HIGHER NUMERALS IN SEVERAL MICRONESIAN LANGUAGES

Sheldon P. Harrison
Frederick H. Jackson

0. INTRODUCTION

In a number of Micronesian (MC)[1] languages, there exist monomorphemic numbers for multiples of ten, reaching 10^9 in the case of Ponapeic (PP) languages. (In one Polynesian (PN) language, Nukuoro (NUK), a similar series reaches 10^{10}.) Reports of the existence of higher number systems of this sort are often treated with a great degree of scepticism. A typical reaction is that of Elbert and Pukui (1979:160) who, in considering the Hawaiian (HAW) numbers lau, mano, kini, and lehu,[2] state that "they are used poetically as nouns indicative of great numbers".

They note (1979:161)[3] that "elsewhere these quantities are somewhat fancifully translated as 400, 4,000, 40,000, and 400,000. It is doubtful that actual counts of this magnitude were ever made.... Of even greater magnitude is nalowale, usually translated *lost* but sometimes considered a number equal to ten lehu, which is four million. It is inconceivable that people counted that many".

While it is undoubtedly true that there were no objects in any society that one would want to count into the hundreds of thousands or millions, we do not find it inconceivable that an abstract mathematics with linguistic means of representing quantities of such magnitude or beyond could have developed in traditional Oceanic societies. In our experience, speakers of MC languages will present the higher number series in order [10^1 through 10^x, where the value of x is language specific] much as we would recite the numbers 'one' through 'ten', making it clear through the counting procedure itself that each member of the series is agreed upon as a ten-multiple of the immediately lower member. Though skepticism on the part of those encountering these systems is perhaps understandable, any a priori rejection of the possibility that such systems could exist in a traditional society is in our opinion unwarranted.

In Section 1. of this paper, we present some background information on the numeral systems of MC languages, with particular attention paid to the status of the category countable base (numeral classifier). Section 2. considers the ten-power bases of MC languages, with some attempt at reconstruction of these bases for earlier periods in the history of these languages (Section 2.1.) and a somewhat more detailed discussion of MC forms for 'ten' (Section 2.2.). In

---

Byron W. Bender, ed. *Studies in Micronesian Linguistics,* 61-79.
Pacific Linguistics, C-80, 1984.

Section 3., we consider very briefly the ten-power bases of other Oceanic (OC) languages. Section 4. provides a speculative first attempt at an account of the evolution of the ten-power base series.

1. MICRONESIAN NUMERAL SYSTEMS

This section gives a brief review of those features of the grammar of numerals in MC languages that have some bearing on the interpretation of the ten-power base series found in many of these languages.

All MC languages reflect the Proto-Oceanic (POC) numbers 'two' through 'nine' as reconstructed in Pawley (n.d.),[4] with the sole exceptions of Marshallese (MRS), which has replaced the numbers 'six' through 'eight' by other forms, and Trukese (TRK), which has similarly replaced the numbers 'three' and 'four' in unit counting, although not in serial or higher-order counting. (These language-specific developments are not directly relevant here, however.) Except in the serial (enumerative) counting systems of some languages, these numbers always appear as first component of a bimorphemic numeral, the second component of which is a *numeral classifier*. All MC languages, again with the exception of MRS,[5] have a numeral classification system, though the number of classifiers varies from language to language. A minimal binary classification system is found in Kosraean (KSR) (Lee 1975). At the other extreme, TRK (Sugita in preparation, Benton 1968) and Kiribati (Gilbertese) (KIR) (Harrison in preparation, Trussel 1979) have some ninety classifiers, not all of which however, are in common use. Rehg (1981) lists twenty-nine classifiers for Ponapean (PNP). Sohn (1975), in a "non-exhaustive list", gives thirty-eight Woleaiean (WOL) numeral classifiers. The following table gives the numerals 'two' through 'nine' in the general counting system (with reflexes of the classifier Proto-Micronesian (PMC) *-ua < POC *pua *fruit*) and in the animate counting system (with reflexes of the classifier PMC *-manu[6] < POC *manu(k) *bird, creature*) in KIR, PNP, and WOL, representative of three of the five first-order MC subgroups.

Though we have employed the widely-used term *numeral classifier* above, it is perhaps not as appropriate for these morphemes, given their semantics, as a more neutral term such as *countable base*. Some countable bases are 'qualitative' (what Lyons 1977:460ff terms 'sortal') selected in terms of a classification of objects in the world on the basis of salient features of the inherent semantics of the objects being counted. For example:

KIR	uoua te boki *two-general art book*	*two books*
	uoman ataei *two*-animate *child*	*two children*
	uakai te nii *two-plant art coconut*	*two coconut trees*
	uaai te tikareti *two-long art cigarette*	*two cigarettes*

Other countable bases do not reflect a hyponymic classification in this sense, but are set labels for individual classes of objects (what Benton 1968 terms 'repeaters'):

