuncle level (say to treat parents and their same-sex siblings as equivalent but distinguish them from parents’ opposite-sex siblings) correlate with syncretisms at the sibling/cousin level (e.g. treating parallel cousins as equivalent to siblings, but different from cross-cousins). The belief that most kinship systems exhibit tight couplings between different parts of their systems lies behind what we may call the ‘holistic typologies’ of kinship systems, employed by the great kinship theorists from Morgan in the C19 to Murdock and others in the C20 – these use terms like ‘Eskimo system’, ‘Hawaiian system’, ‘Sudanese system’ and so forth as shorthand for assemblages of many different syncretisms in the whole system. For that reason, it is useful to begin with some ‘system cameos’, which we do in §2, before looking at some selected individual syncretisms.

Whether things always correlate so smoothly is an open, empirical question (see Passmore et al 2021 for a critical view). In this chapter we adopt a more agnostic method that treats each possible syncretism – between any two or more kin types – as a logically independent variable, even though in practice clusters of syncretisms tend to co-vary. We draw our data from the syncretism database known as Kinbank (www.kinbank.net), a sub-bank of Parabank (insert URL) which assembles data on syncretistic patterning in multidimensionally structurable semantic fields (pronouns, kin terms etc.) for what is now over a thousand languages from around the world. In this paper we make use of the Papuan subset of this database, which includes 112 Papuan languages from 36 maximal clades.

We divide the rest of this chapter into two parts, one focussing on depth and the other on breadth: §2, focusses on systems and their internal logic through four ‘kinship snapshots’, each aiming to illustrate the workings of the overall system in a particular language. In §3 we pull out a subset of interesting syncretisms and examine these across a much broader sample, using the data from kinbank, looking at which syncretisms are common and distinctive among Papuan languages, which are confined to particular Papuan families or areas, and what similarities, if any, they show to other major language families. Finally in §4 we close with some questions for future research.

**2. Four sample systems**

We begin with four ‘system cameos’, to give a more holistic and culturally contextualised snapshot of how some rather overall systems work, before taking apart some of the more interesting components of the system in the rest of the article.

**2.1 Watam**

Watam (Foley 1997) is a language of the Ramu-Lower Sepik family. The Watam kinship system includes terms stretching over nine generations. The most distant two pairs of generations (±4, ±3) have self-reciprocal terms: *bijir* applies equally to great-great-grandparents and great-great-grandchildren, while *ŋgamar* applies equally to great-grandparents and great-grandchildren. The generations closer to ego (+2/-2) are no longer self-reciprocal: all grandparents are *nenkai* and all grandchildren are *rumbun*.

The +1 generation makes four distinctions. Father and all his male blood relatives of that generation are grouped together as *aes* and *aem* groups mother and all her female blood relatives of that generation. Then there is a term for father’s sister (*namkwae*) and one for mother’s brother (*akwae*). This is typical of ‘bifurcate merging’ or ‘Iroquoian’ systems. Typologies of kinship, as mentioned above, generally presupposed a logical consistency or harmony between how terms pattern in different generations. In canonical Iroquoian systems, for example, which treat ‘cross-uncle’ (MB) and ‘cross-aunt’ (FZ) differently from ‘parallel uncle’ (FB) and ‘parallel aunt’ (MZ), who are grouped with one’s parents, there is an assumed correlation in ego’s generation whereby parallel cousins would be grouped with siblings (e.g. FBS=B) but cross-cousins would be given distinct terms (e.g. MBS≠B). In Watam, however, this is not the case: there is no differentiation between siblings and cousins on either side of the family. Rather, everybody in this generation is called *yakai* ‘elder same sex sib’, *yap* ‘younger same sex sib’, *mbi* ‘mZ’, or *ondaŋ* ‘fB’. This pattern is characteristic of a ‘Hawaiian’ system which canonically merges parents, uncles and aunts in the +1 generation and siblings and cousins in ego’s generation. Classic typologies would thus have to say Watam splits between an ‘Iroquoian’-type system in the +1 generation and a ‘Hawaiian’ system in the 0 generation.[[7]](#footnote-7)

One generation below ego, speakers of Watam differentiate son (*iniŋ*) and daughter (*namoŋ*). Most children of ego’s siblings are referred to in the same way. However, mother’s brother calls his sister’s children *amuk* while they call him *akwae*.   NE to WB: what does FZ call her kids, i.e. fBC?

[still needs to add material on affines]

**2.2 Ekagi**

Ekari (also known as Ekagi, Mee or Kapauku) is a Paniai Lakes (TNG) language of the West Papuan highlands. The description given here follows Pospisil’s (1960, 1980) analysis of the Ekari kinship system.[[8]](#footnote-8) Note that Pospisil’s glosses of kin terms are often in conflict with those given in Steltenpool’s (1969) dictionary, a source that is disregarded here. The kin terms are cited here in their 1st person forms, which depending on the lexeme either take preposed 1sg pronoun *ani*, or prefix *n-*. The Ekari kinship terminology showcases several common themes in kinship systems from across the world (such as the Iroquois cousin system) with a few interesting wrinkles. As in Watam, it has self-reciprocal terms for great-grandkin (*ani aija*) and great-great-grandkin (*ani pigoka*), extending this to grandkin as well (*ani muuma*); these also apply to other kin in those generations (e.g. great-uncles or great-nephews would be *ani aija* like grandparents). It also has emphasis on birth order, sorted by mother in polygynous relationships, and for order of marriage among co-wives. In other respects it remains relatively simple, especially in its affinal inventory (cf. Nen below for the opposite situation). We present starting in ego’s generation and moving stepwise towards the more distant relatives, finishing with the in-laws.

Starting with ego’s siblings, we find a distinction between same- and opposite-sex siblings (with *paneka* denoting the latter, i.e. fB and mZ), with a distinction of relative age added among the same-sex siblings (*nauwa* ‘meB’, *anibai* ‘feZ’ and *ani weneka* ‘myB, fyZ’, i.e. ‘younger same-sex sibling’). Having a relative age distinction that is limited to same-sex siblings is a common typological pattern, and it is also found in Watam (see above). When it comes to cousins, it is a sort of split between an Iroquoian system with regard to cross-cousins (*noone*), and paternal parallel cousins (B=FBS=MZS≠MZS=MBC), but a Sudanese system with regard to maternal parallel cousins (MZB≠B); the maternal parallel cousins (i.e. one’s mother’s sister’s children) are covered by a separate term for MZC *ani ijoka* (making no distinctions of sex or relative age). The relative age distinctions made in sibling terms are carried across into paternal parallel cousins, e.g. *nauwa* denotes an elder paternal same-sex parallel cousin of a male ego, i.e. meFBS, and so on).

Moving up to the +1 generation, we find the parent terms *niikai* M and *naitai* F. The morphologically related terms *niika* and *naita* (dropping the final /-i/, of unclear status) express MZ and FB respectively, and extend to any consanguineal relative of the parents’ generation, so that e.g. *niika* also refers to M’s female cousins (MMZD, MMBD etc.), while *ani ooka* is FZ and *naama* is MB. If we take a simple same vs different approach to syncretism, this is a Sudanese system in the +1 generation (distinguishing all three kinds of male kin, and all three kinds of female kin), while if we take a more gradient measure it comes very close to being a bifurcate-merging system, very common in Papuan languages (and the rest of the world).

