THE STATUS OF PROTO-MICRONESIAN

Byron W. Bender
and
Judith W. Wang

1. INTRODUCTORY[1]

Although many of the islands of Micronesia came to the attention of the western world as early as the 16th century, most of the languages of their inhabitants were not discovered by American linguists until after World War II. Work on Micronesian languages has been going on at the University of Hawaii since the 1960s. Much of the earliest effort was directed toward producing language materials for training Peace Corps volunteers for Micronesia (e.g., McCauley 1966, Quackenbush 1966), but some preliminary comparative work was begun in seminars conducted by Byron W. Bender in the late '60s and by George W. Grace in the early '70s. The earlier work concentrated on identifying cognates and plotting their distribution, while Grace's seminars focused on comparing individual Micronesian languages with a higher order proto-language (Proto-Oceanic) reconstructed primarily on the basis of evidence outside Micronesia. Direct comparison between Micronesian languages was begun in 1976-1977 by an informal group of UH faculty and students. This task was greatly facilitated by the use of the computer to compile a unified English to Micronesian 'finderlist' or index for a number of Micronesian language dictionaries that had been processed by computer.

The task of comparison and reconstruction was continued in the summer of 1977 in a Linguistic Institute course conducted by visiting professor Ward Goodenough and in seminars conducted by Byron W. Bender and Robert W. Hsu from 1977 to 1981. Since 1981 there have been no formal meetings of the comparative Micronesian group, but various individuals have continued to work on an independent basis.

A first set of results was presented by Jeffrey C. Marck at the Austronesian Symposium of the 1977 Summer Meeting of the Linguistic Society of America. Marck focused on phonology and lexicon, presenting a preliminary compilation of sound correspondences among the Micronesian languages and a set of tentative reconstructions based on these. A few interesting irregularities were discussed by Marck, but his primary purpose was to describe the regularities that he had observed.

Andrew Pawley and Lois Carrington, eds *Austronesian linguistics at the 15th Pacific Science Congress*, 53-92. *Pacific Linguistics*, C-88, 1985.

Bender, B.W. and Wang, J.W. "The status of Proto-Micronesian". In Pawley, A. and Carrington, L. editors, *Austronesian linguistics at the 15th Pacific Science Congress.*
C-88:53-92. Pacific Linguistics, The Australian National University, 1985. DOI:*10.15144/PL-C88.53*

Much of the effort in the later seminars was devoted to entering the accumulated cognate sets into computer storage in a form that can be used with ALIGN, a computer program developed by Robert W. Hsu and James Tharp to extract sound correspondences from the data and display them according to their environment, so that the factors influencing sound change can be more easily determined. To date members of the UH comparative Micronesian project have compiled approximately 1300 cognate sets. Use of the computer as a data-storage device also makes it easier to edit and update the data.

The purpose of this paper is twofold: 1. to describe in relatively non-technical terms the way in which we have used the computer to aid in the task of comparison and reconstruction, and some of the decisions that have been forced upon us in consequence thereof, and 2. to present a preliminary overview of some of the results that we have obtained. Much remains to be done, however, ranging from simple cleanup work on the file to the exploration of both the internal and external relationships of the Micronesian languages.

2. THE DATA

2.1 The languages

The label 'Micronesian' has at least three uses, each of which refers to a somewhat different group of islands and peoples. Geographic Micronesia extends from the former Gilbert Islands (now part of the Republic of Kiribati) in the east past Belau (formerly Palau) in the west to the atoll of Tobi. In between lie Nauru and the Marshall, Caroline, and Marianas island groups.

Political Micronesia refers to the (former) U.N. Trust Territory of the Pacific Islands (TTPI): the Marshall Islands, the various districts of the Federated States of Micronesia (Kosrae, Ponape, Truk, and Yap), the Commonwealth of the Northern Marianas Islands, and the Republic of Belau. Guam, while geographically part of the Marianas island chain, has been administered separately by the U.S. since the Spanish-American War.

In this paper, however, we use the term Micronesian to refer to those languages that Bender (1971) termed 'nuclear', following Matthews (1950) — that is, the languages of geographic Micronesia excluding Chamorro (spoken in Guam and the Northern Marianas Islands), Palauan, and the Polynesian outlier languages of Nukuoro and Kapingamarangi.

Bender labelled two of these languages 'questionably nuclear' — Yapese and Nauruan. Although more is known about these languages now than in 1971, we still are not able to make a definitive statement on the relationship of either to the unquestionably nuclear Micronesian languages. Where available, Yapese and Nauruan forms have been included in the comparative file, but evidence from these languages has not been taken into consideration in the reconstruction of Proto-Micronesian (PMC).

The nuclear Micronesian languages, on which the reconstructions are based, can be divided into five major branches. Three of these consist of single languages: Kiribati (KIR; formerly Gilbertese), Marshallese (MRS), and Kosraean (KSR; formerly Kusaiean). All of these are well represented in the file. The Ponapeic (PP) subgroup includes Ponapean (PNP), Mokilese (MOK), Pingelapese (PNG), and Ngatikese (NGK). PNP and MOK are well represented, while PNG and NGK data are fewer.

The domain of the Trukic (TK) subgroup extends from Truk lagoon to Tobi. E. Quackenbush (1968) estimated there to be at least sixty distinct Trukic speech communities. However, there is the usual question of how many distinct languages are included in this dialect chain. Bender (1971) cites lexicostatistical evidence to show that while the extreme ends of the Trukic continuum are not mutually intelligible and should therefore be considered different languages, it is not so clear where intermediate language boundaries should be drawn. Bender ends up with three Trukic languages: Ulithian (including Sonsorol, Ulithi, and Woleai), Carolinian (including Saipan Carolinian as well as the central Carolinian atolls of Satawal, Pulusuk, Puluwat, Pullap, and Namonuito), and Trukese (the languages/dialects of Truk lagoon, the Mortlocks, and the Hall Islands).

In contrast, Quackenbush (1968) seems to conclude that eleven[2] languages can be identified (his Fig. 19, pp.106-107) in spite of varying degrees of mutual intelligibility between adjacent languages in the chain. For his study he selected fifteen 'dialect areas' tentatively identified on the basis of available evidence:

1. Sonsorol (representing Sonsorol, Pulo Anna, and Merir)
2. Tobi
3. Falalap, Ulithi (representing Ulithi, Fais, Ngulu, and Sorol)
5. Falalap, Woleai (representing Woleai, Eauripik, Lamotrek, Faraulep, Elato, and Ifaluk)
7. Satawal
8. Saipan (all Saipanese dialects)
9. Puluwat
10. Pulusuk
11. Pullap
12. Ulul, Namonuito
13. Murilo (Hall Islands)
14. Nama (Upper Mortlocks)
15. Moc, Satawan (Lower Mortlocks)
16. Fanapanges (Western Truk)
17. Moen (Eastern Truk)

Mogmog (Ulithi) and Ifaluk were included for 'additional perspective', (p.22) occupying positions 4 and 6, respectively, on the west-to-east list. Saipan Carolinian was given position 8 even though it clearly was the result of emigration from several central Carolinian communities. In spite of an early disclaimer that he will use only the term 'language' "to avoid endless repetition of the phrase 'language or dialect'", Quackenbush's differential treatment of these fifteen dialect areas in his conclusions suggests a position on the dialect language question: Pulusuk (10) is either omitted or hyphenated with Puluwat (9) as "virtually identical"; Saipan (8) is omitted from the idealised map of the areas (cf. Figs. 3 and 4, p.24); Upper and Lower Mortlocks, and Eastern and Western (Lagoon) Truk are each hyphenated as single 'languages' in the chain, in spite of evidence that the members of each pair differ in many features, as would be expected of dialects of the same language. In some ways it appears as though Quackenbush would have preferred to sidestep the language/dialect question altogether.

Lincoln (1981) recognises essentially the same list of eleven languages as did Quackenbush (except for Tobi and Pullap, as distinct from Sonsorol and Puluwat, respectively), while grouping them into Western, Central, and Eastern Trukic in a way that coincides with Bender's Ulithian, Carolinian, and Trukese.

Jackson (1984) assumes there to be at least seven Trukic languages: Lagoon Trukese (TRK), Ulithian (ULI), Pulo Anna (PUA), Mortlockese (MRT), Puluwatese (PUL), Satawalese (STW), and Woleaian (WOL). These plus the two major Saipan Carolinian dialects (CRL and CRN) are relatively well represented in the comparative file. Forms from other TK languages and dialects have been included when available, but no systematic search has been made for such forms.

A fairly careful search has been made for Proto-Oceanic (POC) reconstructions to which the PMC forms may be related. Less effort has been expended in searching for other non-Micronesian forms. We justify this by the fact that our primary goal has been the reconstruction of a plausible ancestral language from which the various Micronesian languages might be descended. Systematic comparisons will have to be made both with its presumed immediate ancestor (POC) and with other presumed daughters of POC in order to determine the exact genetic status and external relationships of Proto-Micronesian.

2.2 Data sources

Dictionaries exist and have been used for KIR, MRS, KSR, MOK, PNP, TRK, PUL, WOL, PUA, and YAP. The Saipan Carolinian dictionary has not yet appeared in print, but the data from the dictionary were available to the project. Other published and unpublished materials have also been consulted. For some languages — PNG, MRT, STW — forms were elicited directly from native speakers.

At this stage in the project most forms in the file have been checked by one or another member of the group. These forms generally have not been coded by source. In many cases this poses no difficulty. For some languages a dictionary or other printed matter is the only source available. When more than one source exists, however, there may be disagreement over form and/or meaning. This is true not only of printed sources, but also of native speakers. Where information about sources has been included it usually takes the form of a person's initials or some similar abbreviation included at the end of a line of data. A partial key has been included at the beginning of the file.

2.3 Representation of forms: computer vs. standard orthography

Ideally, one probably would want to represent all the data in terms of underlying phonological forms. Not all of the languages have been analysed phonologically, however. Among those that have been some use more or less phonemic standard orthographies, while others do not. In at least one case (KSR) the language has resisted several attempts at full analysis.

When the standard orthography is more or less phonemic, we have adopted it — with minor concessions to the limitations of computer processing such as changing characters with diacritics to sequences of characters, e.g., á to A'. Also, because the file originally was entered on punch cards, only upper-case letters were available. Direct communication with the computer via an on-line terminal would permit us to use both upper- and lower-case. Once having chosen to use all upper-case, however, we find it easiest to continue that practice.