	KIR		PNP		WOL		PMC	
	General	Animate	General	Animate	General	Animate	General	Animate
two-	uoua	uoman	riau	riemen	riuwauw	riuwemal	*ruwa-ua	*ruwa-manu
three-	teniua	teniman	siluh	silimen	seliuw	selimel	*telú-ua	*telú-manu
four-	aua	aman	pahieu	pahmen	faauw	faamal	*fa(a)-ua	*fa(a)-manu
five-	nimaua	niiman	limau	limmen	limauw	limmal	*lima-ua	*lima-manu
six-	onoua	onomaun	weneu	wenemen	wolouw	wolomal	*ono-ua	*ono-manu
seven-	itiua	itimaun	isuh	isimen	fisiuw	fisimel	*fitú-ua	*fitú-manu
eight-	waniua	waniman	waluh	welimen	waliuw	walimel	*walú-ua	*walú-manu
nine-	ruaiua	ruaman	duwau	duwemen	tiweuw	tiwemal	*s'iwa-ua	*s'iwa-manu

TRK	rúwéchamw *two-head*	chamwen *head-of*	iik *fish*	*two fish heads*
	rúwéché *two-flat object*	chéén *leaf-of*	wuuch *banana*	*two banana leaves*
	rúwéféw *two-round object*	faaw *rock*		*two rocks*
	rúúfutuk *two-meat*	futuken *meat-of*	kkow *cow*	*two lumps of beef*

However, in most MC languages, a large proportion of all countable bases are 'quantitative' (what Lyons terms 'mensural'), referring to enumerable (measurable) quanta of time or space, to containers, sets or agglomerations, or to parts of a whole. For example:

KIR	uabong	*two days*	(bong *day*)
	uangaa	*two fathoms*	(ngaa *fathom*)
	uamwangko	*two cups*	(mwangko *cup*)
	uarinan	*two rows*	(rinan *row*)
	uaatao	*two layers*	(atao *layer*)
	uamwakoro	*two pieces*	(mwakoro *piece*)

The category *quantitative countable base* in MC languages also includes bases with a fixed numerical value, the ten-power bases that are the focus of this article. The morphology and syntax of these items is, in most languages, not distinct from those of other countable bases. This observation will prove significant in our account of the evolution of the ten-power base series (see Section 4.). Some examples are:

KIR	uangaun	*twenty* *two-unit of ten*	(ngaun	*unit of ten*)
	uabubua	*two hundred*	(bubua	*unit of hundred*)
	uangaa	*two thousand*	(ngaa	*unit of thousand*)

1.1. Serial (enumerative) counting and the number *one*

With the exception of MRS, all MC languages (and many other OC languages as well)[7] possess a serial or enumerative counting system, which, in MC languages, does not involve countable bases (classifiers). The numbers in this system are not used as nominal adjuncts; they are used either in abstract counting or in enumerating a series. The following table presents the serial counting systems from *one* to *ten* of representative MC languages.[8]

Serial/enumerative counting							
	KIR[9]	KSR	PNP	MOK	TRK	WOL	PMP
one	(te)era	sra	ehd	oahd/ehd	eet	yet [yeetA]	*-sa
two	uua	lo	(a)ri	(a)ri	(té)rúú	riuw [rúúwA]	*rua
three	teen	tol	(e)sil	(e)jil	één	yel [yeelI]	*telú
four	aanga	ahng	(e)peng	(oa)poang	fáán	fang [faangI]	*fangi/*fanga
five	niima	luhm	(a)lim	(a)lim	niim	lim [liimA]	*lima
six		on	(o)un	(o)hn	woon	wol [wool0]	*ono
seven		it	(e)is	(e)ij	fúús	fis [fiisI]	*fitú
eight		oal	(e)wel	(a)wal	waan	wal [waalI]	*walú
nine		yuh	(a)du	(a)du	tiiw	tiw [tiiwA]	*s'iwa
ten[10]	ikoa	singuhul	eisek	eijek	engoon	seig	

With the exception of the serial counter for *four*, which apparently reflects a PMC *fangi (KIR aanga suggests an earlier *fanga), the serial counting forms for *two* through *nine* in all the languages reflect established POC reconstructions for the cardinal numbers. The forms for *one* appear to reflect POC *(n)sa (Pawley, n.d.) — with all the languages except KSR showing clear evidence of a prefixed increment. Pawley has also reconstructed POC *ta(n)sa *one*..., but none of the MC prefixes is a clear reflex of *ta-. If, however, the prefix on the forms for *one* were reconstructed as *te-, as suggested by KIR teera,[11] all forms could be accounted for. Moreover, there is substantial evidence, in the Solomons at least, that such a form existed outside MC as well.

A second prefix must also be reconstructed to account for the forms for *two* through *nine* in the PP languages PNP and Mokilese (MOK). This prefix, which apparently is not attested outside PP, can be reconstructed as PPP *a-.[12] The data also suggest the need for yet a third prefix to be reconstructed to account for TRK térúú *two*, but it is our belief that the té- reflects the final *sa of *te-sa *one* copied onto *ruwa *two*: yeeté-rúú. (In TRK, final vowels are deleted at the ends of words, but not before enclitics or suffixes. We suggest that *te-sa-ruwa *one-two* was treated in TRK as a single 'phrase', before final vowel deletion applied.)

