

Chapter 8

The gender system of Coastal Marind

Bruno Olsson

Australian National University

The gender system of Coastal Marind (a Papuan language of the Anim family of South New Guinea; Usher & Suter 2015) is treated in relative detail in Drabbe's (1955) masterful grammar. The division of nouns into four genders (basically masculine, feminine and two inanimate genders) is familiar from various languages around the globe, but the morphology of exponence (gender agreement marked to a large extent by stem-internal changes on targets) is somewhat more exotic and is occasionally cited in the literature. In this paper I provide an overview of the system, combined with discussion of two issues: the origins of stem-internal gender agreement, and the wide-ranging syncretism between animate plurals and the 4th gender (the 2nd inanimate gender). I show that this 'syncretism' makes the status of the 4th gender ambiguous, since the members of this gender also could be analysed as an unusually large class of pluralia tantum. While I argue that the synchronic 4-gender analysis must be maintained for Coastal Marind, I speculate that an erstwhile grouping of pluralia tantum provided the diachronic source of the 4th gender.

Keywords: Gender, number, morphology, diachrony, Papuan languages.

1 Introduction

The idea that gender systems can become more complex (add a gender or two) through the 'reinterpretation' of some non-gender feature as signalling a gender value has a long history in linguistics (e.g. Brugmann 1891 on the origins of the Indo-European feminine gender). In this paper I show that the fourth gender of Coastal Marind could be more parsimoniously described as pluralia tantum in a 3-gender system; however, I will argue that semantic considerations ultimately force us to retain the traditional four-gender description.



Based on its ambiguous status in Coastal Marind, I will speculate that the fourth gender in the languages of the Anim family of South New Guinea could have originated as a grouping of pluralia tantum nouns, and that subsequent changes in the agreement system and attraction of additional nouns to the emerging fourth gender could have led to a present situation where the pluralia tantum analysis is no longer possible, resulting in a 4-gender system.

I also add further support to Usher & Suter's (2015) proposal that one of the main manifestations of gender agreement in the language – stem internal vowel alternations in agreement targets – arose from a process of umlaut triggered by postposed articles, by showing that the synchronic distribution of stem-final vowels in nouns is consistent with gender umlaut affecting a much larger part of the lexicon than just present-day gender-agreeing lexemes. The discussion is based on data from the best known Anim language, Coastal Marind (for a modern reference grammar, see Olsson 2017).

The article is structured as follows. §1.1 is a brief demonstration of the four genders of Coastal Marind. The language is placed in its areal and genealogical context in §1.2, while §1.3 provides information about some relevant structural features of Coastal Marind. §2 describes the interesting correlation between stem-final vowels and gender membership in nouns, showing that it is of limited productivity synchronically, but likely derives from an earlier system of post-nominal gender articles. §3 describes gender agreement across the clause, with emphasis on the systematic correspondence between exponents of Gender IV and the plural of Gender I/II. §4 shows that this correspondence continues in the participant indexing on the verb. This suggests an alternative analysis according to which Gender IV is an unusually large group of pluralia tantum rather than a gender of its own. In §5 I will show that the assignment of nouns to Gender III and IV is largely arbitrary, but that the occurrence in Gender IV of many nouns that are typical pluralia tantum nouns across languages is suggestive of being a remnant of such a grouping. I also show that a similar pattern occurs in Mian, a language that probably is a distant relative of Coastal Marind since the Anim and Ok families (to which Mian belongs) are likely members of the enormous Trans-New Guinean super-family. I conclude that the 4-gender analysis should be maintained for the present state of Coastal Marind, but that the pluralia tantum nouns possibly provided the source for the fourth gender.

1.1 The Coastal Marind 4-gender system

The existence of a 4-gender system in Coastal Marind is evident if one compares the form of the demonstrative *Vpe* (where *V* stands for a vowel) or the adject-

tive *samlayVn* ‘mid-size, neither big nor small’ combined with different nouns in examples (1)–(3). As indicated by the hyphens, attributively used adjectives are compounded with their head nouns. The nouns themselves are invariant.

- (1) a. *samlayen-patul e-pe*
mid.size:I-boy(I) I-that
- b. *samlayun-kyasom u-pe*
mid.size:II-girl(II) II-that
‘that mid-size boy/girl’
- (2) a. *samlayin-patul i-pe*
mid.size:I/II.PL-boys(I) I/II.PL-that
- b. *samlayin-kyasom i-pe*
mid.size:I/II.PL-girls(II) I/II.PL-that
‘those mid-size boys/girls’
- (3) a. *samlayan-da e-pe*
mid.size:III-sago(III) III-that
‘that mid-size sago palm/those mid-size sago palms’
- b. *samlayin-bomi i-pe*
mid.size:IV-termite.mound(IV) IV-that
‘that mid-size termite mound/those mid-size termite mounds’

All nouns denoting male humans behave like *patul* ‘boy’ (in 1a) in combining with a demonstrative with the initial vowel *e-* in the singular; nouns denoting female humans (and all animals) pattern like *kyasom* ‘girl’ (1b) in combining with an *u-* initial demonstrative. As the examples in (2) show, these nouns exhibit a contrast in number. The demonstrative has to be *ipe* in the plural, and the adjective, which is compounded with its head noun, has the exponent vowel *i* in the final syllable of the stem.

The nouns in (3) are inanimate, and trigger different vowels on the demonstrative: *da* ‘sago palm’ triggers *e-*, *bomi* ‘termite mound’ triggers *i-*. Note that the resulting forms are homophonous with demonstratives in the preceding examples: *epe* in (3a) with the demonstrative used for *patul* in (1a), and *ipe* in (3b) with the plural forms in (2). For (3a), the distinct form *samlayan* of the adjective proves that this is indeed a separate gender, although the agreement of the demonstrative happens to be homophonous with that seen in (1a). But the case

in (3b) is more difficult, since the agreement on both the demonstrative and the adjective turns out to be homophonous with the plural forms. I will return to this pervasive syncretism further below.

The four agreement classes – from now on referred to as Gender I, II, III and IV – are summarized in Table 1, as evidenced by the exponence pattern of *samlayVn*.

Table 1: Exponents of agreement on *samlayVn* ‘mid-size’

	SG	PL
I	e	$\left. \begin{array}{l} \diagup \\ \diagdown \end{array} \right\} i$
II	u	
III	a	
IV	i	

These data represent one of the most well-known gender systems in New Guinea. The Coastal Marind system of four grammatical genders has featured in prominent publications such as Corbett (1991: 116) and Aikhenvald (2000: 60) after having been brought to the fore in Foley’s influential compendium on Papuan languages (Foley 1986: 82–83). This attention is due to the description of the gender system provided in Petrus Drabbe’s extensive grammar of the language (Drabbe 1955). Few researchers seem to have had the courage to dive deeper into Father Drabbe’s sometimes quite demanding *Spraakkunst*, so one purpose of this article will be to give a more representative picture of the gender system and its manifestations, and, in particular, the syncretism between animate plurals and Gender IV. The data come from my own fieldwork on the Western variety of Coastal Marind, a dialect that is mutually intelligible with the Eastern variety described by Drabbe.

1.2 Coastal Marind in context

The varieties collectively known as Coastal Marind are spoken in ca. 40 villages along the coast of the Arafura sea and in the adjoining swampy lowlands. I estimate the total number of speakers to be around 14.000 based on government and SIL figures. The Coastal Marind land forms part of the linguistically diverse Trans-Fly area (Evans 2012; Evans et al. 2018) straddling the border of present-day Indonesia (where Coastal Marind is spoken) and the independent country of Papua New Guinea.