In some (but not all) cases where the standard orthography differs not too greatly from the phonemic analysis, we have modified the standard spellings somewhat in the direction of the phonemic representations. Only in the case of

Marshallese, in which the standard spelling differs significantly from the phonemic representations, have we chosen to use the latter instead of the former.

Thus Trukese and Saipan Carolinian are represented in their almost-phonemic standard orthographies.[3] Woleaian, too, appears to have an almost-phonemic spelling system, which we have retained in the file.[4] Elbert's (1972) orthography for Puluwat appears to be phonemic, although he does not say so.[5] The spelling system proposed by Sohn and Bender (1973) for Ulithi also has a different symbol for each phoneme, but no spelling rules were proposed per se, and this orthography has yet to be used in a dictionary or by Ulithian speakers generally. We have, however, used this system for the ULI data in our file. The orthography used by Oda (1977:Appendix) for Pulo Anna and Sonsorolese is phonemic.[6]

Standard spelling systems do not exist for any of the other Trukic languages or dialects, but all appear to be phonologically similar enough that they can be represented within a single set of general spelling conventions.[7] The Trukese Orthography Committee specifically chose to adopt an orthographic system that would serve to represent all of the major eastern Trukic languages/dialects (Lagoon Trukese, Mortlockese, the Hall Islands, and Puluwatese), although the Goodenough and Sugita dictionary includes only the central lagoon dialect. We use a modified Trukese orthography for Trukic languages/dialects which do not yet have official spelling systems of their own. Exceptions are made, however, for forms obtained from historical sources such as Lütke.

Among the Ponapeic languages only PNP and MOK have standard orthographies. Rehg and Sohl (1979:xix-xx) indicate that the Ponapean standard orthography is phonemic except that the vowels /e/ and /ɛ/ are both represented by the letter e. These vowels are distinguished in the Northern dialect, but not in Kiti speech; however, information provided by Kenneth Rehg permits us to mark the distinction in our file even though it is not marked in the dictionary. Glides are represented by i and u, resulting in occasional ambiguity (Rehg 1981:50-51).

Harrison (1976:20) indicates that the Mokilese consonant symbols correspond to the consonantal phonemes of the language. As in Ponapean, the vowel phonemes /e/ and /ɛ/ are both represented by the letter e, but are distinguished in the dictionary by the use of different type-faces. The symbol ɛ not being available on the computer keyboard, we use the ampersand (&) to distinguish /ɛ/ from /e/. Our other major deviation from PNP-MOK standard orthography is in the use of doubled vowels (including doubled digraph oaoa) instead of digraphs with h to represent long vowels. Judging by the information available, PNG and NGK appear to be similar enough to PNP and MOK to be represented within the same overall set of spelling conventions.

Abo et al. (1976) attempted to systematise the various Marshallese spelling practices into a proposed standard orthography for that language. Bender (1968) makes it evident that this orthography is overdifferentiated with respect to the vowels and underdifferentiated with respect to the consonants. Phonemically the language has just three[8] vowels that differ only in height, whereas the orthography uses nine vowels, showing redundantly the three-way allophonic colouring distinction (front, back unrounded, and back rounded) determined by the surrounding consonants (respectively, plain, velarised, and both velarised and labialised). We employ instead a phonemic transcription based on that given following each headword in Abo et al., which shows the underlying glides — /w/, /y/, and /h/ — that are often omitted from the proposed standard orthography.

Like Marshallese, the Kosraean standard orthography appears to be both over- and underdifferentiated with respect to actual sounds present and to probable underlying forms. The letters and rules for combining them would be almost adequate for broad phonetic transcription if the spelling conventions did not eliminate the distinction among plain, velarised, and labialised consonants everywhere except before the two mid front vowels, orthographic e and ac. Kosraean has been described as having twelve surface vowels (Lee 1975)[9]:

	front	central unrounded	back rounded
high	i	ih	u
high-mid	e	uc	o
low-mid	ac	uh	oh
low	ah	a	oa

At least one of these (oa) is actually a diphthong which often loses its rounded on-glide in casual speech. Another supposed vowel (uc) appears not to contrast with the lower mid central vowel (uh). However, no one has yet succeeded in producing an adequate phonological analysis of KSR. For lack of a better means of representation, we have adopted the standard orthography for KSR — with some reservations about what it actually may represent.

For Kiribati we have used more or less the spellings given in the Sabatier-Oliva dictionary. Phonologically KIR appears to be less complex than its sister languages, and the orthography reflects this. However, the official spellings given in Sabatier-Oliva oversimplify in certain respects. The official spelling system does not distinguish between plain and velarised bilabial stops (both are spelled with b), nor does it mark vowel length, but Sabatier also gave unofficial spellings which provide this information.

The decision to use mostly standard orthographies has resulted in a certain amount of confusion over the phonetic values of various symbols. Both types of difficulties exist in our data: the same symbol may represent different sounds in different languages (e.g., d in PNP as opposed to PUA and SNS), or essentially the same sound may be represented by different symbols in different languages (e.g., KSR sr, WOL sh, and CRL sch).

Use of standard orthographies does have the advantage that it makes it easier to locate forms in dictionaries. It is not clear that this outweighs the disadvantage of having to become familiar with a multiplicity of spelling systems. Phonemicisation would help somewhat, inasmuch as it tends to reduce the number of symbols needed for a given language, but even phonemic symbols have some arbitrariness about them — witness the choice of d to represent an alveolar fricative in PUA/SNS, or the diacritic ¨ for the combination of velarisation and labialisation in MRS.

3. THE COMPUTER FORMAT: BASIC CONVENTIONS

3.1 Band format for cognate sets

The format and programs that we use are adapted from a more general format and set of programs designed for processing dictionary materials.[10] Each cognate set is treated as though it were a dictionary entry, with the reconstructed PMC form corresponding to the 'headword' of the entry.

Each entry consists of a series of lines of 'bands', each containing one form, normally a putative cognate. Each band consists of three parts: 1. a band label representing the language name (discussed in section 3.2), 2. the putative cognate, and 3. an English gloss. The first band label of an entry (the headword band) is identifiable by the fact that it always begins with a period; headwords of subentries are marked by two periods preceding the band label.

(3.1)	.MC	K	I	N	I		VI. *PINCH, PICK
	KSR	K	IH	N	–		VI. *PINCH, PICK
	TK	K	I	N	I		*PICK UP, *PLUCK (E.G., FRUIT)
	CRL	GH	I	L	I	(-)	PICK UP, GATHER
	WOL	G	I	L	I		PICK, PLUCK WITH HANDS
	PUA	K	I	N	I		VI. TO PICK UP (FRUIT)

Thus in the sample entry above the leftmost column, consisting of .MC, KSR, TK, CRL, etc., represents the band labels or language names. The middle column lists the putative cognates, headed by the reconstructed PMC form. The blank spaces which have been inserted between segments serve as boundary markers for the ALIGN and DISPLIGN programs, while the underscores and periods function as placeholders to keep the correspondences aligned (see discussion below). The rightmost column gives the English gloss for each form.[11]

The data must be entered in such a way that the computer programs will be able to distinguish how the various segments are to be aligned with the corresponding segments in the putative cognate forms. This would be no problem if each segment could be represented by a single character and if all segments in each form corresponded to segments in the cognate forms. Such is not the case, however. Our decision to use the standard orthographies for most languages requires us to make provision for diacritics and digraphs. In any case, not all portions of all forms turn out to be cognate. In reconstructed forms, too — whether our own or those of others — sometimes it is not clear which of two (or occasionally more) proto-phonemes should be reconstructed in a particular instance. There has to be a way to indicate that there are alternate reconstructions for the same segment.

Human judgment currently is required to decide how the various segments correspond among the different forms in a cognate set. In some cases it is not immediately apparent how the segments should be aligned, even when we are reasonably sure that the forms are cognate. By entering the alignment information as part of the data, rather than doing it automatically by program, we can experimentally adjust the alignments on an item-by-item basis in order to improve the overall 'fit' of the correspondence sets.

The actual mechanics of the aligned format are as follows. Segments to be aligned are entered with single blank spaces between them to serve as boundary markers for the ALIGN and DISPLIGN programs. Underscores are used to indicate zero reflexes (presumed loss of a proto-segment). Non-cognate portions of forms are enclosed in parentheses, which the computer program interprets as a signal to 'ignore' those portions in compiling correspondence sets and when displaying the aligned cognate sets. Periods or 'dots' may be used to indicate that a portion of a reconstructed form is missing. All putative cognates within the same set must have the same number of aligned segments (including underscores and dots).

(3.2)	.MC		F	A	K	A	_	A	F	I	*EVENING <*-TIM>
	KSR		_	E	K	UH	_	_	_	_	EVENING
	TK		F	A	K	A	_	_	_	_	*EVENING <*-TIM>
	MRT	(LE-)	F	A'	_	_	_	A'	F	_	EVENING
	TRK		F	A'	_	_	_	A'	F	_	EVENING MEAL, MAIN MEAL
	PUL	(LE-)	F	A'	_	_	_	_	F	_	EVENING MEAL
	STW		F	A'	_	_	_	_	FF	_	EVENING
	CRL	(LEE-)	F	A'	_	_	_	_	FF	_	EVENING
	WOL		F	E	G	A	_	A	F	I	LAST NIGHT
	OC		.	.	.	.	R.	A	P	I	(R.API) EVENING, DUSK

In the above example the underscore in the reconstructed PMC form serves as a place-holder for the *R (represented by R. in our all-upper-case computer key punch orthography) in the presumed POC antecedent *RapiRapi. The initial underscore in the Kosraean (KSR) form represents what appears to be a regular loss of PMC *f in that language. Mortlockese (MRT) and Puluwatese (PUL) le- and Carolinian (CRL) lee- are separate morphemes. CRL lee- occurs in other time words such as leesor *morning*, leealowas *noon*, etc. The use of blank spaces as boundary markers allows us to include the diacritic ' as part of the segment A' (/á/) in various Trukic forms. In these languages /á/ is a low front vowel, distinct from the low central or back vowel /a/. In Satawalese (STW) and CRL the geminate ff corresponds to the single f in other languages and has been aligned as though it were a single segment.

The *-afi part of PMC *faka-afi appears to come from an unreduplicated POC *Rapi, but PMC *faka does not correspond to anything in the presumed POC antecedent. Thus dots are used in the OC band as place-holders for the non-cognate portion of the PMC form, while the reduplicated portion of POC *RapiRapi is enclosed in parentheses to signal that it is to be ignored for alignment purposes.