The forms for *one* with countable bases involve a prefix that in most languages is also a reflex of PMC *te-:

KIR	teuana[13]	*one* (general object)
	temanna	*one* (animate creature)
MRS	juon	*one*
TRK	eew	*one* (general object)
	emen	*one* (animate creature)
WOL	seuw	*one* (general object)
	semal	*one* (animate creature)

However, although PP forms for *one* also involve a prefixed countable base, the prefix is more likely to be a reflex of PPP *a- than of PMC *te-, unless one assumes irregular assimilatory lowering in forms like MOK apas *one* (long object) (cf. Harrison 1976, Rehg this volume):

MOK	ew	*one* (general object)
	emen	*one* (animate creature)
	apas	*one* (long object)
	ekij	*one* (piece)

One must, then, reconstruct the following forms for *one*:

i) PMC *-sa *one* (serial counting), which is a reflex of POC *(n)sa *one*.

ii) PMC *te- *one*, which is affixed in KIR and the Trukic (TK) languages to *-sa in serial counting, and in MRS, KIR, and TK to countable bases. It is likely that *te is also related to the KIR common article te and to a demonstrative prefix in TRK (see Bender 1981 and Harrison in preparation). Similar forms, though usually ta, are attested in the New Hebrides and the south-east Solomons. The isomorphism between the KIR article and the number *one* is unlikely to be an accident, though it is difficult to determine which function, if either, is historically prior.

iii) PPP *a- 'unit and serial prefix', which may, however, also prove to be related to the a- prefix appearing with demonstratives in KIR (aei *this*, anne *that*, aarei *that*).

2. TEN-POWER COUNTABLE BASES IN MICRONESIAN LANGUAGES

As already mentioned, morphemes for *ten*, *hundred*, etc. in MC languages are countable bases that combine with the number prefixes *one* through *nine* to form numerals.[14] For example, with the countable base *ten*, we find:

	KIR (-bwii *ten*)	PNP (-isek *ten*)	WOL (-ig *ten*)
ten	tebwiina	eisek	seig
twenty	uabwii	rieisek	riuweig
thirty	tenibwii	silihsek	seliig
forty	abwii	pahisek	faaig
fifty	nimabwii	limeisek	likeig
sixty	onobwii	weneisek	woloig
seventy	itibwii	isihsek	fisiig
eighty	wanibwii	welihsek	waliig
ninety	ruabwii	duweisek	tiweig

Higher numbers such as, for example, *four hundred and thirty five* are formed in a left-to-right sequence beginning with the highest appropriate power of ten. In some languages the numerals are conjoined, while in others they are juxtaposed: Thus:

KIR abubua ao tenibwii ma nimaua *435*
4-100 and 3-10 and 5-general

MOK pahpwiki jilihjek limoaw *435*
4-100 3-10 5-general

WOL faabiugiuw me seliig me limauw *435*
4-100 and 3-10 and 5-general

Unlike Indo-European ten-power numerals, which, except for relatively recent forms like 'million', do not exceed 10^3 (*thousand*), the ten-power bases of some MC languages go as high as 10^9. The following chart presents the unit numeral for the ten-power morphemes (i.e., 1 x 10^2, etc.) in a representative sample of MC languages:[15]

KIR	MRS	KSR	PNP	TRK	WOL	CRL	
tebwiina	jo͞noul	singuhul	eisek	engoon	seig	seigh	10^1
tebubua	jibukwi	siofok[16]	epwiki	epwúkú	sebiugiuw	ebwúghúw	10^2
tengaa	jerapen[17]		kid	engérέw	sangeras	sangaras	10^3
terebu			nen	ekit	sen [sennA]	ssel	10^4
tekuri			lopw		selob		10^5
teea			rar		sepiy		10^6
tetano			dep		sengit		10^7
tetoki			sapw		sangerai		10^8
			lik				10^9

Although the extent to which the entire set is known by all speakers of a language seems to vary, in no case is it difficult to find speakers willing to volunteer the full series. Furthermore, as mentioned earlier, speakers are in total agreement as to the values of the numbers.

In some languages, numbers above 10^3 are more usually formed on the English pattern (often with Eng. borrowings like KIR mirion *million* and birion *billion*. For example:

KIR	teningaa	*3,000*
	tenibwii tengaa	*30,000*
	tenibubua tengaa	*300,000*
	teniua te mirion	*3,000,000*
	tenibubua ao tenibwii tengaa	*330,000*
	tenibubua ao tenibwii ma teniua tengaa	*333,000*

In the traditional KIR system, '300,000' would be rendered

KIR	tenikuri	ao	tenirebu	ao teningaa
	3-10^5	*and*	*3-10^4*	*and 3-10^3*

2.1. Reconstructing PMC ten-power bases

Ignoring the base 'ten' itself (see Section 2.2.), it is possible to reconstruct five of the ten-power base series on the basis of cognates found in at least two lower-order subgroups of MC (ULI = Ulithian, PUL = Puluwat):

PMC	*pukua	*k(u,i)si[18]	*lopwa	*sep(u,i)	*nena
KIR	bubua	kuri		rebu	
MRS	bukwi	??jo-kde-n		??dep	
KSP	foko				
PNP	pwiki	kid	lopw	dep	nen
TRK	pwúkú	(k)kit			
CRL	bwúghúw				sse-l
ULI	buxuy				
WOL	biugiuw		lob[lobᵂA]		se-n [sennA]
PUL	pwúkúw	kkit			

PMC *púkua *hundred* is reflected throughout MC. It is problematic only in KIR, where one must assume an irregular change *k > b if the KIR form is in fact cognate. PMC *k(u,i)si is also widely distributed, though its value is not consistent (10^5 in KIR, 10^3 in PP, and 10^4 in TRK and PUL). MRS jokden *ten pairs* is doubtful, since MRS d is not a regular reflex of PMC *s. PMC *lop'a is reflected only in PP and WOL. PMC *sep(u,i) is reflected in KIR as 10^4 and in PP as 10^7.