The dialect situation is complex, and it is probable that ongoing research will show that some of the varieties described in the literature as dialects are in fact distinct languages. Dialectal variation in gender would likely be an interesting area to explore, as there are differences (mainly in assignment) even between villages speaking virtually identical varieties of Coastal Marind. On the whole, however, the basics of gender and agreement are the same in all known varieties, so the data presented here (from the village of Wambi) are representative of all coastal varieties, and probably of the (less well-known) inland varieties as well.

On a higher level, gender has recently emerged as a crucial factor in the genealogical classification of Coastal Marind. Usher & Suter (2015) show that gender ablaut in nouns such as *anem* ‘man’, *anum* ‘woman’ and *anim* ‘people’ recur throughout a number of languages of the Trans-Fly region. This observation, in addition to a large set of lexical cognates showing regular sound correspondences, leads Usher & Suter to propose a hitherto unrecognized language family – the Anim family, named after the recurring word for ‘people’ – of which Coastal Marind so far is the only language for which substantial descriptive work is available. Obviously, more work on the other Anim languages – several of which are rapidly losing speakers – could provide crucial insights into the development of the Anim gender system.

1.3 Typological background

Some of the structural features of Coastal Marind are relevant to the description of its gender system. Coastal Marind displays the relatively rare combination of verb-final constituent order and massively prefixing verb inflection. Based on co-occurrence, a prefixal template with ca. 18 slots can be set up, marking notions such as tense, various aspectual distinctions, applicatives, reciprocal, various adverbial meanings (‘again’, ‘first’, ‘far away’, ‘in contact with surface’) and indexation of (roughly) actor, recipient and affected possessor; undergoer indexation is in turn marked on the verb stem by complicated alternations including pre-, suf-, in-, and circumfixal morphology.

Some of the prefixes occupying the first (i.e. leftmost) positions agree in gender with an argument, although they primarily mark grammatical distinctions other than gender (e.g. tense-aspect). The prefixes devoted to argument indexing, on the other hand, reflect person and number but are insensitive to gender (with some exceptions to be discussed later). The verb stem itself is an important site for the manifestation of gender, so the intricate stem changes will be crucial to the arguments made here.

A relatively straightforward example of how verbs are segmented is given in (4). This verb has two prefixes, of which the first (leftmost) prefix agrees in gender with the subject (plural of Gender I/II). The stem is separated from the prefixal complex by a phonological boundary (indicated in glossing by means of a trailing hyphen followed by a blank). The formative *n-* on the stem marks it as the 1st person undergoer form, which clearly is a mismatch since there is no 1st person participant involved in the event. This idiosyncrasy is part of the reciprocal construction, and such value mismatches are not uncommon in Coastal Marind (cf. §4).

- (4) *ip-enam- n-asak-e*
ABSC:I/II.PL-RECP- 1.U-fight-IPFV
'They are fighting.'

Nominal morphology is sparse: there is no case marking and most nouns do not show overt gender marking. The exception is a handful of nouns (mostly kinship terms) that show alternations in the stem-final vowel according to gender (see below). This marking pattern also occurs on a subset of adjectives which agree with a noun in attributive and predicative use. The majority of adjectives are invariant and fail to show agreement. Instead, the main loci of gender agreement outside verbs are demonstratives and pronominal-like words (emphatic pronouns, question words). In the next section I turn to the reflexes of gender in nouns and what they can tell us about the diachronic development of gender marking in this part of the lexicon.

2 The manifestation of gender in nouns

2.1 Overt gender

A comparison of gender agreement across different word classes confirms that the picture emerging from examples (1)–(3) above is correct. All words that show morphological alternations according to gender follow these four agreement classes, although exponents vary across the targets showing agreement, and although many targets do not distinguish all four classes. Before dealing with agreement proper, we will consider nouns displaying OVERT GENDER. Whereas such alternations are not productive in contemporary Coastal Marind, a closer look reveals that traces of a more wide-ranging system of stem-final vowel alternations can be observed. The origins of this system of overt marking can be reconstructed following Usher & Suter (2015), as will be seen later.

Table 2: Overt gender on nouns

	I SG	II SG	I/II PL	III	IV
a.	<i>anem</i> 'man'	<i>anum</i> 'woman'	<i>anim</i> 'people'	<i>anem</i>	<i>anim</i>
b.	<i>namek</i> 'cousin (m)'	<i>namuk</i> 'cousin (f)'	<i>namik</i> 'cousins'		
c.		<i>namakud</i> 'animal'	<i>namakid</i> 'animals'	<i>namakad</i> 'thing(s)'	<i>namakid</i> 'thing(s)'
d.	<i>amnanggib</i> 'married man'		<i>amnangga</i> 'married men'		
e.	<i>wananggib</i> 'boy'	<i>wananggub</i> 'girl'	<i>wanangga</i> 'children'		
f.	<i>nahyam</i> 'my husband'	<i>nahyum</i> 'my wife'			
g.	<i>eyal</i> 'somebody (m)'	<i>eyul</i> 'somebody (f)'			
h.	<i>nanih</i> 'face (m)'	<i>nanuh</i> 'face (f)'	<i>nanih</i> 'faces'		

Some nouns with overt gender marking are listed in Table 2. Gender membership is reflected by the vowel in the final syllable of the stem (referred to as the 'stem-final vowel'), and the meaning of the noun is largely predictable from the gender. Thus, the skeletal stem *anVm* (a) can be thought of as having the general meaning 'person', which is narrowed down to 'man' when assigned to Gender I (*anem*), 'woman' in Gender II (*anum*), etc.; the stem *nahyVm* 'my spouse' (f) (*nahy* is a 1st person possessive prefix) giving 'husband' (*nahyam*, Gender I) and 'wife' (*nahyum* Gender II) once gender is assigned and vowels plugged into the stem.¹

Assuming that the sets of gender forms derived from the skeletal stems are best treated as members of unitary lexemes, we can say that these lexemes are a proper subset of the nouns having REFERENTIAL GENDER (Dahl 2000), i.e. nouns that lack intrinsic gender and receive their gender value from the referent at

¹Note that 'overt gender' only applies to nouns for which there is at least one other noun differing only in a stem-internal vowel, with a corresponding change in meaning. For example, the Gender IV noun *bomi* 'termite mound' does not have overt gender despite the presence of stem-final *i* (which is the general exponent of Gender IV agreement), since there are no corresponding nouns **bome*, **bomu* etc. to be found in the other genders.

hand. Most such nouns do not show overt gender, e.g. *yunayon* ‘infant’ (which takes agreement in Gender I or II depending on the sex of the referent).

The disassembly of Coastal Marind nouns into skeletal stems with inserted gender markers could appear to be a slightly misleading way of approaching the gender system of the language, since the phenomenon is fairly marginal. Only a dozen lexical items or so display the vowel alternation,² and many of the expected forms are irregular (e.g. plural of *wananggVb* is *wanangga* ‘children’, there is no plural **wananggib*) or simply non-existent (e.g. there is no plural of *eyVl* ‘somebody’). The vowel alternation seems to be complete only for the stems *anVm* and *namakVd*: in addition to the person-denoting triplet man/woman/people, the former provides the forms *anem* and *anim* for inanimate denotanda in Gender III and IV respectively, for example in some compounds denoting fruits (*ambun-anem*, a *Syzygium* species in Gender III), while *namakVd* apparently can be used for non-rational entities (animals, things) of all genders except the masculine I.³

Looking at more nouns from Gender I and II, it seems clear that the pattern of alternating vowels showing gender membership is exception rather than rule. Nouns in Gender I denoting male humans also include *patul* ‘boy’, *ad* ‘father’, *manday* ‘wife’s elder brother, younger sister’s husband’ and so on; these nouns do not participate in any alternation with corresponding plural or female-denoting nouns. Person-denoting nouns in Gender II that likewise show no trace of overt gender are *kyasom* ‘girl’, *nikna* ‘son’s wife’, *ne* ‘mother’s brother’s wife’ etc.