Alternate reconstructions for a particular position in a proto-form would be given as a single complex 'segment' — that is, delimited by blank spaces, with commas (but no blanks) separating the alternatives. For example,

(3.3)	.MC	S,S'	A	K	E	_	TO *RIDE A VEHICLE

where there is no Kosraean cognate to distinguish between Marck's *s and *S.[12]

Metathesis poses a special problem. That is, in order to show how the segments in a presumably metathesised form correspond with segments in the putative cognates, we have had to undo the metathesis. So as not to lose the actual form, we have enclosed it in parentheses, preceded by a percent sign (%_____), and put it in the gloss portion of the band, e.g.,

(3.4)	.MC	S	U,I	K	U,I	M	A,I	*WRAP, *FOLD
	GIL	R	U	K	U	M	A	FOLD, WRAP UP
	MRS	T	I	K	I	M	_	(%KITIM) WRAP THE BODY; MAT USED TO COVER CORPSES, CASKET
	KSR	SR	O	K	O	M	I	WRAP, ENVELOP (VT)
	PNP	D	I	K	I	M	_	(%KIDIM) WRAP (VT)
	TRK	T	U'	K	U'	M	I	WRAP (VT)
	WOL	T	U'	G	U'	M	I	WRAP (VT)
	PUA	T	U'	K	U'	M	I	WRAP (VT)

For the most part the gloss portion of the band gives only enough information to enable us to identify the particular linguistic form being cited. Specially marked 'keywords' in the gloss portion of the headword band are used as input for compiling an alphabetical English to PMC index — what we call a finderlist. For example, the asterisk preceding 'evening' in the headword (.MC) band in example 3.2 will cause the computer to generate an entry.

(3.5) EVENING
FAKA_AFI <MC>

in the finderlist. The notation '<MC>' following FAKA_AFI indicates that this form has been reconstructed for PMC. We adopted this convention when we began including Proto-Trukic reconstructions not only in the data file, but also in a combined finderlist with the PMC reconstructions.

3.2 Band labels

Band labels consist of 2-4 letters: 2-letter abbreviations for reconstructed proto-languages, 3-letter abbreviations for the names of the various present-day languages, with sometimes an additional character (?, L or X) to distinguish various types of doubtful or superfluous forms which are not used in the reconstructions. (See section 4 for further discussion of the significance of the additional characters.)

Within each entry (cognate set) the languages are entered and displayed in order from east to west, with Jackson's Proto-Trukic (abbreviated as TK) reconstructions immediately preceding the Trukic forms. Nauruan and Yapese forms are listed in positions corresponding to their longitude. Forms from non-Micronesian languages, including reconstructed languages such as Proto-Oceanic, follow in no set order.

The most common bands would be represented in the following order:

(3.6)
.MC = Proto-Micronesian
GIL = Kiribati (formerly Gilbertese)[13]
MRS = Marshallese
KRS = Kosraean (formerly Kusaiean)
PNG = Pingelapese
MOK = Mokilese
PNP = Ponapean
TK = Proto-Trukic
MRT = Mortlockese
TRK = lagoon Trukese
PUL = Puluwatese
STW = Satawalese
CRN = northern (Enne; Tanapag) dialect of Saipan Carolinian
CRL = southern (Elle) dialect of Saipan Carolinian
WOL = Woleaian
ULI = Ulithi
PUA = Pulo Anna
SNS = Sonsorolese
YAP = Yapese
OC = Proto-Oceanic
EO = Proto-Eastern Oceanic
FIJ = Fijian

4. FURTHER CONVENTIONS: THE ENCODING OF SUBSIDIARY, NEGATIVE, AND MISSING INFORMATION

4.1 Inflectionally and derivationally related forms - 1

Sometimes a language will contain two or more forms that appear to be reflexes of the same proto-form. For example, many Oceanic languages have a system of 'inalienable' as opposed to 'alienable' possession marking. In the Micronesian languages inalienable possession is marked by putting a personal suffix directly onto the possessed noun. Many of the inalienable nouns also have alienable counterparts that are unsuffixed — e.g., in order to speak of someone's hand in KSR one would add the appropriate possessive suffix onto a stem po- or paho-, but to speak about the word for *hand* one would use just paho. Sometimes the free forms lack the final stem vowel of the bound forms, sometimes not. Often the vowels differ, not only between free and bound forms, but also among the different possessed forms, according to the (historical) final vowel of the possessive suffix. The synchronic analysis of these possessive forms may be far from straightforward (Rehg 1982). Most analysts would argue that syntactically they constitute two distinct subclasses of nouns. In cases such as these all of the variant (stem) forms may be entered, the one assumed to most closely correspond to the proto-form left unmarked and the other forms marked by an X attached to the language abbreviation in the band label. For example,

(4.1)

.MC	_	A	_	F	A	R	A		*SHOULDER <*-BOD>
MRS	H	A	_	Y	E	R	A	(-)	SHOULDER (COMBINING FORM FOR PERSONAL POSSESSIVE SUFFIXES)
MRSX	H	A	_	Y	E	R	A	(Y)	SHOULDER
PNP	_	A	_	P	A	R	A	(-)	SHOULDER (BASE FORM)
PNPX	_	A	_	P	&	R	&		SHOULDER (3PS)
TK	_	A	U	F	A	R	A		*SHOULDER <*-BOD>
TK	_	A	_	F	A	R	A		*SHOULDER <*-BOD>
MRT	_	A	(W)	U'	F	A	T	A (-N)	SHOULDER (3PS)
MRTX	_	A	(W)	U'	F	A	R	_	SHOULDER
TRK	_	A	_	F	A	R	A	(-N)	SHOULDER (3PS)
PUL	Y	E	_	F	A	R	A	(-N)	SHOULDER (3PS)
STWX	_	E	_	F	A	R	_		SHOULDER
STW	_	E	_	F	A	R	A	(-N)	SHOULDER (3PS)
CRL	_	A	(Y)	U'	F	A	R	A (-)	SHOULDER
CRLX	_	A	I	F	A	R	_		SHOULDER
CRL	_	A	I	F	A	R	A	(-L)	SHOULDER (3PS)
WOL	Y	A	_	F	A	R	A		SHOULDER
ULI	Y	A	_	F	A	R	A		SHOULDER
PUA	Y	A	_	D	A	L	A		SHOULDER
SNS	Y	A	_	F	A	R.	A		SHOULDER
AN	.	.	.	B	A	R.	A		SHOULDER
OC	.	.	.	P	A	R.	A		SHOULDER

Note that neither of the variant forms ayu'fara- and aifara- in CRL is marked by an X. Both will therefore appear in the sound correspondences. Often (but not always) these alternate forms are cross-referenced to each other in the gloss portion of each band, to ensure that the information that an alternate exists is carried along in the concordance data that accompany a compilation of sound correspondences.

The use of the X convention allows us a choice as to whether or not to include these forms when using the ALIGN program to sort out the correspondence sets. For example, the program can be variously instructed to ignore the X bands if we wish to exclude these 'extra' forms, to regard them as a 'different language' from their non-X counterparts if we wish to compare these with the other reflexes, or else to regard them as the 'same language' as the non-X forms if we wish all reflexes to appear together.

In the early stages of the project all forms based on the same verb root were collected together in the same cognate set, with liberal use of the X convention. More recently we chose to separate the various forms into sub-entries according to the process by which they are derived, in the hope that they might be of use in reconstructing part of the morphology of Proto-Micronesian. (See sections 4.6, 7.6.) Removing the X's from the band labels — as we have done — does have the disadvantage of inflating the number of occurrences of those correspondences that are contained within the root. The same correspondences are picked up not only for the root, but for each of the derived forms. This difficulty can be overcome by restoring the X's to all but the first occurrence of the same root in each daughter language.

4.2 Non-cognates

Another use of the X convention is to mark forms that we are reasonably sure are not cognate, but which we wish to include in the file, nonetheless. Many of these are forms which appear to be possible cognates, but which were determined upon closer examination of the correspondences not to be cognate. These have been left in the file to prevent co-workers from re-discovering them and having to re-evaluate their cognacy over and over again. Such forms are not aligned as are true cognates. Instead, the aligned portion of the band is filled by a period or 'dot' in each column, and both form and gloss are given in what would normally be the gloss portion of the band. This prevents them from entering into the sound correspondences even if we choose to include the X bands in a concordance.[14]

The same procedure has been used for forms which are clearly not cognate with the rest of a particular set, but which were thought to be of interest for other reasons — e.g., KIR kamea, MRT kamweya, ROT ko'mia *dog* (< Eng. come here ?) as opposed to general Trukic reflexes of PTK *kulaaku (< Cham. gulagu *dog* ?). Most of the clearly non-cognate forms probably should be eliminated from the file eventually.

4.3 Missing or non-cognate portions of forms

Dots have also been used as placeholders in aligned forms which have not been X-ed out, to indicate that only part of the form is cognate. For example Kosraean and Ponapean appear not to have reduplicated the initial CV- of pre-MC *kangi *sharp* as have the other Micronesian languages. This is represented in the following fashion[15]:

(4.2)	.MC		K	A	K	A	NG	I	*SHARP
	GIL		K	A	K	A	NG	–	SHARP
	MRS		K	A	K	A	G	–	SHARP
	KSR	(LAHL–)	.	.	K	UH	NG	–	SHARP
	PNP		.	.	K	&	NG	–	SHARP
	TK		K	A	K	A	NG	I	*SHARP

To the extent that these dotted portions are coextensive with synchronic or reconstructed morphemes, they reconfirm the morphemic analysis.

4.4 Doubtful cognates

There are some forms whose cognacy is questionable but which were not considered dubious enough to warrant use of the 'X plus dots' convention. These have been aligned, but are marked with a question mark, either at the end of the band label or at the beginning of the gloss.[16] Attaching the question mark to the language abbreviation in the band label permits a choice as to whether or not to include these forms when compiling correspondence sets. Putting the question mark in the gloss portion of the band does not allow such a choice.

4.5 Loanwords

A slightly different problem is the question of how to handle known loanwords. Obvious loans from non-Micronesian languages such as English, Spanish, German, or Japanese generally have been excluded. However, many loans are not at all obvious, particularly when the languages involved are closely related. Where we are reasonably certain that a word is a loan, but want to include it nevertheless, it is marked by an L attached to the band label. For instance, KIR has a rather large number of loanwords from one or more Polynesian languages. Some of these can be identified by the fact that they exhibit a different set of reflexes of certain POC/PMC phonemes. For example, the expected reflex of POC/PMC *r in KIR is Ø (loss). In a few forms, however, *r is retained as KIR r. We therefore take many (but not all) of these forms to be Polynesian loans.