The cross-nuncle terms are self-reciprocal with the corresponding niblings (= ‘nephews/nieces’). Thus, I (male or female) refer to my father’s sister as *ani ooka*, and she calls me the same. I refer to my mother’s brother as *naama*, and he uses the same term in return. In the rest of the –1 generation, a general term *ani joka* ‘my child’ exists alongside two sets of birth-order terms: *ibo* ‘first-born son’, *ipouga* ‘2nd son’, etc., and *oumau* ‘1st daughter’, *maga* ‘2nd daughter’ etc. The birth-order terms are reckoned in reference to children born to the same mother, so a man can have as many *ibo* etc. as he has wives. [Any reference to other birth-order term systems in or sample/other Pap langs?]. Birth order also appears, optionally, in grandchild terms: in addition to the self-reciprocal grandkin terms mentioned above, terms for grandchildren can be formed by adding *-pa* to the birth-order terms (e.g. *ibopa* ‘child of first-born son’).

Among the affinal terms, we find the generic *ani* *waka* ‘my spouse’, complemented by marriage-order terms for the wives in a polygynous marriage (*epame* ‘1st wife’, *jupikaame* ‘2nd wife’ etc.). Staying in the 0-level generation, the term (*ani*) *geeto* is used for the spouse of a consanguineal relative of one’s spouse, e.g. WBW, HZH. Two self-reciprocal terms are used between the spouse of an uncle or aunt and the child of a spouse’s sibling: *ani wape* ‘FZH, MZH; WZC, WBC’ and *naamai* ‘FBW, MBW; HBC, HZC’. (Note that *naamai* and the self-reciprocal *naama* ‘MB; ZC’appear to show the same formal relationship as the parent and MZ/FB terms, differing only in the final /-i/.) These terms extend collaterally, so that e.g. *ani wape* also refers to FFBDH or FFZDH, i.e. to one’s father’s cousin’s husband, and so on.

Rather strikingly, the remainder of the affinal kin space is populated by only two terms, *ani geeka* and *ani baaka*, which makes the Ekari affinal system one of the least complex affine terminologies in our database. The first term, *ani geeka*, is used for referring to one’s opposite-sex sibling-in-law, i.e. for ego’s spouse’s same-sex sibling (HB, WZ), the spouses of ego’s same-sex siblings (fZH, mBW), and, more generally, to the spouses of any same-generation, same-sex consanguineal relative (e.g. fFBDH). All other in-laws are covered by the term *ani* *baaka*, which includes in its rather disparate membership set: (a) same-sex siblings-in-law (mZH, fBW), and more generally, same-generation same-sex in-laws (e.g. mFBDH); (b) off-generation in-laws, i.e. in-laws in ascending or descending generations, e.g. WF, WM, HM, WMB, HFFBS, etc.; and (c) the spouses of off-generation in-laws, e.g. WMBW, WBDH, etc. The logic behind this bipartition of the affinal space, as presented in Pospisil’s (1980) analysis, is that *ani geeka* is a marriageable in-law, whereas *ani baaka* denotes the non-marriageable in-laws. Accordingly, the preferred choice for a second wife would be either a sister of the first wife (in a sororate marriage) or a deceased brother’s wife (in a levirate marriage), both of which are referred to as *ani geeka*.

**2.3 Nagovisi**

Southern Bougainville, where Nagovisi is spoken (cf Ch. XX), is renowned as an area of entrenched matriliny, in contrast to the generally patrilineal organisation of most Papuan groups. Dravidian systems, which prescribe marriage with one’s cross-cousin (MBD, FZD, from a male viewpoint), are rare in Papuan languages, to the extent that Scheffler (1971:) claimed there were no reliable reports of Dravidian systems in New Guinea, but in actual fact all four languages of the South Bougainville family (Nasioi, Nagovisi, Siuai and Buin) have Dravidian systems (Hage 2006; Thurnwald 1910, Rivers 1914, Oliver 1955, Rausch 1912), and in this respect resemble the majority of languages in Australia but differ markedly from other Papuan languages. Although the terms for husband and wife are distinct from the cross-cousin terms, the Dravidian nature of these systems is shown by the syncretism of MB=HF (*papa* in Nagovisi) and FZ=WM (*kabo*) since, under cross-cousin marriage, these are potentially the same person in each case.

Traditional Nagovisi social organisation (Nash 1971, 1974) involves a number of nested groupings at different levels, all matrilineal: moieties[[9]](#footnote-9) (*manka* ‘eagle’ vs *komo* ‘hornbill’), clans, and (as subdivisions of some clans) lineages. For each of these, membership is inherited from one’s mother. There are also various other terms for matrilineal groupings: *madawo* ‘matrilineal kin (indefinite range), *malo/matalo* ‘father’s matrilineal kin’, *maniku* ‘wives and daughters of a descent group’, *nuga* ‘male matrilineal kin of wives and daughters’, and *motai* ‘husbands and fathers of a descent group, i.e. men who have married in’.

Note that the logic of exogamy – marrying outside one’s moiety – dictates that both your parallel grandparents are in the same moiety to each other, and to you. For MM this is simply through descent. But for FF, it is because if e.g. you and your mother are ‘eagle’ matrimoiety, your mother must have married a ‘hornbill’ man’; his mother must therefore have been ‘hornbill’, so she must have married an ‘eagle’ man, placing your FF back in the same matrimoiety as you. This creates a ‘two-generation return’ to each moiety, a feature that will become relevant below. Similar logic places your two ‘cross grandparents’ – MF and FM – in the opposite moiety to you.

The first striking property of Nagovisi kin terms is that, like many Australian systems, it displays a patterning that conflates kin two generations apart in the same direction, as long as they are in the same moiety. Thus *ngo* is not just M, MZ, and FBW – women in your own moiety, one generation up – but also MMM and FFM – women in your own moiety, three generations up. Similarly *mma* is not just F, FB and MZH – your father and other men one generation up from you, in his moiety – but also FFF and MMF – men in your father’s moiety, three generations up from you.

Similar logic applies to your cross-nuncles: *papa* is not just your MB and FZH (i.e. men in your mother’s moiety, one generation up) but also your FMF and MFF (also men in your mother’s moiety, but now three generations up), and *kabo* is not just your FZ and MBW (women in your father’s moiety, one generation up) but also your FMM and MFM (i.e. women in your father’s moiety, three generations up).

Turning from odd-generation to even-generation kin, the grandkin terms have some interesting properties, related to the conflation of kin two generations apart that we have just seen. The only grandparent term confined to its own generation is MM (as befits a matrilineal system); this extends to FFZ, who is also a +2 generation female in your matrimoiety. All the other grandparent terms conflate with older siblings or siblings-in-law in ego’s generation: FF = eB, FM =eZ,[[10]](#footnote-10) and MF = HeB. On the other hand, younger siblings in one’s own generation share terms with relatives in the –2 generation: *inalaman* yB, FByS, MZyS also means SS, *inalamanda* yZ, FByC, MZyD also means SD, and *inobe* WyZ also means DS and DD.

As can be seen from the foregoing, there are four sib terms, following the world’s commonest sib term pattern in distinguishing eB, eZ, yB and yZ. However, once one talks of classificatory rather than close siblings, the pattern of sibling terminology changes, and there is a term imari ‘opposite-sex classificatory sibling or cross-cousin’.

These examples show what appears to be an unshakeable consistency in the Nagovisi kinship system, organised around matrimoieties as the boundaries of semantic extension, and regular groupings-together across even-numbered generations. However, we should bear in mind Sapir’s famous dictum (Sapir 1921:39): ‘Unfortunately, or luckily, no language is tyrannically consistent. All grammars leak’. The chink in the Nagovisi system involves sibling-in-laws. *Mama* is not just eZ (same moiety) but also WeZ (opposite moiety), tata not just eB (same moiety) but also HeB (opposite moiety), and *inalaman* not just yB (same moiety) but also HyB (opposite moiety). Generalisation: sibling terms can be extended to one’s sibling-in-laws, their opposite-moiety status notwithstanding. A possible hypothesis is that is an effect confined to address terms, where consanguineal terms are often hospitably extended to relevant affines (e.g. addressing one’s mother-in-law as ‘mother’ in many cultures); more work is needed on the pragmatics of kin term use in Nagovisi to determine whether this, or some other factor, lies behind the discrepancy.