Although overt gender is found only in a very small portion of the nominal lexicon, it should be noted that some of these nouns are high-frequency items, such as the words corresponding to the stem *anVm*, whose combined score makes them more frequent than any other noun in my corpus. Outside the noun inventory, stem-final vowel alternation plays an important role in common agreement targets such as the emphatic pronoun *anVp* (‘-self’), adjectives such as *papVs* ‘small’ and the postposition *IVk* ‘from’. This means that overt gender on nouns, and stem-final vowel alternation in general, is a common feature of Coastal Marind discourse, and obviously not as marginal as it would seem from a dictionary count alone.

²There are a handful of other nouns with overt gender in addition to the ones shown in the table. All of these denote humans of different age-ranks or societal roles that are more or less obsolete today, so the corresponding terms are falling out of use.

³In fact it seems that the stem *namakVd* ‘animal/thing’ can be used in Gender I: speakers reported that *namaked* can be used to refer to a male, although apparently with pejorative overtones, although I have never observed this in spontaneous speech.

A central claim of the comparative work in Usher & Suter (2015) is that the vowel alternations according to gender occur in languages throughout the Anim family, and that its origins can be reconstructed. Consider the forms *aneme(a)* ‘man’, *anumu* ‘woman’, *animi* ‘people’ from the related language Ipiko, another member of the Anim family. Usher & Suter argue that the stem-final vowel in *anVm* and other alternating stems is a residue of an earlier system of postnominal articles marking the gender of the noun, and they reconstruct expressions such as **anem=e* ‘the man’, **anum=u* ‘the woman’, **anim=i* ‘the people’ (2015: 114). In an earlier stage the noun was invariant and it was the presence of the gender article that triggered umlaut in the stem-final syllable (the shape of the invariant stem is beyond what can be reconstructed from the available data).

Usher & Suter’s hypothesis is plausible, especially as it refers to a well-known process leading to stem-internal vowel alternations (cf. Germanic umlaut giving English *mouse* and *mice* triggered by an earlier plural ending *-iz). It can be added that some alternations are likely the result of more recent derivations involving gender-marking morphology. For example, the word *wayuklu* ‘girl’ and its plural *wayuklik* ‘girls’ are probably related to the postposition ‘from’ which has the forms *luk* and *lik* in the feminine and plural respectively, and which seems to be the source of many deverbal nominals in Coastal Marind (see Geurtjens 1933: 335 for the etymology; cf. *dahahiplik* ‘drunkards’ from *dahahip* ‘become drunk (plural subject)’). However, the ultimate source of the vowel alternation in *IVk* ‘from’ is likely not distinct from the umlaut process giving rise to the forms of *anVm*, so the suggestion that some cases of synchronic vowel alternations are of more recent origin than the original umlaut is not intended as a counterexample to Usher & Suter, but as an indication that the alternating pattern propagated indirectly through the lexicon as a result of derivation.

2.2 Simulating the effects of umlaut in the lexicon

Given the observations of alternating nouns showing overt gender, and Usher & Suter’s suggestion that the alternation came about because of umlaut triggered by a postposed article, the following interesting question arises: are there traces of umlaut also in non-alternating noun stems?

If umlaut was a regular process, we would expect it to have appeared with many nouns, as long as they were used with postposed articles. In the ideal case, all nouns in Gender I would have ended up with the stem-final vowel *e*, those in Gender II stem-final *u*, Gender III *a*, and those in Gender IV *i*. This is clearly not the case, as shown by the counts of stem-final vowels in Table 3. The table displays the frequency with which each of the five vowels of Coastal Marind occurs

in the last syllable of nouns whose gender membership has been determined. I have excluded all nouns showing overt gender from the counts, since we already know that their stem-final vowels correlate with gender membership. This is the reason why Gender I has so few members: the remaining male-denoting nouns have overt gender (e.g. *anVm*). Gender II likewise contains only a handful of female-denoting nouns, but has a higher count since it includes all names of animals.

Table 3: Distribution of stem-final vowels in nouns according to gender

	I (<i>e</i>)	II (<i>u</i>)	III (<i>a</i>)	IV (<i>i</i>)	Tot.
/i/	5	29	25	44	103
/u/	0	27	39	19	85
/e/	1	15	31	13	60
/o/	2	22	34	14	72
/a/	4	55	108	29	196
Tot.	12	148	237	119	516

Consider now the possibility that stem-final vowels of nouns and gender membership correlate to some degree, despite there being no one-to-one match. We are particularly interested in the vowels *e*, *u*, *a* and *i*, which Usher & Suter (2015) identify as the vowels of the proto-Anim demonstrative.⁴ The vowels are given inside parentheses after their associated genders at the top of the table. We cannot test the correlation for Gender I, since there are too few nouns assigned to this category. The relevant cells for the remaining three genders have been shaded in Table 3. We now need to ascertain whether these scores could have been produced by a chance distribution of stem-final vowels, or whether they are non-random, thereby providing evidence that the umlaut pattern is found beyond the synchronically attested overt gender nouns.

To test this, I performed a simulation in which the nouns were reassigned randomly to the four genders (keeping the proportions intact), and then counted the frequency with which the vowels turned up in each gender. This procedure was then repeated a total of 200.000 times; the accumulated counts for the occurrence of the relevant vowels in Gender II, III and IV are presented in Figure 1, with the actual frequency of the vowel represented by the cross on the x-axis. The results

⁴In fact, Usher & Suter (2015: 119) tentatively reconstruct both **a* and **o* for the proto-Anim Gender III, but the exponent *o* is rare in Coastal Marind.

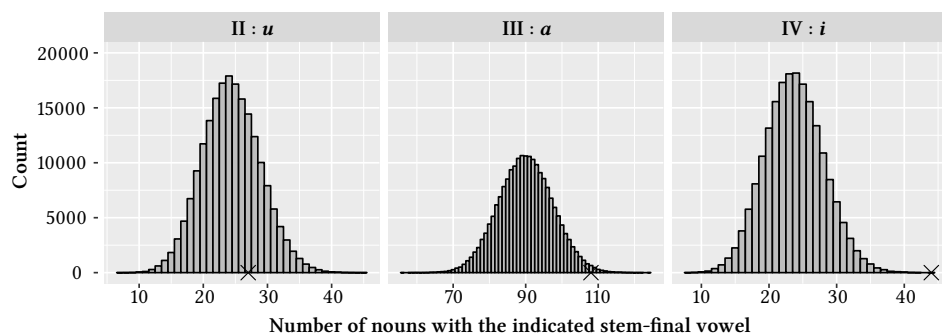


Figure 1: Actual and simulated distributions of stem-final vowels

show that two of the vowels are over-represented to a significant degree: *a* as the stem-final vowel in Gender III ($z=2.40$, adjusted $p<0.05$) and *i* as the stem-final vowel of Gender IV ($z=4.65$, adjusted $p<0.001$). These results support the hypothesis that gender umlaut affected a part of the lexicon that is larger than the set of nouns with overt gender, including many nouns of Gender III and IV.

No other positive skewings were close to statistical significance. This is somewhat surprising for Gender II, which would be expected to show a preference for *u* as the stem-final vowel (cf. the leftmost pane in Figure 1). I have no explanation for this, but it is worth noting that Coastal Marind seems to differ from other Anim languages in the uniform assignment of animals to Gender II: animals turn out to be divided between Gender I and II (the ‘masculine’ and ‘feminine’ genders) in Kuni (Edwards-Fumey 2007: 9), Ipiko (Usher & Suter 2015: 117, examples 16–17), and Bitur (Phillip Rogers, pers. comm.) which belong to three distinct sub-branches of Anim. A possible scenario would be that the reassignment of all animals to Gender II is an innovation present in Coastal Marind, which then would have obliterated any preponderance of *u* in Gender II as the new members entered.