The reconstruction PMC *nena requires some comment. It is reflected in TK and PP with the same value (10^4). In the former, however, the reflexes at first glance might appear to be less than secure. For WOL sen [sennA], we hypothesise loss of *e between identical consonants (i.e., *te-nena > se+nnA), which is a regular development in TK. This same process is assumed in the development of CRL ssel. In this form, however, the mora count has been redistributed, so that it is now the initial consonant that is geminate (where a geminate consonant counts as a mora): *te-nena > se+nna > ssena > ssel. This type of development, too, is widespread in Trukic.

Two additional reconstructions can be made for PTK:

PTK	*t''e-ngarat''u	*(p)piya
TRK	e-ngérew	
CRL	sa-ngaras	ppiy
ULI	sa-ngaras	ppiy
WOL	sa-ngeras	piy
PUL	ye-ngeráy	

PTK *(p)piya is in fact a noun meaning *'sand'* (see Section 4.). PTK *t'e-ngarat'u *10*^{3} has no cognates elsewhere in MC but is widely reflected elsewhere. It is clearly a reflex of PAN *Ratus, POC *Ratu *hundred* (Blust 1972), with the unit prefix PMC *te and reflex of the numerical ligature POC *nga (see Pawley, n.d.). KIR -ngaa *10*^{3} is a possible cognate for PTK *ngarat'u *10*^{3} but only if one assumes a very irregular loss of the historical final syllable. (The expected KIR form would be *ngaati).

It is significant that WOL shows a contrast between sangaras *1,000* and sangerai *100,000,000*. The former, as already discussed, is a reflex of POC *nga-Ratu *100*, while the latter is undoubtedly related to PPN *lau *countless, indefinite number'* (Biggs, et al. 1970), which is taken by Pawley (n.d.) to reflect his POC *dau *hundred, unit of hundreds*. So far as we are aware, WOL is unique in independently reflecting both of these POC reconstructions. The failure of the other TK languages to reflect both forms, however, may be the result of recent phonological developments which would result in *Ratu and *dau being reflected identically: TRK and Mortlockese (MRT) engérew and PUL yengeráy are as likely to be reflexes of one as the other POC reconstruction. (As Codrington (1885:249) noted, the isomorphism between the word for 'hundred' and that for 'leaf' [POC *dau] is probably not accidental. Note, however, that all MC languages reflect a POC nasal grade *ndau *leaf* but oral grade *dau *hundred* where this item is reflected.)

In KIR, three (four if KIR ngaa has TK cognates) of the ten-power bases above 'ten' (10^2 through 10^5) have cognates in other MC languages. Of the remaining three, tano and toki have nominal interpretations in KIR, the former meaning *sand, soil, ground* and the latter *end, limit*. In PP, sapw *10*^{8} also means *land* and lik *10*^{9} means *outside*. PNP rar *10*^{6}, as a noun, has the meaning *finger coral*. It is possible that this latter isomorphism is not accidental, given the nature of 'finger coral', an object that is frequently found in small fragments on the shore. MOK evidence with regard to this item is problematic, however, since *finger coral* is MOK lar while *10*^{6} is MOK rar. All the TK forms have numeral cognates, either intra-MC or extra-MC, except PTK *(p)piya *sand*, WOL -ngit *10*^{7} and CRL -púngút *10*^{5}. (It is possible, however, that these latter forms are related in some way to Uripiv ongut *100* and similar forms in the northern New Hebrides (Ray 1926).)

It is interesting to note that in the case of those higher power numerals with independent nominal interpretations, although the forms themselves are not cognate in a strict sense, they appear to reflect similar semantics. Thus the highest ten-power bases of KIR and PP (KIR toki *end*, PNP lik *outside*) both carry the sense of 'limit, extreme'. In these same languages the immediate lower ten-power base (KIR tano *sand, soil*, PNP sapw *land*) also have similar nominal interpretations, as does PTK *(p)piya *sand*, the highest power base in CRL and ULI and a relatively high one (10^6) in WOL. These observations suggest that a similar semantic may govern the ten-power base systems of all MC languages, even in cases where the forms themselves are not cognate.

2.2. Micronesian forms for 'ten'

The following forms for 'ten', or for groups of ten, are extant in MC languages (PUA = Pulo Annian):

KIR	tengaun	tebwiina	?ikoa	
MRS	joṉoul			
KSR	singuhul[19]			
PNP	ngoul		ehk	eisek
TRK	engoon[20]		-ik	
PUL	yengool		heeyik	
CRL		ääf	seeigh	
WOL	sengaul	seyaf	seig	
ULI			seyex	
PUA			deikI	
MRT	yengool		sɛɛk	

Reflexes of PMC *ngaulu *ten* are found in all subgroups of MC, and serve as the sole forms for 'ten' in MRS and KSR. The KIR reflex is used in all counting systems except the general (-ua) system (in which case KIR tebwiina *ten* is used). The PNP reflex is restricted to counting days, food prepared in an earth oven, multi-stemmed plants, and small pieces.[21] WOL sengaul has the restricted interpretation *ten groups*. Cognates of PMC *ngaulu *ten* are widely distributed outside MC as the number 'ten'. Its base is reconstructed as POC *pulu *ten*, usually reflected with a unit prefix and the ligature *nga, parallel to POC *Ratu and *dau.