**2.4 Nen and Komnzo**

Nen (Evans 2019) and Komnzo (Döhler 2018) of the Morehead district, Southern New Guinea, belong to different branches of the Yam family – the Nambu and Tonda branches respectively – but will be taken together here not only because they exhibit so many similarities, including many cognate forms and parallel semantic structures, but also because they show how multiple languages can be linked together in a regional system of intermarriage which tends to reinforce structural similarities. In the case of the Morehead district the most salient regional features are (a) marriage by direct exchange of sisters across patrilineal clans, so that each woman marries the brother of the other, (b) the existence of three ‘moieties’ which aggregate the dozens of clans in the region into three higher-level groupings for purposes of marriage and alliance, and (c) the absorption of women into their husband’s clan upon marriage (see Ayres 1983 and Williams 1936 for classic ethnographies). There is a strong patrilineal bias: knowledgeable Nen men can recite clan genealogies back nine patrifiliative generations,[[11]](#footnote-11) whereas often people do not know the clan of their mother’s mother since there is no independent practice of talking about matrilineal descent. Beyond two generations no specialised kin terms are found.

Starting with direct lineal kin, in both languages siblings are primarily organised by relative age (N *ani*, K *nane* ‘elder sib’; N *nəŋgən*, K *nəŋgəθ* ‘younger sib’), children are simply referred to as ‘(X’s) child’, e.g. N *tande toge* ‘my child (son or daughter)’, and there is a single term for grandkin of all types (N *kake*, K *zaθ* [[12]](#footnote-12)) or the Nama loan *aki* take in FF, FM, MF, MM and CC). The grandkin terms can be reduplicated to denote more distant relatives in either direction – N *kakekake* can mean either ‘ancestors’ or ‘descendants’.

It is in the terms for the parents’ generation, for cousins, and for affines (in-laws) that the more unusual features of these languages appear.

Beginning with the parents’ generation, the terms for fathers and mothers can be extended out to their same-sex siblings (FB and MZ), optionally modified by an adjective like ‘big’ or ‘little’. Thus N *dede* and K *ŋafe* span both F and FB (but FyB in K could be, more precisely, *ŋafe katan* ‘little father, FyB’), and N *ama* and K *ŋame* can include both M and MZ, again optionally modified e.g. *ŋame katan* ‘little mother, MyZ’. A point whose significance will become clear shortly is that F and FB belong to the same clan, and likewise for M and MZ; in ego’s and parents’ generations the more precise kin terms, even while extending across several kin types, will always stay within members of the same clan.

Working out to the remaining uncles and aunts, who are all cross-kin[[13]](#footnote-13), we encounter specialised terms, and these in turn differentiate according to whether one can trace the relationship just through one parent, the ‘general’ situation, or whether one can trace the relationship through both parents as a result of sister-exchange in the parents’ generation being involved. Thus for MB we have N *baba* (general) or *mitarbe* (in the case of a consummated sister exchange, which means that my MB is also my FZH) and in K *ŋäwi/babai* (general) or *fäŋafe* (in the case of a consummated sister exchange). For FZ we have N *babale* (general) or *mitadma* (in the case of a consummated sister exchange) and K *babai* (general) or *fäŋäme* (after a consummated sister exchange).[[14]](#footnote-14) Komnzo has /fäŋame/ for FZ (after exchange), not /fäŋafe/. It is probably a contraction of /fäms ŋafe/ "exchange father" and /fäms ŋame/ "exchange mother". Both of these are also somewhat archaic and most people use /bäiŋaf/ and /bäiŋam/ (which are the words found in Wära, next door). These terms are all self-reciprocal, so that *baba* or *ŋäwi*, for example, mean either ‘FZ’ or ‘ZC’, i.e. niece through (a man’s) sister.

We can show the logic of these kin terms, and others arising through consummated sister-exchange, though the following diagram showing part of the Komnzo system (Döhler 2018:31) from the point of view of a child (dark square) whose parents married as a result of direct sister exchange. This traces alternate paths of relatedness to those in the parents’ generation. *Fäŋafe* can be traced wither as my mother’s brother (red path) or as my father’s sister’s husband (blue path), and *fäŋame* as my father’s sister or as my my mother’s brother’s wife. *Yamit* (N equivalent *miti*) is used between the cousins born of such a symmetrical exchange – we could call these ‘double cross-cousins’ because there are two possible paths (e.g. MBC or FZC) that could be used to trace the cross-cousin relationship.

Diagram

Description automatically generated

Fig. 1. Illustration of how sister exchange in the parents’ generation evokes the special kin terms *fäŋafe, fäŋame* (uncle, aunt) and *yamit* (cousin), each limited to the case of kin where sister-exchange has created a binuclear family with two descent paths.

This brings us to the other cousins. As we saw in §2.2 with Ekagi, the Nen and Komnzo system doesn’t fit neatly into any of the classic four typologies – Eskimo (the English type, where Sib≠FBC=FZC=MBC=MZC), Dravidian (where FBC and MZC count as siblings, but the cross-cousins MBC, FZC etc. have their own special term), Sudanese (where each possible type of cousin has their own special term, all distinct from sibling terms) and Hawaiian (where all cousins count as siblings). It will help understand the logic of why Nen and Komnzo deviate from these systems if you bear in mind that the main determinant of how cousin terms work is how the intersection of descent and marriage rules impact upon clan membership.

Thus father’s brother’s children (patrilateral parallel cousins) are known by the relevant sibling terms, since they are in the same clan as you, because your father and his brother are in the same clan and each transmits this to his children.

But your mother’s sister’s children are NOT called as siblings, since (normally) your mother and her sister would marry men from different clans, so their children would likewise belong to different clans. Rather, in Nen this sort of cousin calls each other by terms whose basic meaning is ‘wife’s sister’s husband’, i.e. it is primarily a term used between two men who draw their wives from the same clan. This usage is then passed on to the next generation, so that this sort of cousin calls each other by the affinal term that would be used between their fathers, regardless of the fact that it is actually their mothers who are most closely related: in N the term is *nako* or *yakali* (though many maintain *nako* is a loanword from Suki)[[15]](#footnote-15). In Komnzo there is a slight formal distinction between the brother-in-law and cousin terms: the cousins’ fathers would call each other *nakum* (mWZH) while the term *naku* is between the cousins who are children of these men.

In both Nen and Komnzo, the children of your MB – if he has not married your FZ, creating the situation for the special double-cross cousin terms described above – are simply known by the same name as him, N *baba* and K *babai*, so that MBC = MB, an ‘Omaha skewing’ (see §1). Again we see the operation of clan-tracking logic: your relation to this type of cousin is centred on your relationship to the clan relationship which his father (your MB) puts you in, namely your mother’s birth clan.

Summarising the logic of cousin terminology: your FBC, in the same clan as you, are treated as your siblings; your MZC get called by the term your F calls his WZH, reproducing the affinal relationship between clans at the level of your F; your MBC is equated to your MB, which is the clan your M was born into (in Nen, but not Komnzo, she would have changed into your clan on marrying your F). And, just in the case of your cousins born through a consummated sister exchange, there are special terms, symbolising the fact that this binuclear family sets up a special link between the two clans.