3 Gender agreement

I will now consider how gender is manifested across agreeing pronominals, demonstratives and adjectives.⁵ The purposes will be to give an overview of the agreement system, which contains some typologically interesting features, and more specifically to show that the apparent syncretism noted above between

⁵There is one more type of agreement target, viz. the four postpositions *IVk* ‘from’, *nV* ‘without’, *tV* ‘with’ and *hV* ‘like’. They are interesting for a variety of reasons, but I omit them from discussion here.

Table 4: Pronominal and demonstrative targets

Gloss	I SG	II SG	I/II PL	III	IV
‘whats-his/her-name, whatchamacallit’	<i>age</i>	<i>agu</i>	<i>agi</i>	<i>ago</i>	<i>agi</i>
‘who/what’	<i>ta</i>	<i>tu</i>	<i>ti</i>	<i>ta</i>	<i>ti</i>
‘him-/her-/itself/themselves’	<i>anep</i>	<i>anup</i>	<i>anip</i>	<i>anep</i>	<i>anip</i>
‘this/these’	<i>ehe</i>	<i>uhe</i>	<i>ihe</i>	<i>ehe</i>	<i>ihe</i>

Gender IV and the plural of Gender I/II is observed throughout the system. It even turns up in some unexpected places, prompting the question of whether the system is not better analyzed as comprising three genders instead of four, a possibility that will be further explored in §4, §5 and §6.

3.1 Pronominals and demonstratives

The only word classes in which agreement is found on a majority of the members are demonstratives and pronominals. Agreement on the distal demonstrative *Vpe* was seen in (1)–(3) above; some more examples of agreeing targets within these categories are in Table 4. While the small set of personal pronouns in Coastal Marind (*nok* ‘I, we’ *oy* ‘2SG’, *yoy* ‘2PL’) show no gender distinction, gender agreement is pervasive across other pronominal-like elements such as question words (e.g. *tV* ‘who, what’ *Vn* ‘where, which’) and the polyfunctional word *agV*, which has among its uses that of a placeholder ‘whats-his/her-name’ (referring to a person) or ‘whatchamacallit’ (referring to a thing).⁶ Note that, in contrast to the various unpredictable exponents of Gender I and III, the exponents of Gender II (*u*) and Gender IV (*i*) are constant across all targets, with the latter showing homophony with the I/II plural in all four items.

3.2 Adjectives

Coastal Marind adjectives are similar to nouns in that both classes lack the luxuriant inflectional possibilities of verbs. The main morphosyntactic feature distin-

⁶Forcing speakers to choose a gender for words meaning ‘who, what?’ that refer to some unknown entity might seem counter-intuitive since the gender of the referent must be unknown in many cases (since there is no clear semantic basis for Gender III and IV); cf. European languages restricting gender agreement to attributive ‘which’ (e.g. Russian *kotoryj* ‘which (masc.)’ etc.) while pronominal ‘who’ lacks agreement (e.g. Russian *kto* ‘who’). Gender agreement on placeholders appears more common, especially in placeholders of phrasal and/or pronominal origin such as English *whatchamacallit* etc.

Table 5: Gender agreement on adjectives

Gloss	I SG	II SG	I/II PL	III	IV
'light (weight)'	<i>akek</i>	<i>akuk</i>	<i>akik</i>	<i>akak</i>	<i>akik</i>
'short'	<i>dahwages</i>	<i>dahwagus</i>	<i>dahwagis</i>	<i>dahwagis</i>	<i>dahwagis</i>
'thin'	<i>halahel</i>	<i>halahul</i>	<i>halahil</i>	<i>halahal</i>	<i>halahil</i>
'sharp'	–	–	–	<i>yayayay</i>	<i>yayayiy</i>
'dull'	–	–	–	<i>yandayal</i>	<i>yandayil</i>
'old, ancient'	<i>taname</i>	<i>tanamu</i>	<i>tanami</i>	<i>tanama</i>	<i>tanami</i>
'strong'	<i>tage</i>	<i>tagu</i>	<i>tagi</i>	<i>taga</i>	<i>tagi</i>
'ripe'	–	–	–	<i>eho</i>	<i>ihu</i>

guishing adjectives from nouns seems to be the lack of inherent gender. A small subclass of adjectives (13 members are known in the Western dialect) agree in gender, some of which are shown in Table 5. Other adjectives are invariant (e.g. *yaba* 'big', *ndom* 'bad', *waninggap* 'good'). The patterns of exponence largely follow those familiar from nouns with overt gender, with agreement marked by means of changes in the stem-final vowel, except for *VhV* 'ripe' which shows a unique pattern of vowel height harmony. Note that some of the adjectives are semantically incompatible with animates, whence the dashes in the table.

The forms of agreeing adjectives are much more regular than nouns with overt gender: Gender I and II consistently have /e/ and /u/ as their exponents, and their plural indicated by /i/; for inanimates, Gender III is largely indicated by /a/, while the pattern of homophony between the I/II plural forms and the Gender IV forms is observed again.

A remarkable exception from these regularities is the adjective 'small', whose forms are given in Table 6. This adjective is noteworthy for two reasons. First, it is the only word in the language that distinguishes singular and plural for Gender III and IV. This is done by means of the suppletive stems *isahih* and *wasasuy*, neither of which bear any phonological resemblance to the singular stem *papVs*. Following Corbett (1991: 168) we can say that 'small' is OVER-DIFFERENTIATED since it distinguishes a feature (number of inanimates) which is absent elsewhere in the system. However, one could also argue that 'small' does not show true agreement for gender, because the stems involved are suppletive. This is the approach taken by Durie (1986: 362), who – speaking of verbal number suppletion – argues that “suppletive stems select for rather than agree with the number of their argument”. Either way we look at it, 'small' has to be marked as an exceptional item,

and does not detract from the generalization that number as a nominal category is restricted to the animates, e.g. the members of Gender I and II.

Table 6: Gender agreement on ‘small’

	I	II	III	IV
SG	<i>papes</i>	<i>papus</i>	<i>papes</i>	<i>papis</i>
PL	<i>isahih</i>	<i>isahih</i>	<i>wasasuy</i>	<i>isahih</i>

Second, the stems used for ‘small’ in the plural are *isahih* and *wasasuy*, of which the former (which is also used as a noun meaning ‘children, young of animals’) is used not only for animates, but also for plural of Gender IV. This would be quite surprising if the syncretism between I/II plural and Gender IV noted so far (e.g. the demonstrative *ipe* covering I/II plural and IV) were merely a case of accidental homophony. Below we will see other cases where syncretisms between I/II plural and IV suggest a more profound relationship between the forms.

4 Agreement and participant indexing on verbs

The morphology of the Coastal Marind verb is complicated, and nominal gender plays a role within three of the inflectional sites of the verb: in a set of gender-agreeing prefixes, in the person indexing reflecting an UNDERGOER argument, and, somewhat marginally, in the indexing of the ACTOR argument of the verb. The gender-agreeing prefixes are the most straightforward, and behave largely like the non-bound agreeing items that we have seen so far. I will give some examples of gender agreement on the verb below. I contrast gender AGREEMENT with bound person marking on the verb, which I refer to as INDEXING. I will show below that these two phenomena behave quite differently in Coastal Marind, so it is convenient to make the terminological distinction between agreement and indexing in the description of the Marind verb.