CRL ääf *ten* does not appear to be current, though reported in Fritz (1911). WOL seyaf means *ten pieces of copra* and is also used in counting coins and valuable shells. The KIR root -bwii *ten* may be related to MOK (e)-pwi *some, several* and to PP *pwihn and TK pwii *group* (with PP languages showing a reflex of the suffix *-na). (This latter form has been reconstructed as PMC *p'utu(-na) *group, flock, school*, largely on the basis of MRS bwijin *group* (PMC *t > MRS j). Though loss of PMC *t is not unexpected in PP and some TK languages in this environment, it would be unusual for KIR.)[22]

The history of the remaining forms for 'ten' is complex. For the TK forms and PNP ehk we reconstruct a base *-ke *ten*. The PNP form derives from this base, with a unit prefix (probably *a-) by regular historical changes (see further below). The TK forms require us to postulate an optional prefix i-[23] (PTK *t''e-i-ke *ten*) whose source and function is unclear. However, these same elements, in a different order, also yield PNP eisek *ten* (< PPP *a-i-te-ke).

Although TRK, PUL, and MRT do not provide direct evidence for *-ke in the forms for *twenty* (rúwe), *fifty* (lime), *sixty* (wone), and *ninety* (ttiwe) (i.e. when the number morphemes end in a non-high vowel), the loss of *k in this environment is not unexpected: e.g. *ruwa-ke *twenty* > *rúwa-e > rúwee > rúwe. After historical high vowels, PNP -ehk has the alternate -akan (PNP siliakan *thirty*, isiakan *seventy*, weliakan *eighty*), the shape of which is problematic. The final -n may be a reflex of the *-na suffix, but why it should appear in these forms is unclear, since it is usually restricted to units (*one, ten*, etc.) rather than in multiples. The low vowel of -ka- is likewise unexpected. KIR ikoa *ten (pair counting)* is only questionably a reflex of an earlier *i-ke, since no account can be given for the final a of the KIR form.

The reconstructed *-ke *ten* appears to have extra-MC cognates, but with the value *one*; for example, the form ke *one* is widespread in the south-east Solomons. (Note also Roviana (ROV) manege *ten*.) There may be some relationship between these forms and the reconstructed POC *ta-(n)sa-kai *one* (Pawley n.d.), with a unit prefix, though the *-ke ∿ *-kai alternation is perhaps problematic. One might also note, within MC, KSR soko, a possible reflex of POC *(n)sa-kai, which becomes more plausible if one can assume a pre-KSR *sa-kau *one*.

3. TEN-POWER BASES OUTSIDE MICRONESIAN

Ten-power base systems of the sort described here, though perhaps most developed in MC, are attested in many other OC languages.[24] In PN, for example, though the series ends in most languages at 10^3 ('thousand'), Tongan (TON) has numerals up to 10^5. The most extensive ten-power base system thus far encountered is, in fact, that of NUK, which reaches 10^{10}. Other OC languages, especially in the Solomons (e.g. Bugotu (BUG), and Kia (KIA) of Guadalcanal), Motu (MTU), Nguna (NGU), and ROV, have ten-power bases above 10^3. Thus (SAM = Samoan):

TON	SAM	NUK	MAO	BUG	KIA	MTU	NGU	ROV	
hongofulu	sefulu	hulu/mada	ngahuru	sa-lage/ hangavulu	tazo/ (ŋa)fulu	gwauta	(dualima)	manege/ -ŋavulu	10^1
teau/ngeau	selau	lau	rau	hathangatu	gobi	sinahu	põnotia	gogoto	10^2
afe	afe	mano	mano	toga	toga	dahu	manu	tina	10^3
mano		semada		mola	vuro	gerebu	tivilia	vuro	10^4
kilu		seguli		feferi	feferi	domaga			10^5
		seloo		vuthea	tahozea				10^6
		sengaa		vathegila					10^7
		semuna							10^8
		sebugi							10^9
		sebaga							10^{10}

We will not attempt any reconstructions from these data, except to point out that most of the forms for 'ten' reflect the POC base *pulu. The bases POC *dau and *Ratu, considered in Section 2.1., are also reflected in some of the forms above. (For PPN reconstructions up to 10^3, see Biggs et al. 1970; Pawley (n.d.) gives some reconstructions for POC.)

4. EVOLUTION OF THE TEN-POWER BASE SYSTEM

Though extensive ten-power base systems are extant in a number of widely separated areas of Oceania, it is unlikely that forms above 10^3 (or possibly 10^2) can be reconstructed for POC (or for PMC). One observes, for example, that:

i) few of the higher ten-power bases (above 10^3) are reconstructable beyond the very lowest order subgroups.

ii) in many instances, what appear to be cognate items have different ten-power values in different languages.

iii) in some languages, the morphology of ten-power bases above 10^3 is distinct from that of the lower ten-power bases. Thus PP ten-power bases above 10^3 do not take the unit prefix e-, while in NUK those below 10^4 do not take the prefix se-.