It will already be clear that affinal (marriage) relationships play a key part in the system – no surprise, given the general importance of growing ‘ropes’ (wider relationships) and using them to build alliances in Melanesian societies. We have already mentioned the important cluster of terms focussed on the relationship between men married to sisters – *nakum* in K, *nako* or *yakali* in N, with *nako* then extended down in Nen to the relation between their children, whereas in K there is a distinct though formally related term *naku* for the latter. Both etymological sets – the *naku* set and the *yakali* set – have widespread currency across southern New Guinea, including outside the Yam family, and while we have yet to untangle the whole etymological web, it seems likely that this reflects a pattern of borrowing linked to exogamous, bilingual marriage. Beyond this term (and the special double cross-cousin terms, which are simultaneously consanguineal and affinal depending on the path traced), key affinal terms are: N *kamat[[16]](#footnote-16)*, K *ngom* WB (if they are not married to each other’s sisters); N *tampre*, K *fäms* ‘brother-in-law through consummated sister-exchange; WB=ZH’, N *tanat*, K *enat* ‘parent-in-law/son-in-law’, and N *yézeg* (K?[[17]](#footnote-17)) ‘co-wife, HW’.

A final indication of the importance of sister exchange in the kinship comes from contextual shifts in the meanings of the Nen words *sakr* or *är sakr* ‘boy’ and *mleg* ‘girl’. We mentioned above that Nen and Komnzo sib terms are organised on an ‘older sibling’ vs ‘younger sibling’ principle, regardless of sex. But to specify sex, a woman may refer to her brother as *tande är sakr* (‘my boy’) and a man may refer to his sister as *tande mleg* (‘my girl’). However, these are not simply gender-specified sibling terms in the English style, since if a parent uses them, in which case tande är sakr means ‘my son’ and tande mleg means ‘my daughter’. The key to understanding this usage is that the constant is ‘my boy/girl who will be used as a male/female exchange partner, in marriage’. This example illustrates one of the methodological problems in comparing data across languages, since whether to include these terms as alternative sib terms (♂︎Z, ♀︎B) or as a vaguer contextual extension, as argued for here, impacts on the type of sibling system Nen would be claimed to have.

**2.5. Comparison across the sample systems**

This sample represents just 5 of the 800 or so kinship Papuan systems, though it was deliberately constructed to sample as widely as possible, both geographically (north, west, east, south) and genetically (four distinct families, with two from one family and region as a sort of control). It shows clearly that, in kinship as in so many other respects, Papuan languages show extraordinary diversity. Some of the more important features discussed above are summarised in Table 1.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Language | Watam | Ekagi | Nagovisi | Nen & Komnzo |
| Family; Region | Ramu-Lower Sepik; Sepik | TNG; Pania Lakes, West Papuan highlands | Bougainville Island; South Bougainville family | Yam, Trans-Fly region |
| Patterning in +1 generation | Iroquoian | Sudanese (though M/MZ and F/FB only differ by a vowel) | Iroquoian  (Dravidian) | Iroqoian |
| Sib/cousin patterning in 0 generation | Hawaiian | Split: Iroquoian on paternal side, Sudanese on maternal side | Iroquoian | Split: Iroquoian on paternal side, Sudanese on maternal side, special term for cousins resulting from symmetrical direct sister exchange |
| Consistency across +1 and 0 generations | No | No | Yes | No |
| Sibling dimensions | 4-way: mZ, fB, elder vs younger same-sex sib | 4-way: opposite-sex sib, meB, feZ, younger same-sex sib | 4-way: gender of ref x relative age | Basic older vs younger sib distinction, with secondary possibility of male-referent vs female-referent distinction |
| Patrilineal vs matrilineal social organisation | ? | Patrilineal | Matrilineal | Patrilineal |
| Grandkin | Not self-reciprocal; maternal and paternal subtypes not distinguished | Self-reciprocal; maternal and paternal subtypes not distinguished | Paternal grandparents = older siblings; maternal grandparents = older sib-in-laws | Self-reciprocal; maternal and paternal subtypes not distinguished |
| ±3 and ±4 generation terms? | Yes | Yes | Extend upwards on 2-generation cycle, e.g. M=MMM | No |
| Number of affine terms | [Wolfgang to folow up?] | Spouse (+ order of marriage for cowives); co-sibling in law, uncle-in-law, aunt-in-law; opposite-sex-sibling-in-law; one further vaguer term | Affinal terms generally polysemous with structurally equivalent consanguineal terms, e.g. MB=HF, FZ=WM | Rich set, none overlapping with consanguineals: H, W, WZH, WB=ZH (consummated sister exchange), WB (no sister exchange), WF/DH, HW |
| Self-reciprocal cross-nuncle terms? | No | Yes | No | Yes |
| Skewing? | No? |  | No | Yes, from MB, based on term WZH that he would use to refer to your F |
| Other special features |  | Order terms in cowives, children and grandchildren terms | Extension of sib terms to sib-in-law terms, inconsistently with rest of system | Special terminological sets for relations resulting from symmetrical sister-exchanges |

Table 1. Summary of main features of the four kinship systems surveyed

What Table 1 shows clearly is that NO feature in this table is the same across all languages of our mini-sample. But to misquote Animal Farm, we could say that ‘all Papuan languages are different, but some are more different than others’ – this applies clearly to Nagovisi here, which in many respects is much more like a typical Australian kinship system than a Papuan one, and it has a clear matrilineal social organisation where the others are strongly patrilineal. Even looking at the other languages, great diversity is evident: in the parent’s generation, there are typical Iroquoian systems (Watam), mixed Iroquoian/Sudanese ones (Ekagi, Nen/Komnzo) and the specific subtype of Iroquoian known as Dravidian, where FB=HF (Nagovisi); in ego’s generation there are typical Iroquoian systems (Nagovisi), Hawaiian (Watam), and split systems which are Dravidian on the paternal side but Sudanese on the maternal side (Ekagi, Nen/Komnzo). Self-reciprocal nuncle terms (MB=BC) are found in Ekagi and Nen/Komnzo but not the others; self-reciprocal grandkin terms in Ekagi and Nen/Komnzo but not the others; dedicated terms for great-grandkin, and for great-great-grandkin, in Ekagi and Watam but not the others. Sibling terms, though averaging around a 4-way distinction, make this on a number of different criteria. Nagovisi does not have distinct affine terms, recycling consanguineal terms, but all the others maintain a strict distinction between consanguineal and affinal terms and make rich sets of affinal distinctions.

This short comparison of four cameos should give an initial picture of just how different Papuan kinship systems are. But to make this more systematic – and avoid accusations of cherry-picking – we need to broaden out our sample. To do this in anything less than a multivolume encyclopaedia we need to strategically pick out a few syncretisms and compare them across a range of languages. This is what we do in the next section.

**3. Zooming out to a broader sample**

So far, we have focussed in some detail on just four Papuan systems. While this allows us to get a feel for their overall logic, it has the disadvantage of being just a tiny subsample of the diversity observed in Papuan languages. In this section we do the opposite: we draw on our sample of Papuan languages and contrast their diversity against global patterns. Kinbank is a database containing a global sample of kinship terminology from 1,229 languages, built around 115 kin types. Within this dataset are 112 Papuan languages from 36 maximal clades (Figure 1). We use the Kinbank data, and the Papuan subset, to make broad-based statements about the diversity of kinship terminology in Papuan languages relative to global patterns.