Several inflectional prefixes are sensitive to the gender of some argument of the verb, although their main function lies in some other domain (e.g. tense-mode-aspect) so it is not appropriate to call them ‘gender prefixes’; rather, they are prefixes of which a sub-string happens to show agreement in gender. Let us take the prefix *Vp-* ‘ABSconditive’ as an illustration. Simplifying matters drastically, we can say that this prefix is used when the speaker is drawing attention to

some present state-of-affairs that is unavailable to the addressee, either because her attention is on something else, as in (5), or because she made a previous statement contradicting the state-of-affairs that actually holds, as in (6). The question of what argument of the verb controls the gender agreement in the prefixes is complicated, and I will not explore it here. Suffice to note that it is the (intransitive) subject in (5) that is the controller, whereas the Gender I agreement in (6) corresponds to the male recipient-like participant (other constellations would behave differently).

- (5) (Addressee standing facing away:)
kosi-awe up-Ø- kwayita!
 small-fish(II) ABSC:II-3SG.A- be.swimming.inside
 ‘A little fish is swimming in there!’
- (6) (Reply to “You should talk to him!”, female speaker:)
ep-ak-o- lay-e!
 ABSC:I-1.A-3SG.DAT- talk-IPFV
 ‘I am talking to him!’

Morphologically these prefixes are straightforward, since they have the same forms as the distal demonstrative *Vpe* (betraying a historical relationship), minus the final *-e*. The same holds, for example, for the continuative prefix *anVpand-* which most likely derives from the emphatic pronoun series *anVp* (cf. Table 4). Gender agreement in the prefixal complex then seems to be of relatively recent origin, resulting from the integration of free demonstrative and pronominal elements into the verb. Once more, the syncretism between the Gender I/II plural and Gender IV that was encountered in the nominal targets recurs in the prefixal agreement, so the Absconditive prefix *ip-* would be used with an animate plural controller, or with a noun from Gender IV. However, gender of verbal arguments triggers more dramatic alternations elsewhere in the verb, as we will now see.

I refer to bound person markers on the verb as participant indexing since they express person/number of participants of the verb directly – there is no need to say that the affixes in (7) ‘agree’ with some ellipsed or covert argument in the clause.

- (7) *no- y-amuk-e*
 1.A- 2SG.U-kill-IPFV
 ‘I’m going to kill you.’

There are also frequent mismatches ('disagreement') within person indexing of a type that is not found in the gender agreement. For example, many intransitive verbs use a suppletive stem with plural subjects, with the additional quirk that actor indexing then is obligatorily 3SG instead of 3PL. Compare the regular verb *dahetok* 'return', which employs the expected 3PL indexing, with the suppletive stem *nayam* 'come (plural subject)' (cf. *man* 'come (singular subject)').

(8) *na-* *dahetok*
3PL.A- return
'They returned.'

(9) *a-* *nayam*
3SG.A- many.come
'They came.'

For this reason I prefer to maintain a terminological distinction between agreement and indexing in the description of Coastal Marind. I use agreement about the prefixes whose shape reflect gender and which apparently derive from relatively recently incorporated pronominal elements, while indexing is used for the markers that primarily code person/number of various argument roles, and often require construction- or verb-specific rules for their description (as in the case with the suppletive verbs above). Having established this, we are now ready to explore how gender is manifested in person indexing on the verb.

Let us start by the indexing of undergoer participants. Since we will be concerned with the difference between animate and inanimate undergoers, the discussion will be restricted to 3rd person forms (1st and 2nd person are always animate). Undergoer indexing is realized by means of intricate changes in the verb stem, and is mainly pre-, in-, or suffixing depending on the conjugation class. I will not attempt to segment the verb stems in the interlinear examples below into morphemes; the morphological details are not of interest here.

Consider the verb 'put on a string', which has the following forms when the undergoer is animate:

- (10) a. *awe ah- laleh!*
fish(II) IMP- string:3SG.U
'String one fish!'
b. *awe ah- lalah!*
fish(II) IMP- string:3PL.U
'String many fish!'

With inanimates from Gender III, a different stem *lilig* is used (11). Recall that no number distinction is made for inanimates, so *lilig* can be used for one or several pieces of meat, fruits, or other inanimate entities as long as they are in Gender III.

- (11) *muy ah- lilig!*
 meat(III) IMP- string.inanimate
 ‘String the piece(s) of meat!’

With undergoers from Gender IV, however, the stem used with animate plurals, i.e. the 3PL stem *lalah*, is used (12). As in the previous example, there is no number distinction, so the cardinality of *baba* (a kind of grass, seeds of which are used for necklaces) has to be inferred from context.

- (12) *baba ah- lalah!*
 Job’s Tears(IV) IMP- string:3PL.U
 ‘String the *baba* seed(s)!’

It is remarkable that Gender IV nouns trigger the use of verb stems otherwise used for 3rd person animate plurals, since gender agreement is not manifested elsewhere in person indexing. No distinction is made between Gender I and II, and inanimate stems such as *lilig* generally look like separate lexemes rather than inflectional forms of the verb. Some more examples of alternations are given in (13).

- (13) Stem alternations according to undergoer
- | | | | |
|---------------------|------|------------------|-----------------------|
| a. ‘wrap’ | | | |
| Animate | 3SG: | <i>ambeh</i> | 3PL: <i>ambah</i> |
| Inanimate | III: | <i>ambam</i> | IV: <i>ambah</i> |
| b. ‘rub (bodypart)’ | | | |
| Animate | 3SG: | <i>hwahwetok</i> | 3PL: <i>hwahwituk</i> |
| Inanimate | III: | <i>hwahwid</i> | IV: <i>hwahwituk</i> |
| c. ‘eat’ | | | |
| Animate | 3SG: | <i>aheb</i> | 3PL: <i>hi</i> |
| Inanimate | III: | <i>yi</i> | IV: <i>hi</i> |
| d. ‘become’ | | | |
| Animate | 3SG: | <i>win</i> | 3PL: <i>in</i> |
| Inanimate | III: | <i>ay</i> | IV: <i>in</i> |

Such verbs differ in the degree of similarity between the different stems, but all employ the same stem for Gender IV undergoers as for 3PL animates. There seem to be no exceptions to this pattern, so if a verb is semantically compatible with both animates and inanimates, then the 3PL/IV stem sharing occurs, regardless of how the remainder of the paradigm is structured. Note also that there is no morphological resemblance to the agreement patterns that we observed for nominals: with the exception of stems like *hwahwituk* ‘rub many animates’ (e.g. when scaling fish) or ‘rub a Gender IV-item’ (e.g. a knee, *mig*), which shows the high vowels /i u/ associated with gender agreement (e.g. *ihu* ‘ripe:IV’), the vowel alternations seen within the nominal domain are absent. I take this to confirm that gender agreement and participant indexing are two quite distinct phenomena in Coastal Marind, and that they have different histories, which renders the conflation of animate 3PL and Gender IV across the two systems the more remarkable.

Finally, let us consider other types of participant indexing on the verb. There are three varieties of indexing, all realized by prefixes, in addition to the indexing of undergoers by means of stem alternations. These are indexing of actor, seen in examples (7)–(9) above, plus indexing of a recipient-like participant, and what can be described as affected possessor of an argument of the verb. I will not provide examples of the latter two, because inanimate arguments filling recipient- and possessor-like roles are extremely rare in the corpus, and it is not clear whether these indexing mechanisms interact with the gender membership of inanimate arguments. The data from actor indexing are more interesting, so let us have a look at it to see whether Gender IV nouns trigger 3PL indexing in this domain.