In KSR some proto-phonemes appear to be reflected by two or more present-day phonemes, but not enough data exist to establish a pattern of cooccurrences among these multiple reflexes, nor is there any likely donor language in view if we should choose to regard some forms as probable loanwords. It also is possible that the multiple reflexes are a vestige of a former dialect differentiation that no longer exists in KSR. A closer examination of KSR reflexes not only of PMC, but also of POC, may help resolve this question.[17]

4.6 Inflectionally and derivationally related forms - 2

The general dictionary format also allows for the use of subentries. Although we have not been entirely consistent in doing so, it is possible to separate out derived forms that are shared by a large number of MC languages, which then can be used to reconstruct a (presumably) derived form in the proto-language. Where enough sets of forms exhibit the same derivational pattern

— that is, where the daughter languages all appear to share the same derivational process — this may suggest the existence of such a process in the proto-language.

In earlier versions of the file all forms with the same verb root were grouped together in a single cognate set with the derived forms X-ed out. All affixes and reduplicated portions were enclosed in parentheses as things to be ignored. Later on it occurred to us that we might be able to reconstruct something beyond the bare verb root if we could compare some of the material that we had been ignoring. In order to do this, these large mixed cognate sets had to be broken up and realigned in smaller sets representing the derivationally related verb classes. However, we still wished to keep such subsets together rather than scattering them throughout the file alphabetically. The subentry convention allows us to do this.

For the main entry we have chosen to use the simplest form of the verb — that is, those forms that we had been reconstructing as the verb roots. These generally are reflected as intransitive verbs in the daughter languages, often less an initial C(V)- that is the result of reduplication. The morphologically more complex forms are treated as subentries. Thus the set formerly represented as

(4.3)	.MC	K	I	N	I		*PINCH, *PICK
	GILX	K	I	N	I	(-KINI)	N. PINCH
	GIL	K	I	N	I	(-KA)	VT. PINCH
	MRS	K	I	N	_	(J-IY)	PINCH WITH FINGERNAILS
	KSR	K	IH	N	I	(S)	VT. PINCH, PICK
	KSRX	K	IH	N	_		VI. PINCH, PICK
	KSRX	K	IH	N	_	(-KIHN)	VI. PINCH, PICK
	MOK	K	I	N	I		VT. PINCH WITH FINGERNAILS
	MORX	K	I	N	I	(-KIN)	PINCH WITH FINGERNAILS
	PNP	K	I	N	I	(-I)	VT. PINCH
	PNPX	K	I	N	I	(-KIN)	PINCH
	TK	K	I	N	I	(-T"-I)	*PLUCK, *HARVEST
	TKX	K	I	N	I	(-KINI)	*PINCH OFF, *BREAK OFF (AS PIECES)

(in which the various Trukic forms are subsumed under the Proto-Trukic reconstructions for the purposes of this example) would be reorganised as a main entry

(4.4)	.MC	K	I	N	I	VI. *PINCH, *PICK
	KSR	K	IH	N	_	VI. PINCH, PICK
	TK	K	I	N	I	VI. *PICK UP, *PLUCK (E.G., FRUIT)

(plus various Trukic cognates)

with two subentries

(4.5)	..MC	K	I	N	I	K	I	N	I	VI. *PINCH, *PICK
	GIL	K	I	N	I	K	I	N	_	N. PINCH
	KSR	K	IH	N	_	K	IH	N	_	VI. PINCH, PICK
	MOK	K	I	N	I	K	I	N	_	PINCH WITH FINGERNAILS
	PNP	K	I	N	I	K	I	N	_	PINCH
	TK	K	I	N	I	K	I	N	I	ACTION OF CUTTING OR BREAKING OFF PIECES

(plus various Trukic cognates)

and

(4.6)	..MC	K	I	N	I	-T	(-I)	VT. *PINCH, *PICK
	GIL	K	I	N	I	-K	(-A)	VT. PINCH
	MRS	K	I	N	_	-J	(-IY)	PINCH WITH FINGERNAILS
	KSR	K	IH	N	I	-S		VT. PINCH, PICK
	MOK	K	I	N	I	_		VT. PINCH
	PNP	K	I	N	I	_	(-I)	VT. PINCH
	TK	K	I	N	I	-T"	(-I)	*PLUCK, *HARVEST

(plus various Trukic cognates)

Some of the evidence thus accumulated for various grammatical morphemes in PMC is discussed in section 7.6.

5. COMPUTER-GENERATED APPARATUS: LISTINGS AND SIMPLE SORTS

5.1 DISPLIGN

The DISPLIGN program is designed to produce an easily-read aligned printout of the data. Thus data that are entered in the form

(5.1)	.MC	M A M A T A	*A*WAKE
	GILX	. . . . .	MAMATA = INTUITION, SAGACIOUS
	MRS	M _ M E J _	KEEP AWAKE, STAY UP
	TK	M A M A T" A	*A*WAKE
	MRTX	M _ M A S _	VI. GUARD
	MRT	(A-) M _ M A S A' (A'TA')	VT. WAKE HIM
	PUL	(YA-) . . M A H A (A-LO')	VT. TO AWAKEN
	PULX	. . M A H A (A-TA')	VI. TO WAKE UP
	STW	M _ M A S _	AWAKE
	CRLX	M _ M A S _	AWAKE
	CRL	M _ M A S A (-TA')	WAKE UP
	WOL	M _ M A T A	AWAKE
	PUA	M _ M A T A	AWAKE

with just enough blank spaces to separate the different parts of the band, would appear in the format

(5.2)	.MC	M A M A T A	*A*WAKE
	GILX	. . . . .	MAMATA = INTUITION, SAGACIOUS
	MRS	M _ M E J _	KEEP AWAKE, STAY UP
	TK	M A M A T" A	*A*WAKE
	MRTX	M _ M A S _	VI. GUARD
	MRT	(A-) M _ M A S A' (A'TA')	WAKE HIM
	PUL	(YA-) . . M A H A (A-LO')	VT. TO AWAKEN
	PULX	. . M A H A (A-TA')	VI. TO WAKE UP
	STW	M _ M A S _	AWAKE
	CRLX	M _ M A S _	AWAKE
	CRL	M _ M A S A (-TA')	WAKE UP
	WOL	M _ M A T A	AWAKE
	PUA	M _ M A T A	AWAKE

with additional blank spaces inserted for legibility.[18]

DISPLIGN also inserts dividing lines between entries, numbers each entry, and counts up the number of aligned segments in each band. When the number of

aligned segments changes within an entry (usually an indication of misalignment and therefore probably incorrect correspondences), this is flagged by the program.

5.2 BANDSORT

If an alphabetical listing of all the forms from a particular daughter language is desired, the BANDSORT program can be used. This program sorts the bands alphabetically, first according to band labels and then by the content of the band, effectively producing a listing in which the data from each language are grouped together in alphabetical order by language name (abbreviation) and, within each language, by form. There is a provision in the program to cause it to select only those bands specified by the user. This kind of printout is useful in proofreading not only in the comparative dataset, but also in ordinary dictionary-making, to ensure correct assignment of band labels and consistency in both labeling and content of bands. An example is given as Appendix A.

5.3 Finderlists

Another program, called INVERT, takes specially marked keywords in the gloss portion of a band and creates an English index or 'finderlist' (see Appendix B). Although the program can be instructed to look in any band or set of bands for its keywords, we have found no particular need for finderlists keyed to anything but the reconstructed Proto-Micronesian and Proto-Trukic forms, since the file is organised according to the reconstructed headwords — .MC in the case of apparently pan-Micronesian cognate sets, .TK in the case of what appear to be exclusively Trukic cognate sets.

A finderlist is not quite the same as the reverse dictionary to be found in many bilingual dictionaries. For one thing, it generally does not give a definition or examples as one would expect in an ordinary dictionary entry, but functions mostly as an index to direct the user to the appropriate entry. More importantly, it may be keyed so as to group together all of the pronouns, for example, or all of the directional suffixes, in addition to listing them according to their nearest English counterparts. It also is possible to code entries so that all forms pertaining to a certain semantic or cultural domain can be listed together under some heading such as geographical terms, parts of the body, names of plants, etc.

6. COMPUTER-GENERATED APPARATUS: CONCORDANCES OF CORRESPONDENCES

As anyone knows who has ever had to compile correspondences by hand, this is a laborious, monotonous, and time-consuming chore — precisely the sort of task one wishes to relegate to a computer. It was for this purpose that the ALIGN program was conceived.

The principle according to which the ALIGN program compiles sets of sound correspondences is the same as is used to produce text concordances. That is, one must locate all occurrences of the item in question, determine what constitutes a relevant context or environment, and sort the various occurrences

of that item according to contexts. In doing ordinary text concordances this means listing all the phrases or sentences in which the desired word occurs. In doing historical-comparative phonology this means listing all the cognate sets in which the desired correspondence occurs.

The ALIGN program recognises as a 'segment' any character or sequence of characters delimited by single blank spaces on either side. Thus, long vowels and geminate consonants may be regarded as different from their short counterparts — e.g., X YY Z vs. X Y Z — or one may choose to represent the long/geminate segments as sequences of two identical short segments — X Y Y Z. This format also accommodates (other) digraphs and diacritics — when these are represented as sequences of characters (e.g., A') — as well as alternate segments (e.g., .MC A S,S' A *rub* — cf. POC *asa *grate, sharpen by grating or rubbing*). Parentheses may be used to enclose non-cognate portions of forms; such material will be 'ignored' — i.e., omitted from the correspondences.

The output of the ALIGN program consists of an index of the correspondence sets, plus a concordance or listing of the forms grouped according to correspondences. Within each group of forms the segments being concorded — that is, the correspondences being displayed — are vertically aligned, and the forms ordered alphabetically according to following segments.[19] This permits the reader to examine all occurrences of a correspondence in a given environment without having to search through the entire file for the forms. A sample concordance page is given as Appendix C, and a few excerpts are given in section 7.4.

A concordance may involve any number of languages, in any order. The number of formally different correspondences increases appreciably with the number of languages being concorded. This is partly explained by the fact that the ALIGN program treats a correspondence with a gap (one that is lacking a cognate in a particular language) as something different from a correspondence which is identical except that the gap is filled (one for which all languages have a putative cognate), and partly by the fact that the different reflexes of a proto-segment in one daughter language will not, in general, be divided among the etyma in the same way as the reflexes in another daughter language (one instance of this is discussed in section 7.5). Some of the increase in the number of correspondences is simply 'noise' — one or two occurrences of an irregular correspondence which may represent anything from a keypunch error or error in alignment to a loanword whose non-native origin is revealed thereby.