The theoretical possibilities of kinterm systems form a vast design space. In the preceding section we described four unique kinship systems to exhibit Papaun kinship diversity, but there are only four of a possible 4,140 possible sibling organisations (Nerlove & Romney, 1967).[[18]](#footnote-18) If we consider all 115 kin types in Kinbank, then there are 1.2 x 10138 possible systems of kinship terminology. This huge possibility space is mostly unpopulated but allows us to characterise kinship systems on many dimensions. Examining all such dimensions is a vast enterprise, beyond what can be done here. Instead, we examine features of kinship terminology that characterise Papuan kinship, and examine the level of diversity in Papuan kinship, relative to other language groups.

A picture containing black, screenshot

Description automatically generated

**Figure 2.** A map of the 112 Papuan languages from 36 clades stored in Kinbank, with labels for the four system cameos.

**3a. Common features of Papuan kinship**

The most prominent feature in Papuan kinship is the bifurcate merging pattern. As described in the introduction, this pattern separates (bifurcates) parent’s opposite-sex siblings and merges their same-sex siblings with parents, but corresponding patterns of splitting-and-merging can occur through the kinship system (Godelier, 2011). We look at the father/nuncle set, mother/aunt set, brother and male siblings, and sister and female siblings.[[19]](#footnote-19) For each of these three relative sets, there is a “design space” of five possible organisations, of which we observe the maximum of four in both Papuan and non-Papuan languages.[[20]](#footnote-20)

First, we look at the parent/nuncle set, specifically: F = FB ≠ MB (Figure 3). The bifurcate merging system makes up 55% of the total sample of Papuan languages (ntotal = 103), almost double the next most frequent organisation (F ≠ FB ≠ MB). The bifurcate merging pattern only makes up 28% of non-Papuan kinship systems (ntotal = 879). The bifurcate merging system is most prominent amongst TNG languages, found in 21 languages, but occurs at least once in 27 of the 44 language clades for which we have data. The equivalent pattern in mother/aunt terms (M = MZ ≠ FZ) shows a similar level of frequency (45%; n = 45) and diversity (24 of 44 language clades) amongst Papuan languages, and infrequency amongst non-Papuan languages (23%; n = 189).

Within sibling and cousin terms, the corresponding pattern to a bifurcate merging terminology occurs in the form of either B = FBS ≠ MBS or Z = MZD ≠ FZD. Amongst the male set, 60% of Papuan languages contain a bifurcate merging pattern (ntotal = 67), compared to only 17% of the global sample (ntotal = 663). 43% of female cousin systems contain the equivalent pattern (ntotal = 66), compared to only 18% of the global sample (ntotal = 621). As with the parental structures, most languages displaying bifurcate merging pattern are from the Nuclear Trans-New Guinea family, but the bifurcate merging pattern occurs in fewer clades than the equivalent pattern in parents. The male pattern occurs in 18 language clades, and the female in 16. There are fewer Papuan languages that have complete cousin data, which may explain this discrepancy.

Despite the prevalence of the bifurcate merging pattern, it does not appear to be a overarching organising principle of Papuan kinship diversity – and it need not be as we saw in our description of Ekagi. Of the 62 Papuan languages for which we have data, only 16 present a bifurcate merging system in male and female parents, and male and female cousins. Of the 97 languages for which we have father/nuncle and mother/aunt data, 38 have bifurcate merging patterns in both subsets of kin, and 21 show the pattern in either male or female relatives. Of the 64 languages for which we have cousin data, 27 have the bifurcate merging pattern in both male and female cousins, and 13 show the pattern in either male or female cousins. We see similar patterns of heterogeneity when assessing the relationship across generations, but within genders: father/nuncle bifurcate merging systems co-occur with male cousins bifurcate merging systems in 28 of 66 languages, and mother /aunt and female cousin systems in 19 of 62 languages. As was discussed in the Nen/Komnzo case, and as we have discussed elsewhere (Passmore et al., 2021), kinship terminology do not have to be coherent throughout the system for a system to be functional. The inconsistency of merging patterns appears to be a feature of Papuan kinship.

A picture containing text, screenshot, diagram, square

Description automatically generated

**Figure 3:** Bar plots showing the proportion of languages that show kinship syncretisms. Numbers in boxes show the total number of non-Papuan and Papuan languages used in each calculation. In the bottom row of graphs, only select structures are shown but the numbers reflect total counts.

Asides from bifurcate merging systems, Papuan languages stand out in other ways. First, Papuan languages show a higher frequency of two particular sibling systems, relative to the global sample. One system contains three terms, one each for elder brother, younger brother, and a single sister term (27 Papuan languages or 28% of the sample). The other is the use of a single kinterm for all siblings (14; 14%). The three-term system is found across 15 clades but is most common in the Trans-New-Guinea family (7 languages), with a secondary cluster in North Halmaheran languages (4 languages). The single-term sibling system is found in five clades. It is mostly common in the TNG clade (11 languages), with more than half of those languages coming from the Timor-Alor-Pantar clade (6).

Another common feature of Papuan kinship terminology is self-reciprocity in grandkin terms, as described in Watam, Nen and Komnzo, and Nagovisi (Table 1). In taking a broader view, we see that around 34% of our Papuan sample contains a single self-reciprocal term for ‘grandfather’ and ‘grandson’, compared to only 15% of the non-Papuan sample. There is a similar pattern between ‘grandmother’ and ‘granddaughter’ (36% in Papuan languages, 15% in non-Papuan). Although this feature is proportionally more common in Papuan languages than non-Papuan, it is not the most common organisation of grandkin – with most Papuan languages (57%) having separate words for grandfather and grandson. Grandkin self-reciprocity occurs in 13 Papuan clades.

Finally, Papuan kinship contains a disproportionate occurrence of matrilineal skewing (MB = MBS), another feature that is commonly affiliated with patterns of exchange and with bifurcate merging patterns. 16% of Papuan languages contain this feature (n = 12), compared to 5% of non-Papuan languages. Nine of the 12 languages are from the Nuclear Trans-New Guinea clade, largely in Central West Papuan, with one Yam language,

**3b. Is Papuan kinship disproportionately diverse?**

While above we have looked at particular features of kinship terminology that make Papuan languages stand out, another interesting question is where kinship terminology is generally more diverse amongst Papuan languages than other language groupings. We look at this in two ways. First, we look at how many distinct kinship structures Papuan languages have relative to global variation. Secondly, we look at how the range of diversity covered by Papuan languages compares to other language groupings. Because the possible space is so vast, we restrict our analyses to three subsets of kin types: (a) parents and nuncles, (b) siblings alone, and (c) siblings plus cousins. Unlike the above section, here we analyse male and female relatives simultaneously. For each of these subsets, we consider genealogical categories that are differentiated by relative age (e.g., elder brother vs. younger brother) or by the relative age of connecting kin (father’s elder brother vs. father’s younger brother) and gender of speaker (a man talking to their sister vs. a woman talking to their sister), creating a larger possibility space. To analyse kinship terminology structure, the list of kinterms applied to kin types are converted into a string of 1’s and 0’s by comparing all kin types within our subset of interest to each other and asking if they have the same kinterm (1) or not (0). We call the binary string a “structural vector”. The structural vector is an abstract representation of a language's kinship terminology and ignores the language-specific formal instantiations contained in the kinterms. For example, English *grandfather* and Russian *dedushka*, and English *grandmother* and Russian *babushka*, have unrelated forms but cover identical sets of kin types, i.e., they have the same patterns of syncretism, so they would receive the same structural vector. Using the structural vector, we can ask: how many different structures are amongst our Papuan sample? How does that compare to the totality of global variation? Finally, how does Papuan variation compare to the variation from a random worldwide sample of languages (the same size as our sample of Papuan languages)? We summarise the answers to these questions in Table 2.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | Parents &  Nuncles | Siblings | Siblings & Cousins |
| Papuan | N | 94 | 102 | 61 |
| Distinct Structures | 40 | 33 | 53 |
| Indo-European | N | 88 | 99 | 81 |
| Distinct Structures | 12 | 7 | 22 |
| Austronesian | N | 263 | 329 | 191 |
| Distinct Structures | 50 | 52 | 99 |
| Pama-Nyungan | N | 68 | 98 | 26 |
| Distinct Structures | 36 | 18 | 20 |
| Kinbank | N | 870 | 1,011 | 612 |
| Distinct Structures | 154 | 118 | 327 |
| Distinct structures in Rand. Sample | 32.37  (28.93 - 35.82) | 29.26  (25.94 - 32.59) | 45.04  (41.63 -48.45) |