Sentences with inanimate nouns functioning as semantic agents are also exceedingly rare in my corpus, since argument NPs headed by such nouns mostly fill patient-like roles. I have made several attempts to elicit sentences in which various things belonging to Gender IV are in violent contact with an animate undergoer (such as fruit falling from a tree, hitting a bystander), i.e. verbs that usually provide a good frame for testing all person/number combinations of agent and patient. Speakers were consistent in reporting that only 3SG actor indexing is compatible with IV agents, as in (14).

- (14) *saley* *a-* *n-asib*
 inflorescence(IV) 3SG.A- 1.U-hit
 ‘The coconut inflorescence (fell and) hit me.’

If this were the whole story, agent indexing would finally provide an environment where Gender IV nouns were distinguished from animate plurals. However,

the generalization only seems to hold for the transitive agent-patient configuration: a small number of examples of agentive intransitives in my corpus, such as *esol* ‘make noise’ (15), unambiguously show 3PL actor indexing IV nouns (this has also been confirmed in elicitation).

- (15) *yaba-mesin i-pe t-i-k-at-n- esol-e*
 big-machine(IV) IV-that GIV-IV-PRS-PRSTL-3PL.A- make.noise-IPFV
 ‘The generator is making noise.’

Not even actor indexing is immune to the IV-as-animate-plural pattern, then. I take the difference in indexing between (14) and (15) to reflect semantic restrictions on what participants may be indexed on the verb, so that the inanimate coconut inflorescence in (14) is not enough of an agent to be properly indexed (with actor indexing then defaulting to 3SG, which is also the default for avalent verbs). The verb *esol* ‘make noise’ is less picky and admits its sole argument to be fully indexed, thus giving the 3PL prefix. (Recall that agreement is insensitive to number of inanimates, which means that ex. (15) is equally fine referring to one or more than one generator.)

Whatever the explanations for the subtleties of person indexing turn out to be, the data presented above are roughly consistent with the main point of this and the previous section: in all contexts where Coastal Marind, by various grammatical means, distinguishes between gender, number and animacy, nouns of Gender IV systematically pattern with plurals of Gender I and II. This is quite strange given the fact that inanimates do not show grammatical agreement according to their referential cardinality in the language (cf. example (3) above), which makes it difficult to claim that Gender IV should be considered ‘fixed plural’ nouns (*pluralia tantum*) instead of a gender. Below I will show that some tendencies in the assignment to Gender IV also are consistent with the *pluralia tantum* analysis, because they involve nouns that are *pluralia tantum* cross-linguistically. However, I will argue that this can at most be regarded as suggesting a diachronic relationship with *pluralia tantum* nouns, and that synchronically we must reject the description of the Gender IV nouns as *pluralia tantum* (§6).

5 Assignment and *pluralia tantum* as a possible origin for Gender IV

The basic principles behind the assignment of nouns to the four genders were given above: male humans are Gender I, female humans and all animals are Gender II, while inanimates are mostly in Gender III with a (large) residue in Gender

IV. I do not believe that there are any clear semantic rules for deciding which of the inanimates go into Gender IV, but there are some tendencies. The only semantic fields that are completely restricted to Gender III seem to be abstracts (e.g. *mayan* ‘language, issue, problem’, *sal* ‘taboo’), names of places and geographical features (*milah* ‘village’, *mamuy* ‘savannah’), and various intangibles (*matul* ‘shade’, *usus* ‘afternoon’). Other large semantic fields such as bodyparts and flora are split between Gender III and IV, with very few obvious subdomains assigned to one or the other (flowers is a subdomain that seems to belong to Gender IV). Artifacts are also divided between III and IV, with the only discernible patterns being that almost all bodily decorations are in Gender IV (*segos* ‘rattan girdle’, *himbu* ‘feathered hairdress’), as well as most recently introduced technology (airplanes, ballpoint pens, diesel generators).

Looking closer, we can see that some of the domains that Koptjevskaja-Tamm & Wälchli (2001: 630) identify as typically including pluralia tantum nouns show overlap with the members of Gender IV. These domains are: VARIOUS HETEROGENEOUS SUBSTANCES (“with many subdivisions”, e.g. Lithuanian *putos* ‘foam’), corresponding to Coastal Marind IV nouns such as *ndalom* ‘foam’, *ndakindaki* ‘bioluminescence’, *kangging* ‘layer of crushed seashells on the beach’ and *katal* ‘money’⁷; ARTIFICIAL OBJECTS WHICH ARE CLEARLY INTERNALLY COMPLEX (e.g. English *trousers*), corresponding to Coastal Marind decorations and modern technology in Gender IV; DISEASES “[that] manifest themselves as multiple visible symptoms/spots” (e.g. English *measles*), corresponding to names of skin diseases in Coastal Marind, which all turn out to be in Gender IV, such as *kambi* ‘tinea imbricata’, *dapadap* ‘tinea versicolor’ and *apupin* ‘pimple’.

While suggestive, these findings do not form any consistent pattern. The overlap is not found with other pluralia tantum domains such as names of festivities in Coastal Marind (e.g. German *Weihnachten* ‘Christmas’), and there are numerous exceptions, e.g. some artifacts that clearly qualify as internally complex (e.g. *kipa* ‘net’) are in Gender III rather than IV. It is also clear that – even allowing for some semantic latitude – the majority of nouns in Gender IV do not fit into any of Koptjevskaja-Tamm and Wälchli’s categories. I have found no reason why some names of trees are in Gender III, others in Gender IV, and it seems unlikely that plurality should have anything to do with the classification. Similarly, while it is conceivable that many bodyparts in Gender IV are somehow ‘plural’ (e.g.

⁷The noun *katal* has a primary use as a Gender III noun, then with the meaning ‘stone’. South New Guinea is almost completely devoid of stones, and it is extremely unlikely that one encounters two or more naturally occurring stones at the same occasion. The Gender IV noun ‘money’, on the other hand, usually occurs in collections of more than one rupiah banknote. This is an interesting case of cross-classification seemingly involving a difference in plurality.

put ‘feather’, *tatih* ‘hair’, *tiwna* ‘gums’, *halahil* ‘lungs’) there are plenty that are not (*ambay* ‘uvula’) and some bodyparts seem quite plural but belong to Gender III (*lul* ‘fur’). As pointed out by an anonymous reviewer, however, most languages with pluralia tantum have a fairly idiosyncratic assignment to the class, so the lack of consistency can hardly be an argument *against* the possibility of Gender IV being related to pluralia tantum.

If we consider there to be at least some tendency for ‘pluralia tantum concepts’ to be in Gender IV, this situation could be seen as consistent with a diachronic scenario where Gender IV started out as a class of pluralia tantum, but then acquired new members through some unknown (analogical?) process, resulting in a large, semantically heterogeneous residue gender, with a small core that still reflects the ‘plural semantics’ of the original pluralia tantum grouping. This scenario is only plausible if (pre-)proto-Anim (as-opposed to present-day Coastal Marind) had a number distinction among inanimate nouns, since this would be required for inanimate pluralia tantum nouns to come into existence. Also, we would expect to find some other Anim language that has been more conservative in this regard, and maintains a clearer semantically plural basis for the cognate fourth gender. Unfortunately, there is no systematic data on gender available from other Anim languages to see whether such semantics can be associated with Gender IV, nor is there any indication that proto-Anim had a number distinction among inanimates. For now this hypothesis remains purely speculative, and it can only be evaluated once there is more data on gender systems in other sub-branches of Anim. Still, I believe it is worth spelling out this hypothesis, since it has the merit of providing an explanation to the recurrent pattern of homophony between Gender IV and animate plurals, as well as the surprising phenomenon of the suppletive plural stems triggered by all Gender IV nouns.