These observations suggest that the higher ten-power bases have to some degree a history distinct from that of the lower. In our view, they may have developed as numbers at a more recent historical period. Similarly, the ligature *nga is found only (but not always) with the bases *pulu *ten*, *Ratu *hundred*, *dau *hundred, large number* (where these bases are reflected in extant languages), suggesting an older numeral morphology fossilised in some reflexes of the oldest ten-power bases. Though it is possible that an older extensive ten-power base system may have broken down and subsequently been reintroduced with new items, it is more likely that the systems found are the result of a number of independent innovations. That such systems developed is not in our opinion completely accidental, as we will attempt to demonstrate below.[25]

As stated above, evidence from MC (and throughout OC) suggests that the ten-power bases, even those of the simplest systems, were morphosyntactically not numbers of the same sort as the numbers 'one' through 'nine' but were countable bases (quantitative classifiers) that were themselves counted. That is, the ten-power bases can best be interpreted as 'unit of 10^x' and counted as 'one unit of 10^x', 'two units of 10^x', etc. The ten-power bases themselves are, then, countable bases distinct from other countable bases only in that their sole interpretation is a mathematical abstraction,[26] a fixed quantity, and in that they form a semantic set based on another mathematical abstraction, a ten-multiple series.

However, most of the items in the ten-power base series may not always have had a numerical interpretation, but are likely to have begun as quantitative classifiers with meanings like 'pile', 'group', 'heap', or more specific meanings implying quantity, such as 'sand' or 'leaf'. There is ample evidence in the ten-power base systems of MC and other OC languages that many of these items had such interpretations; for example, PTK *(p)piya *sand*, functioning as a higher ten-power base in a number of TK languages; KIR tetano *ground, soil*, 10^7, tetoki *end, completion*, 10^8; PNP sapw *land*, 10^8, lik *outside*, 10^9; PPN *lau *leaf*, 10^2.

The dynamic posited for the evolution of the ten-power base system is a simple one: a countable base (quantitative classifier) came to be incorporated into a mathematical series based on increasing powers of ten, the basis of which (on present evidence, as far as 10^3) was already present in the proto-language. In some instances, the base in question may have first become associated with a specific numeral value, as possibly in the case of KIR tebwiina *ten* (MOK -pwi *several*, TK pwii- *group, school, flock*), which may have been associated with a group whose conventional extent was 'ten', and later came to represent the abstraction 'ten' alone. (Note that WOL sengaul means *ten groups (of eight or ten)*, possibly its original interpretation, or possibly the converse development to that described above.)

On the other hand, we also find it plausible to conceive of individual ten-power bases being incorporated into the existing series simply by being conventionally 'tacked on at the end', as would be the case were English to develop a word for 'ten thousand' from a sequence 'ten, hundred, thousand, a heap'. The capacity for mathematical abstraction of this sort is certainly within human potential.

The sole linguistic prerequisite for the development of ten-power base systems in the manner described above is the existence of a grammatical category 'countable base (numeral classifier)'. The larger and more open this category is, the more likely it is that new ten-power bases will be added to the series through time, though it does not follow that this development need take place. It is interesting to note that it is precisely those MC languages that have largely abandoned the numeral classifier category (MRS and KSR) that have the most limited ten-power base systems. It is not clear whether more extensive systems simply never developed in these languages, or have eroded through time along with the classifier system.

In order to account in this way for the existence of ten-power base systems outside of MC we must assume that the category 'countable base (numeral classifier)' is not a MC innovation, but existed also at earlier periods in the history of other OC languages. There is some evidence that this was in fact so. Residual sets of countable bases are a feature of many PN languages (for example, TON tekau *score*, uangakau *two score;* tetula *score (of thatch)*, uangotula *two score (of thatch);* tekumi *ten fathoms*, uangokumi *twenty fathoms*).[27] Similar quantitative countable bases are attested elsewhere — Nggela na kua *ten eggs*, na banara *ten baskets of food*, na gobi *ten canoes*, na paga *ten non-human animates*.

Qualitative countable bases, on available evidence, are not widespread in OC. Outside of MC, this category has been reported on San Cristobal and Malaita (Pawley 1972), on Bougainville (Ross 1981), and in the Trobriands. The number of qualitative classifiers reported in these languages is, however, apparently considerably smaller than in MC languages.

Extra-MC evidence for the existence of a 'countable base' category is so fragmentary that it is difficult to draw more than the most speculative conclusions regarding the level for which such a category should be reconstructed, or regarding the nature of the system (for example, whether qualitative countables evolved after quantitative ones, or whether there was an extensive system of qualitative classifiers at some earlier period in the history of OC languages that has been eroded to a greater or lesser degree everywhere but in MC). Such observations as the fact that the countable bases of TON appear with reflexes of the ligature *nga suggest that the category may have had a long history.

NOTES

1. The term 'Micronesian' is used here in the same sense as the term 'nuclear Micronesian' of Bender (1971), including Nauruan (relevant data from which is, however, not available), Kiribati (KIR) (Gilbertese), and all languages of the United States Trust Territory of the Pacific Islands except Chamorro, Yapese, Palauan, Nukuoro (NUK), and Kapingamarangi. As a working hypothesis we will assume the existence of a Micronesian (MC) subgroup of Oceanic (OC), much as described in Marck (1977), though the integrity of such a subgroup has yet to be demonstrated conclusively. Within MC, we assume the existence of a Trukic (TK) subgroup (see Jackson pp. 259-280 in this volume), and a Ponapeic (PP) subgroup, including Ponapean (PNP), Ngatikese, Mokilese (MOK), and Pingelapese. We take the liberty here of making PMC reconstructions on the basis of cognates between any two first-order subgroups of MC, in full recognition of the fact that this procedure may be methodologically questionable.