**Table 2.** For each subset of kin types (columns), we show the number of languages and unique structures for our sample of Papuan languages, for Indo-European, Austronesian, and Pama-Nyungan languages held in Kinbank, and within the entire sample of Kinbank. The final row shows the average number of unique structures in a random worldwide sample of languages, drawn from 1000 random samples, with one standard deviation above and below the mean.

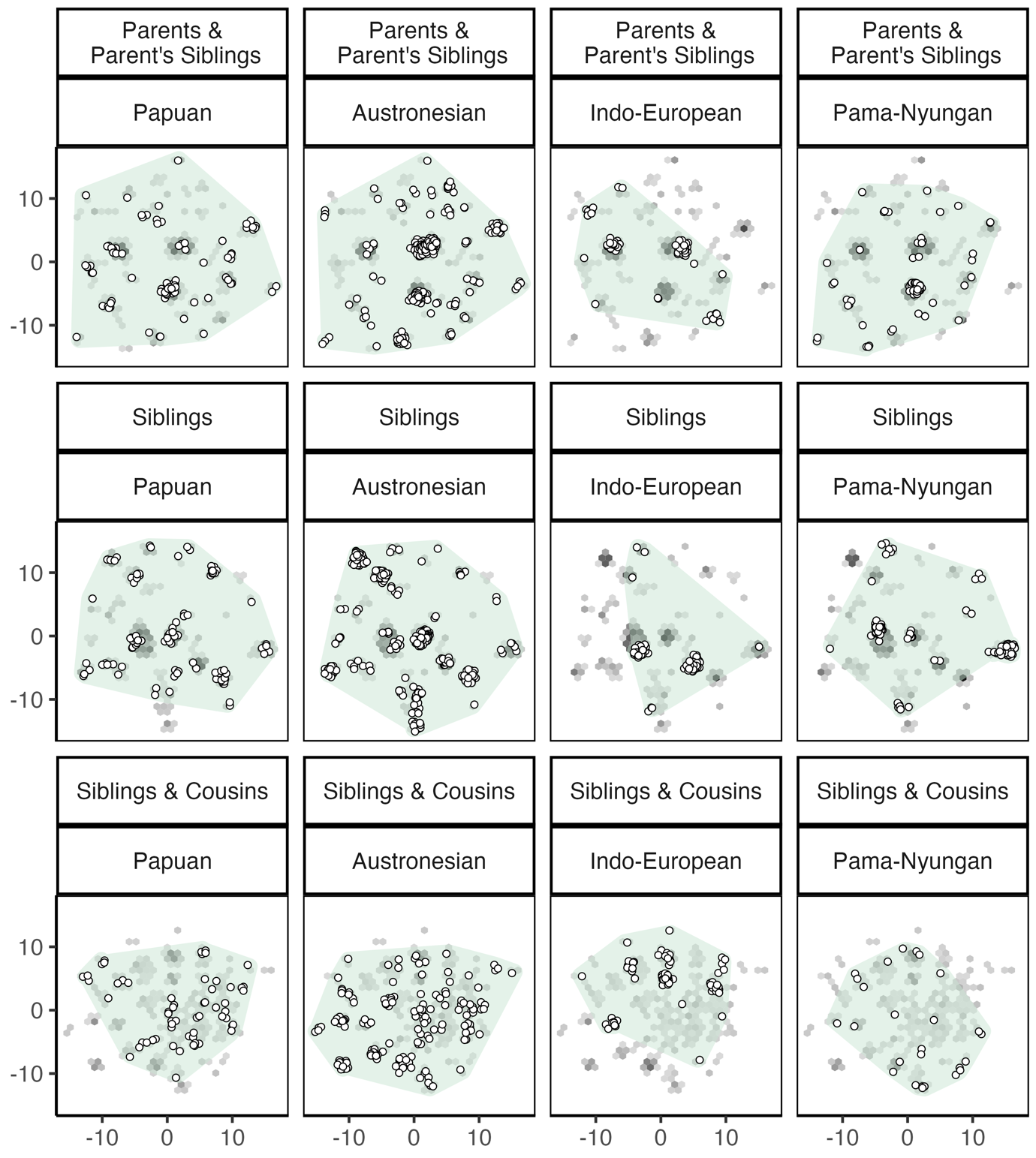
The most striking outcome of Table 2 is that in all our subsets, kinship terminology diversity in Papuan languages is more than one standard deviation higher than we would expect in a random sample. In fact, Papuan languages contain between 16% and 25% more diversity than a random selection of languages of the same size.

To compare the level of diversity in Papuan languages to the global pattern, and to other language groupings, we use the ratio of structures to languages. The ratio of kinship terminology structures to language tells us, on average, how many languages are their per distinct structure where a value of 1 indicates one language per structure and the maximum level of diversity, and every increase in the number of languages per structure indicates a decrease in diversity. In the entire Kinbank sample Parents/Nuncles have a ratio of 5.6 languages per structure, while the Papuan sample has a ratio of 2.3, meaning Papuan languages show more than twice the diversity of the general sample. This effect is greater when compared to Indo-European, with a ratio of 7.3. Austronesian languages have a ratio of 5.3, and Pama-Nyungan shows the lowest ratio, 1.9, and therefore the highest level of diversity in this comparison.

Siblings show a ratio of 8.5 languages per structure in the global sample, but only 3.0 in Papuan languages. Indo-European again shows limited diversity with a ratio of 14.1, Austronesian has a ratio of 6.3, and Pama-Nyungan has a ratio of 5.4. Finally, Siblings and Cousins have a ratio of 1.9 languages per structure in Kinbank, but Papuan languages have a ratio of 1.1, in other words almost each language has a unique structure. Again, Indo-European shows the most limited diversity, with a ratio of 3.7, Austronesian has a ratio of 1.9, and Pama-Nyungan contains a level of diversity approaching Papuan, with a ratio of 1.3. In general, Papuan kinship shows a higher level of diversity than global, Indo-European, and Austronesian samples, and levels of diversity slightly higher than Pama-Nyungan languages.

An implicit assumption within the analyses of the ratios is that each formal difference, even if small, counts as an equally important difference in semantic structure, but we might not always think this is the case. As an example, consider the four possible kin types referring to male siblings (elder, younger, and whether a male or female is speaking), and the systems used in English, Nagovisi, and Watam. In English, there is one term that glosses over all four possible categories*.* In Nagovisi there are two terms, one for elder brother and one for younger brother. Finally, In Watam there are three terms elder same-sex sibling, younger same-sex sibling, and opposite-sex sibling. These are three distinct sibling systems, but we also might think that a system with a term for ‘brother’ is more similar to a system with two words for ‘elder brother’ and ‘younger brother’ (without regard to the sex of ego) than it is to a system with terms based on relative sex. This example extends the idea of possibility space to include distances, where we might assume small differences are indicative of similarity, and that smaller changes are more likely to occur.