Interestingly, a striking parallel to the Coastal Marind case is found in the Ok family, located in the New Guinean highlands. The Ok languages are probably very distant relatives of Coastal Marind and the other Anim languages as both families are proposed members of the large Trans-New Guinea phylum (Fedden 2011; Usher & Suter 2015). I believe that the Ok data support the idea that the similarities between the fourth gender of Coastal Marind (and other Anim languages) and what is described as pluralia tantum nouns in other languages are not coincidental, and perhaps that a diachronic relationship between these categories is plausible.

The best described Ok language, Mian, has a 4-gender system distinguishing Masculine, Feminine, and two inanimate genders – this is the same division as

in the gender systems of the Anim languages.⁸ The exponents of Masculine and Feminine resemble the ones found on demonstratives in Coastal Marind (Fedden 2011: 170, Usher & Suter 2015: 118): the Mian Masculine article =*e*, the Feminine =*o*, and M/F plural =*i* correspond to Coastal Marind Gender I *epe*, Gender II *upe* and Gender I/II plural *ipe* respectively. The phonological similarities might be due to chance, however, and I am not aware of any other evidence that the gender systems of the two families are cognate. Neuter 1 (the third gender) differs from the Coastal Marind inanimates in distinguishing singular and plural (SG =*e*, PL =*o*). The most interesting gender is the fourth (“Neuter 2”) which is invariant for number, and shows homophony with the plural of Neuter 1 (SG/PL article =*o*).

It is interesting that both Coastal Marind and Mian have one gender that shares their exponents with plurals, but note that the pattern of syncretism is different (homophony with inanimate plural in Mian, but with animate plural in Coastal Marind), and could have arisen by chance since both languages have relatively few vowels to choose from (5 in Coastal Marind, 6 in Mian). Speaking against accidental homophony is the fact that even in cases where several paradigm slots are filled by unpredictable gender exponents, Neuter 2 invariably patterns with the plural of Neuter 1 (Fedden 2011: 178–179).

A further argument against the possibility of chance homophony between the Mian Neuter 2 and the plural of Neuter 1 is the fact that the nouns that are assigned to Neuter 2 match the pluralia tantum domains listed by Koptjevskaja-Tamm and Wälchli quite well – better than the Coastal Marind Gender IV nouns do. Assigned to Mian Neuter 2 we find: places (e.g. *bib* ‘village, place’), heterogeneous substances (e.g. *difb* ‘rubbish’, *moni* ‘money’), body decoration (e.g. *amún* ‘hole in nosetip’), various abstracts and temporal nouns (e.g. *am* ‘day’), illnesses (e.g. *klō* ‘ringworm’), various artifacts (e.g. *itó* ‘tongs’, *aiglas* ‘glasses’) and bodyparts, most of which seem to consist of multiple parts (e.g. *abó* ‘testicles’, *amuntêm* ‘intestines, belly’, *wanáan* ‘feather’).⁹

Fedden does not consider the alternative analysis according to which the Neuter 2 nouns are pluralia tantum nouns belonging to Neuter 1, and I will not pursue that issue here.¹⁰ However, I interpret the parallelism between Coastal

⁸Sebastian Fedden (pers. comm.) adds the caveat that little is known about the gender systems of other Ok languages, so we do not know how representative the Mian system is for Ok in general. More descriptive work will be necessary for a fuller picture of the similarities and differences between the Anim and Ok gender systems.

⁹One instance of cross-classification is striking: Mian *bém* ‘worm’ (masculine gender) can also mean ‘noodles’, and then belongs to Neuter 2; cf. Coastal Marind *alalin* ‘tapeworm’ (Gender II), meaning ‘noodles’ in Gender IV.

¹⁰The reader is referred to Corbett et al. (2017).

Marind Gender IV and Mian Neuter 2 as further evidence that the connection between fixed plural and fourth gender in Coastal Marind is no coincidence, as this pattern would not arise independently in the two languages by chance. At this stage it is impossible to tell why the gender systems of Ok and Anim share these similarities. The two families are most likely related as members of the Trans-New Guinea stock, but this relationship is extremely distant and must go back long in time. There is at present no evidence that the gender systems were inherited from some common ancestor, although this would account for the similarities in the gender exponents mentioned above. One could also speculate that the gender systems evolved in parallel at a time when speakers of Ok and Anim languages were in closer contact, but more research remains to be done before we can say anything about the contact between these ancestral populations.

Regardless of whether the similarities between Ok and Anim are the result of common inheritance or contact, it seems to me that the simplest explanation is that both the Anim fourth gender and the Mian Neuter 2 developed from pluralia tantum nouns, which explains e.g. the use of suppletive agreement targets in Coastal Marind and the fact that many of the Mian Neuter 2 nouns (and some of the Gender IV nouns in Coastal Marind) have meanings that are found among pluralia tantum cross-linguistically. This hypothesis can be tested only through more descriptive and comparative work on the two families. Even if it is correct, it would still remain to be shown in detail how a 3-gender system with a large number of pluralia tantum nouns can develop into a 4-gender system lacking number distinction in inanimates, as in present-day Coastal Marind.

6 The synchronic analysis of Gender IV

Having suggested that the Coastal Marind Gender IV originated as a pluralia tantum class, we now need to address the synchronic status of Gender IV. Should we maintain the 4-gender analysis, or opt for the more economical 3-gender analysis according to which the members of the former fourth gender are Gender I or II nouns that just happen to be lexically specified as plural? I believe that this is an important analytical question – not a mere question of which labels to stick where – since the two possible descriptions result in wildly different systems in terms of assignment.

The literature contains some discussion of the possibility of analyzing pluralia tantum as a separate gender, in various languages. Corbett (2012: 233–239) provides instructive discussion of such suggestions for Cushitic, Chadic and Russian, and argues that the pluralia-tantum-as-gender analysis is untenable for all

the proposed cases (i.e., the opposite of the established descriptions of Coastal Marind and Mian). For example, Zaliznjak (1964) proposed to describe Russian pluralia tantum nouns such as *sani* ‘sledge(s)’ as making up their own gender, since they form a unique agreement class within the system. Corbett (2012: 237–238) points out that the same analysis applied to Bosnian/Croatian/Serbian would produce no less than three extra genders, since this three-gender system (as opposed to Russian) has separate plural forms for each gender, each of which contains pluralia tantum that would be reanalyzed as separate genders. This is unacceptable, so Corbett rejects the analysis for Russian as well.

On a more general level, Corbett argues that pluralia-tantum-as-gender analyses are misinformed, since “the special behaviour which creates the extra agreement class is not *gender* but *number*” (Corbett 2012: 238; emphasis in original). According to Corbett, proponents of pluralia-tantum-as-gender analyses mistakenly think that since pluralia tantum nouns need to be lexically specified for a morphosyntactic value (in this case number), they are just like other nouns – which are also lexically specified, for gender – and therefore belong to a gender of their own. Instead, the correct way is to treat them as exceptionally specified for number, and leave the gender system as it is. I interpret Corbett’s remarks as a principled stance against analyses claiming that pluralia tantum nouns make up a gender.¹¹

In spite of Corbett’s reservations, I prefer to maintain the Drabbian analysis of Gender IV as a gender, and not as pluralia tantum of Gender I or II, although I concede that the morphosyntactic evidence for this analysis is somewhat nebulous. We saw that the exponents of Gender IV agreement are identical to the ones marking the plural of Gender I and II, no matter how irregular the alternations of the relevant target are. Verb stem alternations indexing undergoers likewise treat Gender IV and plurals of I/II identically, despite being seemingly unrelated to the agreement patterns of demonstratives and other categories in the non-verbal domains. The only domain where Gender IV nouns do not always pattern with I/II plural is actor indexing (and, possibly, recipient and possessor indexing) on verbs; however, I suspect that this reflects some general constraint against inanimates filling such participants roles, so the diagnostic role of these constructions is unclear.