7. THE RECONSTRUCTIONS

7.1 Marck (1977)

Marck (1977) includes approximately 300 tentative reconstructions based on the following correspondences:

PMC	*p	*p	*m	*m'	*f	*t	*t'	*s	*S
GIL	p	p'	m	m'	Ø	t,Ø[2]	r	r	r
MAR	p	p'	m	m'	y	j	d	t	t
KUS	p	f	m,Ø	m	Ø	t,s[3]	ś	t	Ø
MOK	p	p'	m	m'	p[1],Ø	j,Ø[4]	s	t	t
PON	p	p'	m	m'	p[1],Ø	s,Ø[4]	t'	t	t
MUR	p	p'	m	m'	f	s,Ø[5]	r',cc	t	t
TRK	p	p'	m	m'	f	s,Ø[5]	c	t	t

PUL	p	p'	m	m'	f	h,Ø[5]	r',cc	t	t
CAR	p	b',pp'	m	m'	f	s,Ø[5]	ʂ,cc	t	t
WOL	p	b',pp'	m	m'	f	t,s[5]	ʂ,cc	t	t
UTH	p	b',pp'	m	m'	f	t,s[5]	c	θ	θ
SON	p	b',pp'	m	m'	f	t,θ[5]	s,cc	t	t

PMC	*l	*n	*r	*k	*x	*ŋ	*ñ
GIL	n	n	Ø	k,Ø[2]	Ø	ŋ	n
MRS	l,l	n,ṅ[1]	r	k	Ø	ŋ	n
KUS	l	n	l,r[8]	k	k	ŋ	Ø,l[1]
MOK	l	n	r	k	r	ŋ	Ø
PON	l	n	r	k	r	ŋ	Ø
MUR	l	n	r	k,s[6],Ø[1]	Ø	ŋ	n
TRK	n	n	r	k,s[6],Ø[1]	Ø	ŋ,n[7]	n
PUL	l	n	r	k,Ø[1]	Ø	ŋ	n
CAR	l	l	r	g,kk	Ø	ŋ	n
WOL	l,nn	l,nn	r	g,kk	Ø	ŋ	n
UTH	l	r,l	r	g,kk	Ø	ŋ	n
SON	l	r	l	g,kk	Ø	ŋ	n

1 /__ a
2 See Marck's discussion (p.17) of *t and *k deletion in Kiribati.
3 /__ (i and a)
4 /__ (i,u, and e)
5 /__ (non-low vowels)
6 /__ i (occasionally)
7 /__ i
8 less prominent than l but not rare.

Most of these correspondences are reasonably well-attested. Marck did, however, reconstruct two proto-phonemes on the basis of relatively few reflexes. *x was reconstructed solely on the basis of reflexes of *waka *canoe*, but is supported by the presence of prenasalised *-ŋk in Proto-Oceanic. The first person singular possessive suffix was subsequently reconstructed as *-xi (cf. POC *-ŋku).

Marck also reconstructed a distinction between PMC *s and *S on the basis of Kosraean, in which *S has been lost. This he justified by the apparent correlation between KSR loss and POC *ns. It now appears that the correlation is not as straightforward as Marck had thought (see section 7.5). Part of the difficulty is due to disagreement among the various Oceanic languages. Prenasalisation is problematic in Oceanic, in any case. The POC *s/*ns distinction is even more problematic because there is evidence to suggest that what has been reconstructed as *ns may, in fact, not be simply the prenasalised counterpart of *s (Ward Goodenough, p.c.).

Another problematic correspondence set is one which Marck thought might turn out to represent a distinctively Micronesian third palatal, different from that reconstructed by Blust (1976) for Proto-Fijian-Polynesian and apparently also different from Milke's POC *nj. More data are available to us now which suggest that Marck may have been on the right track (see section 7.4). Marck also found some evidence for a rounded k (*k') in Proto-Micronesian.

For the vowels Marck reconstructed the same five that others have reconstructed for Proto-Oceanic: *i, *e, *a, *o, *u — with back and central allophones of the last conditioned by the preceding sound. Preceding *p', *m', *t', *S, *k, *r, *ŋ, *w, *o, and *u allowed *u to remain back, while preceding *t, *s, *l, *n, *ñ, *i, and *e caused *u to be centralised and possibly unrounded. The labials *p, *m, and *f apparently did not occur before round vowels. Other consonants such as *s, *k', and the putative third palatal were left unclassified due to lack of relevant data. Many of the daughter languages have a high central vowel which must be regarded as a separate phoneme synchronically.

7.2 Reconstructions since Marck (1977)

Marck's reconstructions and the cognate sets on which they were based form the core of the computerised comparative Micronesian file. As the file expanded additional reconstructions by Jackson, Trussel, and Wang (among others) have been based on the correspondences compiled by Marck.

The file has grown appreciably since 1977. It now consists of approximately 1300 cognate sets with considerably more supporting data than were available to Marck. With the help of the ALIGN program it is possible to re-examine the correspondences observed by Marck and perhaps to discover some that he missed. As a result it may prove necessary to revise a number of reconstructions. A clearer picture of the ancestral language should result from this exercise.

The first step in this re-examination consists of a careful scrutiny of the correspondence sets compiled by the ALIGN program.

7.3 NEWMIC4A concordances: procedure

Following Bender (1971), we chose to concord on the five presumed major branches of the Micronesian language family: Trukic, Ponapeic, Marshallese, Kiribati, and Kosraean. Trukic is represented by Jackson's Proto-Trukic reconstructions. Ponapeic is represented by Ponapean in the absence of a reconstructed Proto-Ponapeic. The other branches consist of a single language each.

Dyen's (1965) 36.1% cognate percentage between Trukese and Ponapean suggests a close relationship between these two languages, so we expected to find greater agreement between them than among the other languages. A two-way TK-PNP concordance produced approximately 200 formally or mechanically different correspondences.

Marshallese was the next language to be included in the concordances. MRS has merged the proto-vowels to four phonemes which differ only in height, their colour (front-/backness and rounding) being determined synchronically by the surrounding consonants. The consonants, on the other hand, had multiplied themselves through splits as a result of the 'reading off' of colour from the proto-vowels. There are now three distinct series of consonants in MRS: plain but phonetically palatalised, velarised, and both velarised and labialised (see Bender 1968 for details). These historical changes are reflected by an increase in the number of different correspondences to approximately 500 in the three-way concordance.

Kiribati has undergone a number of mergers among the consonants, but it appears to have preserved an earlier vowel system, as may be seen if one compares present-day KIR forms with the forms reconstructed for Proto-Oceanic. Nonetheless, adding KIR to the concordance caused the number of separate correspondences to increase to over 1300.

We chose to put Kosraean last because of our suspicion of the existence of multiple reflexes in that language (even more than are to be found in KIR as a result of Polynesian borrowings). Adding KSR increased the number of different correspondences to over 2500. If KSR had been put first the irregularities in this language would have scattered some of the regularities among the other languages. The total number of distinct correspondences would have remained the same, but it would have been more difficult to perceive areas of agreement among the four other languages if these were split among a number of different reflexes in KSR. (One instance of this is described in section 7.5).

There is inevitably a certain amount of apparent conflict in these results. Part of this is systematic variation among alternate reflexes. Examination of the forms in the concordance data may reveal whether this is systematic (conditioned) or not. Lack of agreement in the environments in which alternate reflexes occur may be interpreted in a number of different ways — unconditioned split is one possible explanation. Dialect differences, or other variation in the daughter language are other possible explanations.

Another kind of apparent conflict is caused by missing cognates where a form has been lost or simply not discovered in a daughter language. We make the assumption that unless these incomplete correspondence sets conflict with the full correspondences in which all languages being concorded are represented, they should be considered to agree with the latter and can be incorporated into them, thus increasing the number of examples that can be referred to in attempting to discover conditioning factors for sound changes.

The real 'noise' — typographical errors, misalignments, highly dubious cognacy, etc. — usually manifests itself in the form of one or two occurrences of a relatively improbable-looking correspondence. These we have simply ignored.

To arrive at the results presented in 7.4 we began with the four-way concordance of sound correspondences involving Proto-Trukic, Ponapean, Marshallese, and Kiribati. This included some 1300 formally different correspondences.

From these we selected those consonant correspondences for which all four languages showed a cognate and which occurred at least three times.[20] Correspondences which occurred only once or twice were dismissed as probably 'noise'.

To our selected correspondences we then added those from which one or two cognates were lacking — provided that the remaining reflexes did not conflict with a previously established correspondence. In order to examine all of the environments in which these correspondences occurred, we had to re-group and re-order the relevant portions of the concordance data. A special program called REGROUP was written for this purpose.

From our examination of the environments we concluded that some correspondences could be merged, while others should remain apart for the time being. The correspondences remaining after all plausible mergers were made have been labelled with proto-phonemes. In some cases subnumbers have been assigned to different subcorrespondences within a proto-phoneme — e.g., *s, $*s_1$, $*s_2$ (discussed in section 7.4).

The next step was to add in the fifth major branch, Kosraean. The five-way concordance included approximately 2500 formally different correspondences. Again we began with those correspondences involving all five languages and occurring at least three times. In addition, some full correspondences which occurred only once or twice were included on the basis of their greater frequency in the four-way concordance. Non-conflicting correspondences with only one or two gaps were added in cases where Kosraean showed a different distribution of reflexes from the other languages. In cases where Kosraean had a single reflex in agreement with the other languages additional data were not sought out. A portion of the results from the five-way concordance is sketched out in section 7.5.

7.4 NEWMIC4A concordances: preliminary results from the four-way correspondences

We present here those consonant correspondences among the four languages (Proto-) Trukic, Ponapean, Marshallese, and Kiribati that recurred at least three times and for which each of the four languages had cognates. The quantities given in parentheses following each correspondence are for such full correspondences. Also available to us for detailed study of a given correspondence are the instances that show one or two gaps, that is, correspondences in which only three of the four languages, or two of the four, showed cognates. The proto-phonemes used to label each correspondence are generally those used by Marck (1977), although the subnumbers have been added by us. The left-to-right order of the languages is TK-PNP-MRS-KIR.

Labial Obstruents

*p	p/p/p/b	(25)
*p'	p'/pw/b/b'	(9)
	p'/pw/b/b	(10)
*f	f/p/y/Ø	(10)

The second *p' correspondence is primarily an artifact of the neutralisation of the b/b' contrast in Kiribati before u and o — the velarised variety occurs in such environments but is unmarked in the transcription. However, there are eight occurrences of this correspondence with non-round vowels for which the KIR facts need further checking.