Papuan kinship contains more distinct systems than we would expect, given the number of languages, but do Papuan kinship systems cover more of the kinship design spacethan we might expect? A metric that approximates this idea in ecology is called “functional richness”, which quantifies the area occupied by a particular species defined by a set of measured features (Villéger et al., 2008), and has been used in linguistics to quantify the amount of grammatical diversity at risk from languages no longer being spoken (Skirgård et al., 2023). We calculate the functional richness of Papuan languages by first calculating the Manhattan distance between all pairs of languages. Manhattan distance is the sum of the difference between the two vectors and thus gives us a measure of the structural distance between any pair of languages. If a language has a 1 in the position where another language has 0, that increases the distance between them by 1. We then summarise the distance matrix into two dimensions using an algorithm called Uniform Manifold Approximation and Projection (UMAP; (McInnes et al., 2018)). UMAP projects the distances into a two-dimensional space, using manifold learning techniques and logic drawn from topological data analysis (Figure 2). We use the projection of the distance matrix to calculate the area covered by each language group, in the form of convex hulls. The area of the convex hulls (green) quantifies how much of the total space is covered by a subset of languages, compared to the total diversity. We standardize the scores so that the volume of all languages would give a value of 1. Scores less than one can be considered a proportion of the total area covered.

****

**Figure 2:** A UMAP projection of global kinship system diversity for Parent and Parent’s Siblings, Siblings, and Siblings and Cousins. Underlying the plot are a tessellation of grey hexagons. The hexagons show the distribution of languages across the whole database. The shade of the hexagon indicates the density of languages (a darker shade indicates more languages). White circles show the distribution of languages within the family named. The green convex hull shows the totality of the space covered by language groupings.

Figure 2 provides us with a visual description of global kinship terminology diversity, and how different language groupings (white circles) are distributed across the observed possibility space. Although there are clusters of Papuan systems, Papuan languages are scattered widely across the observed space. In Parents and Parent’s siblings, Papuan languages cover approximately 0.74 of the space, which is comparable to the coverage seen in Pama-Nyungan (0.75). Comparatively, Austronesian languages cover almost the entire diversity space (area = 0.91), but Indo-European languages only cover around half (0.53). Papuan sibling diversity covers a large portion of observed diversity (0.96), as does Austronesian sibling diversity (0.92). Pama-Nyungan sibling terminology only covers around half the observed diversity (0.52). Indo-European again shows the lowest level of coverage (0.30). Expanding the sibling category to include cousins creates a much larger theoretical and observed possibility space, and consequently reduces the diversity seen in Papuan languages (0.55). The same level of diversity is seen in Pama-Nyungan languages (0.55). Surprisingly, Indo-European languages cover a large proportion of the sibling and cousin space (0.69). However, coverage across this space is sparse, meaning the high functional score could be attributed to a few languages containing unusual patterns, rather than a language family maintaining high diversity. If we consider the range of Indo-European languages, in Europe but also through Asia, the extent of cousin diversity makes sense. In Western Europe (English, French) it is common to see a lineal pattern, separating lineal relatives (mother, father) from collateral relatives (parent’s siblings, cousins), but in Indo-European languages east of Europe, we observe more diversity in cousin organisation as a result of a higher prevalence of cross-cousin marriage (Schulz et al., 2019). Finally, Austronesian sibling and cousin terminology, again, covers a large portion of the observed diversity (0.90).

Testing the functional richness of kinship has shown that Papuan kinship is more diverse than Indo-European but has similar levels of diversity to neighbouring language groupings, in Austronesian and Pama-Nyungan. Since Austronesian and Pama-Nyungan are clearly defined language families and the Papuan category contains many unrelated language families, the diversity in kinship systems within Papuan languages is lower than we might expect if we assumed kinship terminology variation was unbounded.

Anthropologists and cognitive scientists point to the influence of social and cognitive constraints on kinship terminology diversity (Kemp & Regier, 2012; Murdock, 1949; Passmore & Jordan, 2020). These theories can explain why it is unlikely that we will ever observe all theoretical possibilities of kinship organisation and give reason to believe that kinship variation is not unlimited. However, it leaves open the question of why Papuan languages show an unusually high ratio of languages to distinct structures (which we will call structural diversity), but moderate levels of functional diversity. These are seemingly contrasting findings ~~conclusions~~. We propose three contributing factors. In part, the analyses are limited by the available samples. We analyse a small subset of Papuan languages (~100 languages from a total of ~850, ~12%), whereas we have a large sample of Austronesian languages (~~330 languages from a total of ~1,200, ~28%). Sample sizes between Papuan, Pama-Nyungan, and Indo-European languages were comparable. It may be that functional diversity is higher in the population of Papuan languages than it is in our sample. Making predictions on how diversity might increase with more data relies on two factors for which we have limited knowledge. First, increasing the sample size will increase the likelihood that we sample related languages, and since kinship systems show strong patterns of inheritance, will also decrease the likelihood of showing more functional diversity. Secondly, as the proportion of diversity observed increases, the likelihood we observe previously unseen diversity decreases. How it decreases with respect to cognitive and social pressures is currently unknown.

A second consideration for the contrast between high structural diversity and moderate functional diversity is the impact of neutral evolution. A neutral change, regarding kinship terminology, is the introduction of a new category that has little to no influence on communication or cognitive load. Since the neutral categories have negligible impact to the communication system, they can accumulate over time, creating more unique structures, without capturing more functional space. It is established that Papuan languages are much older than their Austronesian or Indo-European counterparts, which allows time for neutral changes to accumulate. Comparatively, Austronesian languages, a young language family, have similar levels of functional diversity to Papuan languages, but much lower levels of structural diversity. A parallel force to neutral evolution is adaptive evolution. In the context of Kinship terminology, adaptive evolution reflects the change of kinship terminology to a change in social organisation (the gain or loss of cousin marriage for example). Adaptive evolution is what is likely to cause a language grouping to cover more of the possibility space, and so we would posit that kinship behavioural diversity is similar between Papuan and Austronesian groups, explaining the similar levels of functional diversity, but due to the time-depth of Papuan languages, there are more distinct structures amongst Papuan languages than Austronesian.

**4. Conclusions**

**References**

Kemp, C., & Regier, T. (2012). Kinship Categories Across Languages Reflect General Communicative Principles. *Science*, *336*(6084), 1049–1054. https://doi.org/10.1126/science.1218811

McInnes, L., Healy, J., Saul, N., & Grossberger, L. (2018). UMAP: uniform manifold approximation and projection. *The Journal of Open Source Software*, *3*(29), 861.

Murdock, G. P. (1949). *Social structure* (Vol. xvii). Macmillan.

Nerlove, S., & Romney, A. K. (1967). Sibling terminology and cross-sex behavior. *American Anthropologist*, *69*(2), 176–187.