But consider the consequences of abandoning the gender analysis in favour of the pluralia tantum analysis. If the members of Gender IV are considered plu-

¹¹In fact, Corbett says explicitly that this is what he means: “Having not accepted Zaliznjak’s careful and considered analysis of certain Russian pluralia tantum nouns as an additional gender value, I am even less ready to entertain other less convincing proposals along similar lines.” (p. 238).

ralia tantum, they would make up an unexpectedly large portion of the lexicon. Assuming that the currently available numbers (Table 3) are representative of gender membership, one out of five nouns would be pluralia tantum. This seems strange from the European perspective, but sheer frequency can hardly be a decisive argument. More seriously, the system of semantic assignment (males in I, females and animals in II, inanimates in III and IV) would break down, since we would have to claim that Gender I and II contain a fairly random mix of animates and inanimates (all of which happen to be pluralia tantum), with non-pluralia tantum inanimates confined to Gender III.

The resulting system would also be typologically odd in the way it fails to align with the Animacy Hierarchy (Smith-Stark 1974, Corbett 2000: 55ff.). The hierarchy states that if there is a difference in the availability of a number distinction between e.g. animates and inanimates, then it will be animates that make the distinction and inanimates that lack it. Corbett (2000: 59) cites Coastal Marind as an example of a language with a clear split between animates (which trigger singular/plural agreement) and inanimates (which make no distinction according to number). In the new system, we would have to say that number is relevant for a fifth of the inanimates, although these happen to be lexically specified for plural only.

I take these consequences to be unacceptable, so the 4-gender analysis must be preferred. This comes at the price of not adhering to a strictly morphosyntactic approach to the identification of genders in Coastal Marind, because the formal facts alone do not provide clear evidence that the four-gender description is to be preferred over a three-gender description with a large number of pluralia tantum.

7 Conclusion

Besides the descriptive contribution of this paper (most of which can be extracted, with some effort, from Drabbe's grammar), I consider the main points to be (1) the evidence that Usher & Suter's (2015) suggestion that overt, stem-internal gender marking originated from umlaut also explains patterns in the distribution of stem-final vowels of invariant nouns within Gender III and IV; and (2) the description of the ambiguous status of the nouns in Gender IV, which led me to speculate that an earlier 3-gender system was extended into a 4-gender system, and that the 4th gender originally was a grouping of pluralia tantum nouns. As mentioned above, the idea that gender systems can be extended through the reinterpretation of a non-gender feature as gender is not new, and if the suggestions based on Coastal Marind data are correct, the Anim languages (and the distantly

related Ok family) would provide a clear case where a gender system became more complex because of a very specific type of interaction with number.

Acknowledgments

I am very grateful to the Coastal Marind speakers with whom I work, especially Petrus Kilub and Rafael Samkakai who spent many hours with me rechecking the gender of nouns. I wish to thank Matthew Lou-Magnuson for suggesting the use of resampling methods in §2 and Thomas Hörberg for implementation; I alone am responsible for mistakes in the interpretation of the data. I also acknowledge the extremely helpful comments given to me by Edgar Suter, Sebastian Fedden, Bernhard Wälchli, Francesca Di Garbo, Lea Brown, and the anonymous reviewers.

Special abbreviations

The following abbreviations are not found in the Leipzig Glossing Rules:

A	actor	PRSTL	presentational
ABSC	absconditive	U	undergoer
GIV	givenness marker		

References

- Aikhenvald, Alexandra Y. 2000. *Classifiers: A typology of noun categorization devices*. Oxford: Oxford University Press.
- Brugmann, Karl. 1891. Zur Frage der Entstehung des grammatischen Geschlechts. *Beiträge zur Geschichte der Deutschen Sprache und Literatur* 15. 523–531.
- Corbett, Greville G. 1991. *Gender*. Cambridge: Cambridge University Press.
- Corbett, Greville G. 2000. *Number*. Cambridge: Cambridge University Press.
- Corbett, Greville G. 2012. *Features*. Cambridge: Cambridge University Press.
- Corbett, Greville G., Sebastian Fedden & Raphael Finkel. 2017. Single versus concurrent feature systems: Nominal classification in Mian. *Linguistic Typology* 21(2). 209–260.
- Dahl, Östen. 2000. Animacy and the notion of semantic gender. In Barbara Unterbeck (ed.), *Gender in grammar and cognition*. Vol. 1: *Animacy and the notion of semantic gender: Approaches to gender*, 99–115. Berlin: Mouton de Gruyter.
- Drabbe, Peter. 1955. *Spraakkunst van het Marind: Zuitkust Nederlands Nieuw-Guinea* (Studia Instituti Anthropos 11). Wien-Mödling: St. Gabriël.

- Durie, Mark. 1986. The grammaticization of number as a verbal category. In Vasiliki Nikiforidou, Mary VanClay, Mary Niepokuj & Deborah Feder (eds.), *Proceedings of the 12th annual meeting of the Berkeley Linguistics Society*, 355–370. Berkeley: Berkeley Linguistics Society.
- Edwards-Fumey, Deborah. 2007. *The verb subject prefix in Kuni*. Universität Bern MA thesis.
- Evans, Nicholas. 2012. Even more diverse than we had thought: The multiplicity of Trans-Fly languages. In Nicholas Evans & Marian Klamer (eds.), *Melanesian languages on the edge of Asia: Challenges for the 21st century* (Language Documentation & Conservation Special Publication No. 5). Honolulu: University of Hawai'i Press.
- Evans, Nicholas, Wayan Arka, Matthew Carroll, Yun Jun Choi, Christian Döhler, Volker Gast, Eri Kashima, Emil Mittag, Bruno Olsson, Kyla Quinn, Dineke Schokkin, Philip Tama, Charlotte van Tongeren & Jeff Siegel. 2018. The languages of Southern New Guinea. In Bill Palmer (ed.), *The languages and linguistics of the New Guinea area* (The World of Linguistics 4), 641–774. Berlin: De Gruyter Mouton.
- Fedden, Sebastian. 2011. *A grammar of Mian*. Berlin: Mouton de Gruyter.
- Foley, William A. 1986. *The Papuan languages of New Guinea*. Cambridge: Cambridge University Press.
- Geurtjens, Hendrik. 1933. *Marindineesch-Nederlandsch woordenboek* (Verhandelingen van het Koninklijk Bataviaasch Genootschap van Kunsten en Wetenschappen 71). Bandung: Nix.
- Koptjevskaja-Tamm, Maria & Bernhard Wälchli. 2001. The Circum-Baltic languages: An areal-typological approach. In Östen Dahl & Maria Koptjevskaja-Tamm (eds.), *Circum-Baltic languages: Typology and contact*, vol. 2, 615–750. Amsterdam: John Benjamins.
- Olsson, Bruno. 2017. *The Coastal Marind language*. Singapore: Nanyang Technological University dissertation.
- Smith-Stark, Thomas Cedric. 1974. The plurality split. In Michael W. La Galy, Robert A. Fox & Anthony Bruck (eds.), *Papers from the 10th regional meeting of the Chicago Linguistic Society*, 657–661. Chicago: Chicago Linguistics Society.
- Usher, Timothy & Edgar Suter. 2015. The Anim languages of Southern New Guinea. *Oceanic Linguistics* 54(1). 110–142.
- Zaliznjak, Andrej A. 1964. K voprosu o grammatičeskix kategorijax roda i oduševlennosti v sovremennom ruskom jazyke. *Voprosy jazykoznanija* 4. 25–40.