Apical Obstruents

*t'	c/t/d/r	(21)
*s	t/d/t/r	(47)
$*s_1$	t/d/j/r	(3)
$*s_2$	t/d/y/r	(3)
*t	t"/s/j/t	(34)
$*t_1$	$t"_1$/s/j/t	(4)
$*t_2$	t"/Ø/j/t	(15)
$*t_3$	t"/./t/.[21]	(5)
$*t_4$	t"/./y/.	(3)

The *t' correspondence is relatively stable in comparison with $*^s$ and $*^t$, which proliferate into seemingly related but deviant correspondences for which conditioning factors have yet to be found, or operate sporadically, and to which we assign subnumbers for the purposes of this discussion.

*t_2 results from the tendency in Ponapean to delete s before high vowels. It has not run its course through the lexicon, as PNP suk *to pound*, sing *fart*, etc. attest, while at the same time it has begun to affect s's before mid vowels, as for example in mɛɛ(-la) *to die* < PMC *mate (cf. TK *mat"e *dead, die*, MRS me'j *dead*, KIR mate *dead, death*). In cases of this sort, which seem to need the concept of lexical diffusion (Wang 1979) to explain contrasting reflexes in a given environment, subnumbers can serve to record the progress of a tendency in each etymon.

The same tendency has progressed further in the eastern Trukic languages, so that the expectation is that s will be lost (or replaced by a glide) before all non-low vowels, as in TRK máá(-nó) *to die* (cf. WOL mase *to die*), although here again it is still to be found in some words, e.g., TRK so *alight* < PTK *t"$_1$oko (cf. PNP sok, MRS jok *alight*, KIR toka *be placed on*). As this example shows, we use the subnumber one to label those PTK *t"'s that are exceptional in having s reflexes before non-low vowels in eastern TK languages such as Lagoon Trukese and Mortlockese. This in turn accounts for the existence of PMC *t_1.

The present treatment of Trukic and Ponapeic is obviously not parallel. Eventually we would hope to reconstruct a Proto-Ponapeic and to confine to those reconstructions — at least initially — the proliferations resulting from internal Ponapeic developments. Such parallel treatment would reveal more clearly the progress of lexical diffusion in each of the two subgroups, and whether or not given etyma were affected in parallel fashion.

*s_1 and *s_2 may result from the PAN palatals having had more than one reflex in PMC, or other factors may have been involved. The three full correspondences of *s_1 (e.g., TK *p'otau *small basket*, PNP ohdow *basket*, MRS bejaw *pocket, pouch, hand basket of small weave*, KIR b'ara *small cap-like basket made of coconut leaves*) might seem little more than sporadic exceptions, were they not augmented by others having gaps, as shown in the following excerpts from the four-language concordance[22]:

(7.1) t/./j/. (5)

FARM, CLEARING	TK	M A _ A T A	
NO COGNATE	PNP		
CLEARED SPACE, OPEN FIELD	MRS	M A H A J _	
NO COGNATE	GIL		
VT. HELP	TK	T A P A -NG (-I-)	
NO COGNATE	PNP		
VT. HELP SOMEONE	MRS	J I P A -G (-EY)	
NO COGNATE	GIL		
SHINE, LIGHT	TK	T I N A _	
NO COGNATE	PNP		
START A FIRE	MRS	J E' N _ _	
NO COGNATE	GIL		
STING, SMART	TK	T O NG O	
NO COGNATE	PNP		
PAIN IN ARM...	MRS	J E G" _	
NO COGNATE	GIL		
INSIDE, CORE, INTERIOR	TK	_ U T U	
NO COGNATE	PNP		
INTERIOR OF AN ISLAND	MRS	(YA-) W A J _	
NO COGNATE	GIL		

(7.2) t/d/j/. (2)

BECOME SHALLOW	TK	P E T E P E T E
SHALLOW	PNP	P & D & P & D _
SHALLOW, SUPERFICIAL	MRS	P E' J _ P E' J _
NO COGNATE	GIL	

BECOME SHALLOW	TK	P E T E P E T E
SHALLOW	PNP	P & D & P & D _
SHALLOW, SUPERFICIAL	MRS	P E' J _ P E' J _
NO COGNATE	GIL	

MORAY EEL	TK	L A P' U T O
SALTWATER EEL (GENERIC)	PNP	L A PW & D _
MORAY EEL MRS	(KIDE'DDE'L)	_ _ B I J _
RABONO = EEL	GIL	. . . . .

(7.3) ./d/j/. (1)

NO COGNATE	TK	
TO SAVE FROM HARM	PNP	D OA R E
REANIMATE, RESTORE, REVIVE	MRS	J A R I (Q)
NO COGNATE	GIL	

(7.4) ././j/r (1)

NO COGNATE	TK	
NO COGNATE	PNP	
HUSK A COCONUT W/ TEETH	MRS	H E JE -K
LOP BRANCHES, TRIM, PRUNE	GIL	_ E R E _

We see here one of the many advantages of computerisation for tasks such as these: the ability to organise and gain ready access to large sets of complex data.[2 3]

*t_3 and *t_4 further exemplify this point. Although attested by no full correspondences, they were brought to light as deviations from the *t correspondence by holding constant the TK and MRS reflexes while permitting gaps in either PNP or KIR or both. Five instances of the former and three of the latter were thereby unearthed: e.g., TK *t"aro- *near*, MRS tir"i- *near*; TK *kut"u' *louse*, MRS (ya-)kit *delouse* (cf. MRS kij *louse* in the regular *t correspondence): TK *ku'u't"a, PNP kiis, MRS qe'ye't *octopus*; TK *t"aru *oyster*, *oyster shell*, MRS yar" *oyster*; TK *pet"i *float*, MRS pey *drift*, KIR beibeti *float*. The explanations for these MRS deviations, like those in *s_1 and *s_2, remain to be uncovered.

Although the PAN palatals generally are reflected as *s, several instances of their being reflected as *t (or one of its subcorrespondences) have appeared in addition to PMC *$t_1$iri *to spurt, urine* (< PAN *cirit *spray out*) noted by Dyen (1949) and discussed by Goodenough (1961) and Blust (1978). These include PMC *$t_1$up'e *catch* (TK *t"$_1$up'e(-li) *catch*, MRS jibe(-y) *hold, grasp, seize, capture*) < PAN *zambat *carry, hold*, and PMC *tuli *shoot* or *sucker from a root crop* (TK *t"ili *sprout, shoot from a root crop*, PNP ili *sucker of a banana, breadfruit, taro*, etc., MRS jil" *shoot, bud, sprout*) < PAN *suli *sucker*. For the latter MRS presents a doublet in the regular *t correspondence: yil *taro sprout*. PMC *$t_1$iri is attested in three of the four languages: TK *t"iri

urinate, masturbate, MRS (j-)jir *slippery, lubrication*, KIR tii *to spurt, to spout*, and has the doublet PMC *t$_4$iri *masturbate*: TK *t"iri(-i) *masturbate*, MRS yiri(-y) *wipe* (cf. KSR iri *masturbate*).

Velar Obstruents

*k	k/k/k/k	(66)
	k/k/k/Ø	(8)
*k$_1$	k/k/q/k	(12)

The second *k correspondence results from the tendency in KIR to delete k or t to prevent their cooccurrence in the same root (or more than one instance of either in a root) — seemingly a sporadic matter to which there are exceptions (Marck 1977, Trussel n.d.). *k$_1$ shows considerable complementarity with *k, the former generally appearing before round vowels, but there are sufficient counterexamples to warrant caution in merging them. For example, only KSR koet shows evidence of former vowel rounding in the word for 'octopus'(cited above) in which MRS shows q, and both *k and *k$_1$ are found before *u.

Nasals

*m	m/m/m/m	(43)
*m'	m'/mw/m'/m'	(7)
	m'/mw/m'/m	(8)
*n	n/n/n/n	(22)
	n/n/n'/n	(0)[24]
	n/n/n"/n	(1)
*ŋ	ŋ/ŋ/g/ŋ	(19)
	ŋ/ŋ/g"/ŋ	(8)

The situation for the *m' correspondences parallels that for *p', except that there are fewer counterexamples. The second and third *n correspondences result form the read-off of vowel colour onto consonants in the development of Marshallese, and can probably be merged as part of a fairly elaborate scenario that has been proposed by Kenneth Rehg (UH seminar presentation, 1979). As nearly as can be determined, the two *ŋ correspondences seem more capable of being merged than do *k and *k$_1$.

Liquids

*l	l/l/l/n	(33)
	l/l/l'/n	(19)
	l/l/l"/n	(0)[24]
*r	r/r/r/Ø	(22)
	r/r/r"/Ø	(4)
*r$_1$	r/r/r/r	(3)

The *l correspondences parallel the *n correspondences in every respect. The two *r correspondences seem capable of merger in a way paralleling *ŋ, based on the complementarity of rounding in neighbouring vowels. *r$_1$ marks the seemingly sporadic retention of r in KIR. The three full correspondences are augmented by five others with one gap each, e.g., *p'aro *box, container* (TK *p'aro *box, crate, strong container*, PNP pwoahr *hole, cave*, KIR baro *provision box, overflow, hole, depression where water stays* (cf. baron te wa *forepart of the canoe which dips up water in rough seas*).

Glides and Zero

*w	w/w/w/w	(8)
	w/w/w/Ø	(7)
	w/Ø/w/Ø	(9)
	Ø/w/w/Ø	(7)
*Ø	Ø/Ø/w/Ø	(27)
	Ø/Ø/y/Ø	(45)
	Ø/Ø/h/Ø	(26)

These correspondences shade from the *w of considerable antiquity in etyma such as PMC *wakara *root* (< POC *wakaRa) and *maSawa *ocean* (< POC *masawa), to what are clearly prothetic or epenthetic glides in Marshallese alone, which has gone farthest among the nuclear Micronesian languages in reading off vowel colour onto adjacent consonants. Much remains to be done in Austronesian generally to settle questions having to do with the status of onsets and glides. Further progress for Micronesian than what we sketch here must await resolution of the many questions regarding the vowel correspondences that still confront us.

7.5 NEWMIC4A concordances: five-way correspondences

The most significant effect of adding Kosraean to the concordance was to increase the number of different correspondences. Some of this was just 'noise', but there were also a number of correspondences that met our criteria (completeness and frequency) for further consideration. In several instances Kosraean split (or further split) proto-phonemes established on the basis of the four-way correspondences. Only one such instance will be presented here.