2. In citing examples, the orthographies used in the standard reference works (see References) are employed. Where deemed necessary, phonetic detail is given in square brackets. (Data cited from languages for which no reference works are available are given in the transcription of the source.)

3. Cognates of the first four HAW numbers have base-ten interpretations; for example, Maori (MAO) rau *100*, mano *1000*, Fijian (FIJ) tini *10*, Motu (MTU) ge-rebu *10,000*. That 'four' is a common factor in the interpretations of the HAW forms is related to the particular status of 'four' in Hawaiian culture. HAW forms for 'hundred' and 'thousand' are Eng. borrowings, hanele and kaukani, respectively.

4. The number 'one' is historically problematic — see Section 1.1.

5. MRS numerals 'one' through 'three' appear to reflect the Proto-Micronesian (PMC) general classifier *-ua: juon *one* (see Section 1.1.), ruo *two*, jilu *three*, albeit somewhat opaquely. MRS emān *four* appears to reflect the PMC animate classifier *-manu. We can thus assume that earlier stages of MRS had a numeral classifier system.

6. Other classifiers are also reconstructable for PMC.

7. Codrington (1885) notes the existence of serial counting systems in several languages of the Solomons. Such systems are also found in Roviana (ROV) and Rotuman.

8. In MC languages, once 'ten' has been reached in serial counting, the series begins again at 'one'. The same procedure is followed once 'twenty' has been reached, etc.

9. The KIR forms are used in pair counting.

10. The forms for 'ten' will be considered in Section 2.2.

11. The long vowel of KIR teera, as well as the long vowels in other KIR, PNP, MOK, TRK, and WOL forms is the result of a regular phonological process. See Rehg (this volume).

12. MOK ehd *one* almost certainly reflects *te-sa; oahd, on the other hand, may reflect PPP *a-sa; that is, the *a- prefix found before number roots 'two' through 'nine' in PNP and MOK may have been extended to this form for 'one'. It is noteworthy that MOK ehd is used when counting out the ten-power series (*one-ten-hundred-thousand...*), while oahd is used only in unit counting (*one-two-three...*).

13. The suffix -na accompanies all unit value countable bases in KIR except those of the ten-power series. (KIR tebwiina *ten* takes -na except when a conjoined numeral follows: for example, KIR tebwii ma teuana *eleven*). MRS juon *one* is cognate with KIR teuana *one* (see Bender 1981).

14. Countable bases above 'hundred' in PP (as well as PNP ngoul *ten*) do not take the unit prefix e-, although they do take other number prefixes: PNP epwiki *one hundred*, riepwiki *two hundred*, but kid *one thousand*, riekid *two thousand*.

15. All the languages represented here, except MRS and KSR, have other bases with the value 'ten' — see Section 2.2. CRL stands for Saipan Carolinian.

16. KSR -foko *hundred* appears after the number prefixes 'two' through 'nine'.

17. MRS -rap(e)- *thousand* is archaic, now largely replaced by the Eng. loan toujin. Note the form MRS limādep *five thousand*.

18. The correspondences between KIR -kuri 10^5 and the PP and TK forms are not problematic. It should be pointed out, however, that the KIR form might be a PN loan, corresponding to NUK seguli 10^5 and TON kilu 10^5. Goodenough and Sugita (1980) give TRK e-kit 10^4. Our checking of this form with a native speaker revealed a geminate /kk/, as in PUL -kkit.

19. KSR -nguhul *ten* has the variants -ngoul and -ngaul after some number prefixes.

20. TRK -ngoon *ten* is not used for multiples of ten ('twenty', 'thirty', etc.); TRK -ik (∿ -e — see below) appears in ten-multiples with the prefixes 'two' through 'nine'.

21. PNP koadoangoul, a serial counting *ten* or *hundred*, reflects the same base, with a fossilised causative prefix (PMC *ka-) and a reflex of the unit number PMC *sa (> PNP doa).

22. WOL also fails to show a reflex of PMC *t in this form, which suggests the reconstruction PMC *p'u'(t)u' (-na); that is, it is MRS, not KIR or WOL, that is aberrant.

23. Neither ULI nor MRT shows evidence of the prefix *i-. There is also evidence (see below) that the prefix is not reflected in all TRK or PUL forms for 'ten'.

24. Extended ten-power base systems may be found in Western Austronesian languages, as well. Although we have not been able to systematically check the data for such languages, we find it very interesting that two sources have provided the following Ilokano forms: sangapulo *ten*, sangagasut *one hundred*, sangaribu *one thousand*, sangalaksa *ten thousand*, riwriw *one million*. (Constantino, 1971)

25. Girschner (1906), after describing the PNP ten-power base system, comments that 'expressions for still higher numbers have been introduced by missionaries but are not current with the natives'. If Girschner's remarks are in fact directed towards the forms in question here, it should be obvious that there

is no evidence for such an origin for the ten-power bases. We are grateful to Ken Rehg for bringing this reference to our attention.

26. It should be pointed out again, however, that some MC languages also show a 'qualitative' distinction in forms for 'ten' (see Section 2.). In addition to those cases already cited we might note MRT yengool *ten* (inanimate) and seek *ten* (animate). Similar contrasts are apparently found outside of MC, particularly in the Solomons; for example, Nggela na gobi *ten canoes,* na paga *ten puddings, ten pigs, ten birds, ten fish, etc.,* na pigu *ten coconuts, ten breadfruit, ten crabs, ten shellfish,* na banara *ten baskets of food,* na mola *ten baskets of nuts,* na gaibata *ten bunches of bananas,* as compared with e hangavulu *ten (counting)* (Codrington, 1885). Such contrasts are not found with bases above 'ten'. In MC at least, the primary meaning of the bases in question remains quantitative, in that the MC forms do not have *specific* lexical interpretations, in apparent contrast to at least some of the forms Codrington cites.