Passmore, S., Barth, W., Quinn, K., Greenhill, S. J., Evans, N., & Jordan, F. M. (2021). Kin Against Kin: Internal Co-selection and the Coherence of Kinship Typologies. *Biological Theory*. https://doi.org/10.1007/s13752-021-00379-6

Passmore, S., & Jordan, F. M. (2020). No universals in the cultural evolution of kinship terminology. *Evolutionary Human Sciences*, *2*. https://doi.org/10.1017/ehs.2020.41

Schulz, J. F., Bahrami-Rad, D., Beauchamp, J. P., & Henrich, J. (2019). The Church, intensive kinship, and global psychological variation. *Science*, *366*(6466), eaau5141. https://doi.org/10.1126/science.aau5141

Skirgård, H., Haynie, H. J., Blasi, D. E., Hammarström, H., Collins, J., Latarche, J. J., Lesage, J., Weber, T., Witzlack-Makarevich, A., Passmore, S., Chira, A., Maurits, L., Dinnage, R., Dunn, M., Reesink, G., Singer, R., Bowern, C., Epps, P., Hill, J., … Gray, R. D. (2023). Grambank reveals the importance of genealogical constraints on linguistic diversity and highlights the impact of language loss. *Science Advances*, *9*(16), eadg6175. https://doi.org/10.1126/sciadv.adg6175

Villéger, S., Mason, N. W. H., & Mouillot, D. (2008). New Multidimensional Functional Diversity Indices for a Multifaceted Framework in Functional Ecology. *Ecology*, *89*(8), 2290–2301. https://doi.org/10.1890/07-1206.1

1. In general we use the absolute binary criterion below, because of its simplicity, but it obviously has drawbacks since it would treat Spanish *hermano* ‘brother’ as no closer to *hermana* ‘sister’ than it is to *primo* ‘male cousin’, whereas a gradient scale such as a Levinshtein distance might assign a value of 0.83 syncretism to *hermano/hermana* but of only 0.17 to *hermano/primo*. Levinshtein distances are calculated by seeing how many symbol substitutions, as a proportion of the average length between the two words, are needed to convert one to the other. In the *hermano/hermana* example the *h* is ignored since it is not pronounced in modern Spanish. [↑](#footnote-ref-1)
2. Cross-kin are those where the chain of genealogical connection crosses an opposite-sex sibling link. Thus a cross-uncle is MB and a cross-aunt FZ, and a cross-cousin is MBS, FZS etc. The term ‘cross’ is opposed to ‘parallel’, where the chain of connections stays in the same sex, e.g. a parallel uncle is FB and FBS is an example of a parallel cousin. [↑](#footnote-ref-2)
3. We employ the standard symbols for representing etic kinship relationships in a compact manner, namely: B=brother, e=elder, C=child, D=daughter, F=father, H=husband, M=mother, S=son, W=wife, y=younger, Z=sister, ♂︎=male ego/anchor, ♀︎=female ego/anchor. These can then be concatenated to yield expressions like FBS ‘father’s brother’s son’, ♂︎BW ‘man’s brother’s wife’, ♀eZ ‘woman’s elder sister’. [↑](#footnote-ref-3)
4. We follow established practice in using *nuncle* to mean ‘uncle or aunt’ and *nibling* to mean ‘nephew or niece’. [↑](#footnote-ref-4)
5. Among the many types of syncretism that are found, ‘skewing rules’ refer to those which equate members across generations, typically linked by membership of some lineage. (Though note that there is another type of cross-generational syncretism resulting from self-reciprocal terms like ‘grandkin’, such as Nen *kake* ‘grandparent, grandchild’). The canonical skewing rules discussed in the literature (Omaha and Crow) collapse a nuncle and a cousin term, i.e. +1 and ego generations, then often continuing down through the relevant clan lineage. But taken more broadly, there are many other types. For example in Yêli-Dnye the term [X] recurs down one’s mother’s brother’s matriline every two generations, and in Nen the terms *nako/yakali* skew down through a lineage, taking its point of departure from a type of affine (WZH) on down to the cousins linked by this relation (i.e. MZC) – see §2.4. We lack the space to investigate these thoroughly here but a broader investigation of Papuan skewing rules seems likely to reveal many interesting types, reflecting the importance of maximising affinal ties in many Papuan societies. [↑](#footnote-ref-5)
6. Bamford argues that ‘Kamea do not rely on physiological reproduction as a means of tracking social relationships through time. Despite my repeated efforts to ground intergenerational relations in a procreative bond, Kamea were quite insistent on the fact that neither a mother nor a father shares substance in common with their offspring. Instead, the parent-child tie is imagined as an inherently disembodied one.’ We confess to some skepticism about this claim, which contradicts our own experience across a number of societies, which though they extend kinship relations outward, through polysemy, to categories of relation not grounded in biological procreative bonds, nonetheless take procreative bonds as their semantic core.

   ’ [↑](#footnote-ref-6)
7. In fact, this particular mix is by no means rare in the world’s languages. Passmore et al (2021) report that, in their global sample of 571 languages, 35 languages showed this mix – more, in fact, than exhibited consistently Hawaiian systems (15), and a bit over half the number exhibiting consistently Dravidian systems (62). [↑](#footnote-ref-7)
8. As well as his transcriptions, which omit tonal marking; cf. Hyman & Kobepa 2013. [↑](#footnote-ref-8)
9. Throughout this section, ‘moiety’ is to be taken to mean ‘matrimoiety’. [↑](#footnote-ref-9)
10. Fide Nash (1974:132). The equation of eZ with FM seems puzzling, since eZ is a moiety mate with ego while FM is cross-moiety. Similar conundrums attach to the [↑](#footnote-ref-10)
11. Though the degree of genealogical depth varies significantly, even across small distances. While I have recorded genealogies of nine generations from Nen speakers, in nearby Komnzo Ayres (1983: 266) describes the genealogies as "shallow", often going back only two generations between ego and the apical ancestor. Christian Döhler (p.c.) confirms this from the genealogies he collected, which rarely go back further than 3-4 generations until the first ancestor. Of course even the nine generations recorded in Nen is extremely shallow compared to the forty or fifty generations regularly recorded from Polynesian societies. [↑](#footnote-ref-11)
12. *zaθ* also means ‘daughter-in-law’, reflecting the expectation that a daughter’s daughter will typically ‘return’ in marriage to the clan of her grandmother, so that one’s daughter-in-law is frequently the granddaughter of one’s mother (Ayres 1983:226, Döhler 2018:29-30). [↑](#footnote-ref-12)
13. Cross-kin are those where the chain of genealogical connection crosses an opposite-sex sibling link. Thus a cross-uncle is MB and a cross-aunt FZ, and a cross-cousin is MBS, FZS etc. The term ‘cross’ is opposed to ‘parallel’, where the chain of connections stays in the same sex, e.g. a parallel uncle is FB and FBS is an example of a parallel cousin. [↑](#footnote-ref-13)
14. The Komnzo post-sister-exchange nuncle terms are probably contractions of *fäms ŋafe* "exchange father" and *fäms ŋame* "exchange mother" (C. Döhler, p.c.). Both terms are somewhat archaic and most people use *bäiŋaf* and *bäiŋam*, also found in neighbouring Wära. [↑](#footnote-ref-14)
15. Suki has a word *naku*, meaning either mWZH or mWZ (Charlotte van Tongeren p.c.) and also – as reported by Van Nieuwenhuijsen-Riedeman (1979) – a term *nakima* for mZSW, fHMB, fHMBS, fHMBSS.   
     [↑](#footnote-ref-15)
16. The K cognate *kaimat* has a slightly different meaning: ♀︎BW. [↑](#footnote-ref-16)
17. CD: I am not sure about this. I will check in November. My guess would be that they use the term for 'wife' /fzenz/. [↑](#footnote-ref-17)
18. Elder brother, elder sister, younger brother, younger sister, and, for each term, whether it is a man or woman speaking. [↑](#footnote-ref-18)
19. Relative age of parents’ siblings is also a common feature, but it is not important in Papuan languages. [↑](#footnote-ref-19)
20. The fifth possibility is F = MB ≠ FB which violates the rule of colaterality and has never been observed. [↑](#footnote-ref-20)