In section 7.4 we tentatively reconstructed PMC *s on the basis of the correspondence (TK/PNP/MRS/KIR) t/d/t/r. Two subcorrespondences (t/d/j/r and t/d/y/r) were labelled *s_1 and *s_2, respectively. Marck (1977) had recognised two correspondences (TRK/PNP/MRS/KIR/KSR) t/d/t/r/t and t/d/t/r/Ø which he labelled *s and *S, respectively. The five-way concordance yielded the following correspondences:

PMC	TK	PNP	MRS	KIR	KSR	N[25]
*s	t	d	t	r	t	19
	t	d	t	r	Ø	17
	t	d	t	r	s	12
	t	d	t	r	sr	8
	t	d	t	r	y	3
*s_1	t	d	j	r	t	3
	t	d	j	r	s	2
	t	d	j	r	sr	1
*s_2	t	d	y	r	Ø	2
	t	d	y	r	sr	1

If we ignore the different MRS reflexes and recast the distribution in terms of the KSR reflexes we still encounter certain difficulties:

*s (?)	t	d	r	t	22
*s (?)	t	d	r	Ø	19

*__?__	t	d	r	s	14
*__?__	t	d	r	sr	10
*__?__	t	d	r	y	3

Both MRS and KSR show evidence of crossover between reflexes of PMC *s and *t, but not necessarily in the same etyma. KSR sr is the expected reflex of PMC *t'. The synchronic status of y in KSR is not clear. Some y's may be phonemic, while others seem to be either prothetic or epenthetic. We have not yet arrived at an explanation for these splits.

7.6 Grammatical reconstruction

In the process of splitting up the large cognate sets containing several different forms of the same verbal root (see section 4.6) we accumulated evidence that might prove useful toward the reconstruction of the PMC grammatical system.

Transitives are formed by the addition of a transitivising suffix, usually -i (see example 4.6). These suffixes remain productive in some of the daughter languages. In other languages they are no longer productive, but their former presence has left its trace in the presence of final vowels on transitive verbs — only the suffixal vowel was lost in the historical process of final vowel deletion. So-called 'thematic consonants' have been preserved in a number of forms by the former presence of the transitive suffix — these are marked in example 4.6 by a preceding hyphen. There is evidence that final consonants were deleted, too, in Micronesian languages, if there was no vowel (which may have been deleted subsequently) to protect them.[26] The Micronesian evidence for transitivising *-i is reinforced by Pawley's (1973) reconstruction of the same form in Proto-Oceanic.[27]

All Micronesian languages have causatives. Forms such as KIR ka-, MRS ka-, PNP ka-, MOK ka-, TRK a- (reduplicated as kka-), WOL ga-, and PUA ka- point toward a PMC form *ka-. KSR is peculiar in having ahk- as the productive causative prefix, although MRS does preserve fossilised yak- (< *faka-?) and ya- (< *fa-?). Whether KSR ahk- is *faka- with the regular zero reflex of *f and sporadic loss of *a at the end of the prefix, or a metathesised reflex of *ka-, is not clear. What is unusual about the KSR causative is not only the form of the causative prefix, but also the presence of a transitivising suffix -ye on these forms (Lee 1975:187-189). Lee (pp.178-183) states that -i is used to change certain types of noun and adjectives into transitive verbs. -ye occurs only when the causative prefix is present. Its historical antecedents are not clear.

The other transitive suffix in KSR, -khin (Lee 1975:183-186), presumably reflects a reconstructed 'remote transitive' suffix *aki(ni) (Pawley 1973, Pawley and Reid 1979). This suffix is also found as KIR -akina, MRS ke'n and -Vk, PNP ki or kin (the latter when the following word begins with a vowel), PNG kin, PUL -(y)akin or -(y)ekin, in Carolinian as -ghili, -ghini, or -gini (depending on the dialect), WOL (y)agilli, ULI yixili (-xili as a suffix on some verbs), and PUA akini . MOK -ki reflects proto *-aki.[28]

Our data also permit the reconstruction of an ancestral passive suffix PMC *-aki which is reflected as, e.g., KIR -aki, MRS -ak/-e'k, KSR -yuhk, MOK -ek, CRL -ágh, and WOL -ag/-eg. Harrison (1982:202) also derives this from

POC *aki(ni). He suggests that this agentless passive suffix may be a Micronesian innovation.

Reduplication also seems to have been present in PMC. Complete reduplication of former CVCV forms is found throughout the Micronesian family. In some languages, e.g., Trukic and Ponapeic, the reduplicated forms reflect medially the final vowel of the reconstructed PMC form, suggesting either that the forms are frozen or that a final vowel is still present in synchronic underlying forms. In Kosraean, on the other hand, complete reduplications have the form CVCCVC, which suggests that such derivations must have occurred after the historical final vowels were lost from the simple forms, or that reduplication was performed on an abstract root that did not contain them.

Micronesian languages appear generally to have formed intransitive verbs by a process of initial CV- reduplication which is reflected synchronically by initial geminate consonants in the Trukic languages and Marshallese, following loss of the vowel between like consonants (Goodenough 1963). The same is true for the more sonorous consonants (e.g., m and ng) in Ponapean, but initial geminate n, l, and r have been reduced to single consonants, and the first members of geminate obstruents have undergone nasal substitution, yielding homorganic nasal-oral consonant clusters, which are preceded by a prothetic high vowel that agrees in rounding with the following segment(s). What appear to be frozen traces of a similar process of initial CV(C)- reduplication — but without syllable reduction — can be found in KSR. Forms still showing this initial CV- reduplication unaltered also can be found in KIR.

The presence of at least the residue of this process in most or all of the daughter languages points to its existence in the ancestral language. Both complete (CVCV) and initial CV(C)- reduplication appear to have existed in PMC. Kiribati uses the latter, and sometimes the former, to form distributive verbs often glossed "abounding in N" or "frequentative of V" in Bingham (1908). The most regular means of forming such verbs in Marshallese is by a combination of the two types of reduplication. So, for example, kkarjinjin *reek of kerosene* from karjin *kerosene*. Kosraean, Mokilese, and Woleaian (and probably many of the other MC languages) also allow final syllable reduplication.[29]

The existence of a number of reconstructions beginning in *ma- points toward a stative prefix *ma- which may or may not have been productive in PMC, but appears not to be productive in any of the present-day daughter languages.

A word of caution does need to be said on the subject of reconstructing grammatical morphemes in isolation from the rest of the grammatical system. Harrison (1982:181) points out that "slavish devotion" to the principle of reconstructing a unique innovation in the case of shared development in a number of related languages "can easily lead to gross errors in grammatical reconstruction". This he attributes to the practice of applying methods designed for phonological and lexical reconstruction to the reconstruction of grammatical systems. In particular, he claims that POC *aki(ni), which he suggests was a lexical verb appearing in serial construction with a preceding verb, underwent the change to a suffix 1) at different times in different branches of the Oceanic language family, and 2) at different times in different functions — these changes being interrelated with other changes in the grammatical system of each daughter language. This he feels is the only way in which one can account for the multiplicity of functions associated with present-day reflexes of this proto-form.[30]

8. THE GENETIC STATUS OF PROTO-MICRONESIAN

8.1 Introductory

It is generally recognised that the Polynesian languages constitute a very clearly defined family. There may be some disagreement over the exact number of languages (as opposed to dialects) in the family, but there is no language whose membership is at all open to dispute, and the internal subgrouping is relatively well understood. The existence of a close external relationship between Polynesian and Fijian is also quite evident.

The situation in Micronesian is not nearly so clear. Bender (1971) had to include a "questionably nuclear" category for Yapese and Nauruan. We are still uncertain about the genetic affiliation of these two languages. Internal subgrouping is very shallow. External relationships are also unclear.

8.2 Status of Micronesian as an exclusive (?) subgroup

Phonological and grammatical innovations exist which distinguish Proto-Micronesian from Proto-Oceanic, but none have yet been found that are shared by all and only the nuclear Micronesian languages. Nor have any uniquely shared lexical innovations been established to date. No exhaustive search has been made for morphological or semantic innovations.

The limited amount of work that has been done on comparative and reconstructed Micronesian grammar by Harrison (1973, 1978), Sohn (1973), and Sugita (1973) suggests a system that bears considerable resemblance to that reconstructed by Pawley (1973) for Proto-Oceanic. The so-called numeral classifiers do appear to be a Micronesian development, although possibly not uniquely shared[31] and not equally distributed among the MC languages. In particular, Marshallese has only vestiges, and Kosraean would appear to have missed out altogether in this respect. KSR does have two sets of numbers, but nothing like the elaborate classificatory systems of some of the other languages. Harrison (1976:95-97) describes four classifiers in Mokilese, while Rehg (1981: 124-137) cites some thirty classifiers in Ponapean. Kiribati appears to have the largest number of known classifiers; Trussel (1979:appendix) lists 66 numeral classifiers.

8.3 Internal relationships

Internal relationships within Micronesian, too, are less clear than one might like. The Trukic languages/dialects are obviously closely related. E. Quackenbush (1968) describes them in terms of a dialect chain or continuum. Jackson (1984) gives a list of phonological innovations shared by the Trukic languages which he claims are not shared — as a combined set, although individual innovations may be shared — by any other language or language group:

(1) Loss of POC *p before round vowels
(2) Loss of POC *ŋk in all environments
(3) Loss of POC *q in all environments
(4) Merger of POC *n with *ŋ in the environment /a __ i
(5) Merger of POC *n and *ñ elsewhere
(6) Merger of POC *s, *ns, and *j

(7) Separate reflex of (POC?) *nj
(8) Merger of POC *nt and *nd
(9) A unique pattern of loss of *R and/or merger with *d
(10) Loss of POC *y.

Rehg (1981:7-12) indicates that Ponapean and the languages of the nearby atolls — Mokil, Ngatik, and Pingelap — are all mutually intelligible, although the exact degree of intelligibility may vary. Not enough is known about Ngatikese to enable anyone to say much more than that it appears very similar to Ponapean. The 100-word list yields the following percentages of shared cognates among the other Ponapeic languages:

1. PNP - MOK 73%
2. PNP - PNG 79%
3. PNG - MOK 83%

all of which are well above the cognacy rates with the other Micronesian languages. Rehg (pp.9-11) cites other evidence for a closer relationship between PNG and MOK, suggesting the possibility that these may be considered dialects of each other, with Ponapean regarded as a distinct language.

At a slightly higher level, Dyen (1965) proposed a Trukic-Ponapeic subgroup on the basis of lexicostatistical evidence. This evidence ought to be reexamined, however, in the light of our present knowledge of these and other Micronesian languages.