27. PPN *kumi *ten fathoms* is reflected as HAW 'umi *ten*.

BIBLIOGRAPHY

BENDER, Byron W.

1971 Micronesian languages. In Thomas A. Sebeok, ed. *Current trends in linguistics,* vol.8: *Linguistics in Oceania,* 426-465. The Hague: Mouton.

1981 A fossilized article in Marshallese. In Jim Hollyman and Andrew Pawley, eds *Studies in Pacific languages & cultures: in honour of Bruce Biggs,* 209-228. Auckland: The Linguistic Society of New Zealand.

BENTON, Richard A.

1968 Numeral and attributive classifiers in Trukese. *Oceanic Linguistics* 7/2:104-146.

BIGGS, Bruce, D.S. WALSH, and Jocelyn WAQA

1970 *Proto-Polynesian reconstructions with English to Proto-Polynesian finder list.* Interim listing January 1970. Working Papers in Linguistics, Department of Anthropology, University of Auckland.

CODRINGTON, Robert H.

1885 *The Melanesian languages.* Oxford: Clarendon Press.

CONSTANTINO, Ernesto

1971 *Ilokano dictionary.* PALI Language Texts. Honolulu: University of Hawaii Press.

ELBERT, Samuel H.

1972 *Puluwat dictionary. Pacific Linguistics,* C-24.

ELBERT, Samuel H. and Mary Kawena PUKUI

1979 *Hawaiian grammar.* Honolulu: University Press of Hawaii.

FRITZ, Georg

1911 *Die Zentralkarolinische Sprache: Grammatik, Übungen und Wörterbuch der Mundart der westlich von Truk liegenden Atolle, insbesondere der Saipan-Karoliner.* Berlin: G. Reimer.

GIRSCHNER, Regimental Physician Dr.

1906 Grammatik der Ponapesprache. *Mitteilungen der Seminar für Orientalische Sprachen: I Abteilung,* 73-126. Berlin. (Translated by S. Tully, 1975, as 'Ponapean grammar'. Mimeo.)

GOODENOUGH, Ward H. and Hiroshi SUGITA

1980 *Trukese-English dictionary.* Memoir 141 of the American Philosophical Society. Philadelphia: American Philosophical Society.

HARRISON, Sheldon P.

1976 *Mokilese reference grammar.* Honolulu: University Press of Hawaii.

in prep *A grammar of Gilbertese.*
aration

JACKSON, Frederick H.

1983 Reflexes of Proto-Oceanic in the Trukic languages of Micronesia. See pp. 259-280 in this volume.

LEE, Kee-dong

1975 *Kusaiean reference grammar.* Honolulu: University Press of Hawaii.

LYONS, John

1977 *Semantics,* vol.2. Cambridge: Cambridge University Press.

MARCK, Jeffrey C.

1977 A preliminary manual of nuclear Micronesian sound correspondences. Paper presented to the Austronesian Symposium, Linguistic Institute of the Linguistic Society of America, University of Hawaii, Summer 1977.

PAWLEY, Andrew K.

1972 On the internal relationships of Eastern Oceanic languages. In R.C. Green and M. Kelly, eds *Readings in Oceanic culture history,* vol.3:1-142, Pacific Anthropological Records 13. Honolulu: Bernice Pauahi Bishop Museum.

n.d. A Proto-Oceanic grammar. MS.

RAY, Sidney H.

1926 *A comparative study of the Melanesian languages.* Cambridge: Cambridge University Press.

REHG, Kenneth L.

1973 On the history of Ponapean phonology. ***Working Papers in Linguistics, University of Hawaii*** 5/8:17-55. See also Rehg 1983.

1979 The origins of 'compensatory lengthening' rules in Micronesian languages. (Paper presented at the Linguistic Society of Hawaii Diachronic Linguistics Festival, November 1979.) See also Rehg 1983.

1981 ***Ponapean reference grammar***. Honolulu: University Press of Hawaii.

1983a On the history of Ponapean phonology. See pp. 281-316 in this volume.

1983b The origins of 'compensatory lengthening' rules in Micronesian languages. (Paper presented at the Linguistic Society of Hawaii Diachronic Linguistics Festival, November 1979.) See pp. 52-58 in this volume.

ROSS, Malcolm

1981 Proto-Oceanic from the top down. MS. University of Papua New Guinea.

SOHN, Ho-min

1975 ***Woleaian reference grammar***. Honolulu: University Press of Hawaii.

SUGITA, Hiroshi

in prep-aration Trukese reference grammar.

TRUSSEL, Stephen

1979 ***Kiribati (Gilbertese) grammar handbook***. Peace Corps Language Handbook. Brattleboro, Vermont: The Experiment in International Living.

Harrison, S.P. and Jackson, F.H. "Higher numerals in several Micronesian languages". In Bender, B.W. editor, *Studies in Micronesian linguistics.*
C-80:61-79. Pacific Linguistics, The Australian National University, 1984. DOI:*10.15144/PL-C80.61*