Harrison (1982) cites two apparent innovations that might be used as grammatical evidence to argue for a subgroup consisting of Kosraean and the Ponapeic languages: 1. the use of reflexes of POC *aki(ni) in an instrument-flagging function, e.g.,

(8.1) KSR Nga owokihn sop ah.
I washed with soap.

MOK Ih pihnki parnijjo.
He's painting with the varnish.

(pp.204-206), and 2. the use of reflexes of POC *aki(ni) to derive denominal and deadjectival/stative transitives, e.g.,

(8.2) KSR Eltahl sengseikihn kom.
They consider you a teacher.

MOK Ngoah jamanki woallo.
I regard that man as a father.

(pp.212-213). Both of these functions of reflexes of POC *aki(ni) appear to be restricted to Ponapeic and Kosraen. However, Harrison (personal communication) does not consider this adequate basis for hypothesising the existence of a Ponapeic-Kosraean subgroup.

Until stronger evidence can be offered for higher-level subgroups, then, we continue to follow Bender's (1971) five-way subgrouping: Kiribati, Marshallese, Kosraean, Ponapeic, and Trukic.

As Marck (1977) pointed out, this apparent shallowness of internal subgrouping makes it more difficult to decide when one is justified in reconstructing a particular form as belonging to Proto-Micronesian rather than to some lower-level proto-language. It would be unreasonable to require that cognate forms occur in all five putative major branches before we attribute the reconstruction to PMC. Marck compromised by using a single star for forms

which were reflected in either Ponapeic or Trukic and at least two of the other three branches, since the probability is relatively high that these forms occurred in PMC, while other reconstructions were marked by double stars to indicate that the distribution of reflexes was defective. Marck reconstructed such forms when he deemed them important for examining the histories of particular languages.

In some cases these double star forms have cognates outside Micronesia. Marck did not distinguish these from other double star forms since he was primarily concerned with the evidence from within Micronesia. Judging from the practices of other comparativists, however, we may consider ourselves justified in reconstructing a form as PMC if we find MC cognates for a POC etymon, even if the form is not found in all major branches of Micronesian, on the grounds that this may be a retention from an earlier stage. The only other plausible explanation for such a distribution is borrowing from outside the MC family. Such borrowings are known to have occurred in languages such as KIR (< PN) and CRL/CRN (< Chamorro). In both instances the borrowings are easily identified.[32] In the other MC languages, however, there is no clear evidence of borrowing from non-MC Austronesian languages. It has been suggested that such borrowing may have occurred in KSR, but the suggestion remains unproven.

In general, we have followed Marck's lead and reconstructed Proto-Micronesian forms for cognate sets that included forms from at least three — and preferably four — branches, of which one had to be Trukic or Ponapeic. A number of PTK forms have been reconstructed by Jackson on the basis of exclusively Trukic cognate sets. In some cases, however, a primarily Trukic cognate set may include one non-Trukic witness. We have tentatively reconstructed PMC forms for these sets in the expectation that further search may turn up other non-Trukic reflexes.[33]

8.4 External relationships

It is, of course, possible to reconstruct a plausible ancestral language for a set of languages which have not been shown to constitute a valid subgroup. Blust (1984) asserts that Levy (1979) has, in fact, done so for the languages of the South-East Solomons, conglomerating together two groups of languages which Blust believes do not form an immediate subgroup. Instead, Blust presents a case for an immediate subgrouping connection between the Cristobal-Malaitan languages of the South-East Solomons and the languages of nuclear Micronesia.

Both Pawley (1972) and Levy (1979) have presented evidence for the existence of a Cristobal-Malaitan language group. Blust is willing to assume the existence of a nuclear Micronesian subgroup: nonetheless, he is careful to state that his hypothesis would not be seriously damaged if nuclear Micronesian were shown not to be an exclusive subgroup. Blust's difficulty is that he can find no phonological or grammatical innovations that are uniquely shared between Cristobal-Malaitan and nuclear Micronesian languages. Loss of POC *q and of original final consonants, cited by Blust as characteristic of both Malaitan and Micronesian languages, are shared with a number of other Oceanic languages.

Blust's evidence for a Malaitan-Micronesian subgroup falls into three categories: 1) lexical, 2) morphological, 3) semantic. All of these are subject to the danger that shared innovations may be difficult to distinguish from shared retentions. Blust has made what appears to be an exhaustive check of the available data from other Oceanic languages in order to minimise the

'avoidable error' of failing to discover existing external cognates. There is, however, no way to avoid the error that may occur if a retention should happen to be restricted to just the set of languages that constitute one's putative subgroup, or in the case of morphological and semantic innovations, if drift has occurred in the same direction only among those languages. Blust suggests, however, that while coincidences may occur on occasion, a large number of such coincidences is not likely unless these languages do, in fact, form a subgroup. Thus, he takes the number of common features — even though these cannot be shown definitively to be uniquely shared innovations — to be a kind of sub-grouping evidence, nonetheless. The numbers are relatively small, however; one wishes that more data were available from a larger number of languages.

Aside from Blust's Malaitan-Micronesian connection, attempts have also been made to link the nuclear Micronesian languages with the languages of the Admiralty Islands (Smythe 1970) and with the North Hebridean - Central Pacific languages (Pawley 1972). Both of these are rather thoroughly demolished by Blust.[34]

8.5 Speculations on Micronesian prehistory

The archaeological evidence for Micronesia is scanty, but Cordy (1982) cites a number of dates in the 400 B.C. (Truk lagoon) to A.D. 400 range for the Marshalls, Ponape, Truk lagoon, Ulithi, and Yap. No early dates are available for Kosrae. Cordy feels that sites earlier than the ones reported on should exist in Ponape, Kosrae, and the Marshalls, but just have not been discovered yet.

The available dates suggest relatively rapid dispersal and settlement, but do not give any particular indication as to direction of dispersal, whereas the linguistic evidence points to a dispersal from the east, the area of greatest linguistic diversity being in eastern Micronesia — the Marshalls, Kiribati, Nauru, Kosrae, Ponape, and Truk lagoon.

The earliest archaelogical dates available for geographic Micronesia are from the Marianas. These are considerably earlier than the dates for the eastern islands. The material — as well as linguistic — evidence points toward settlement of the Marianas by a different group of people coming from the other direction, possibly from the Philippines.

These people may have explored the islands to the east as well, but if they did settle them, it was considerably later than their settlement of the Marianas — and the settlements may not have persisted, or else may have been absorbed by later-arriving nuclear Micronesian speakers. In all probability, these early western Micronesians found the atolls inhospitable and never did get as far as the high islands at the eastern end of the Carolines.

The earliest dates cited by Cordy for the western Caroline atolls of Ulithi and Ngulu are in the same range as the eastern island sites, as are early dates from Palau (and Yap?). The Palauan language, at least, appears to be more closely affiliated with western Austronesian languages than with nuclear Micronesian. The linguistic affiliations of Yapese have not been established, although Bender (1971) is willing to admit Yapese as Oceanic and possibly even nuclear Micronesian.

It is possible that detailed examination of the sound correspondences in our present data may lead to the discovery of shared innovations that would

allow us to determine the relationships among the various branches of nuclear Micronesian. If archaeologists could get funding for excavations rather than just site surveys and salvage work, they might find more evidence bearing on settlement and origins. At present, however, the only prehistoric sequence of any sort available for anywhere in Micronesia is for the Marianas.

Of course, it may be that the Proto-Micronesians — even assuming that there was only one group of original settlers who all came from the same 'homeland' — spread so rapidly across nuclear Micronesia that there was virtually no period of common Micronesian development, and thus no uniquely shared linguistic innovations, either within nuclear Micronesian or between major branches of the putative subgroup. Unlike Polynesia, which consists largely of geographically distant island groups, the islands of nuclear Micronesia lie relatively close together. Certainly people who, in Blust's view, were capable of making the long voyage north from the South-East Solomons would have had no difficulty sailing from one Micronesian island to another.

In that case, the ancestral language of the nuclear Micronesians may not have been Proto-Micronesian, but Proto- _____ -Micronesian. Blust has suggested that the blank should be filled by 'Malaitan'. It behooves us, as proponents of the nuclear Micronesian hypothesis, to examine Blust's proposal carefully — and also to reexamine the other subgrouping hypotheses that he alludes to (Smythe 1970, Pawley 1972) — to make comparisons with the languages that have been put forward as immediate relatives of the nuclear Micronesian languages, and perhaps in that way discover what, if anything, sets nuclear Micronesian apart from the rest.

NOTES

1. Neither this report nor the project as a whole would have been possible without the effort of a number of people. The core group consisted of Byron W. Bender, Robert W. Hsu, Frederick H. Jackson, Jeffrey C. Marck, Kenneth L. Rehg, Ho-min Sohn, Stephen Trussel, and Judith W. Wang. Some of these people are no longer at the University of Hawaii; all have made significant contributions to the project. We also gratefully acknowledge the contributions of visiting colleagues from other institutions: Paul Geraghty, Ward H. Goodenough, and Sheldon P. Harrison. A number of graduate students have lent temporary assistance to the project, including Martin Combs, Layla Ebrahim, Elaine Good, Gregg Kinkley, and Meryl Siegal. Sue Archibeque did much of the initial data entry.
2. Or twelve, depending on the status assigned to Saipan Carolinian.
3. See Goodenough and Sugita (1980) and Jackson et al. (to appear) for details.
4. Sohn (1975) does not say it in so many words, but it appears that the Woleaian alphabet is essentially phonemic, with the exception of certain geminate consonants and possibly also the long vowels represented by eo and oa, and the ambiguity of the digraph iu with respect to vowel length.

Historically, eo and oa come from sequences of vowels, but they are simply long vowels in the modern language (p.18).

5. See Elbert (1974) for a description of the phonemic inventory of Puluwatese. Nowhere in either work does Elbert discuss spelling conventions.

6. See Oda (1977:9) for the phonemic inventory of PUA; Oda states in the prefatory material to the dictionary/appendix that SNS forms are generally identical to PUA except that SNS distinguishes two fricative phonemes /f/ and /d/, which have merged as /d/ in PUA.

7. H. Quackenbush (1970) describes the phonemic inventories of four Trukic dialects: Moen (Truk lagoon), Pullap, Satawal, and Sonsorol. According to Quackenbush, Moen, Pullap, and Satawal all have the same nine-vowel inventory which is also found in Saipan Carolinian. Seven vowels are given for Sonsorol — Quackenbush uses a different set of symbols and arranges them slightly differently, but basically agrees with Oda. Quackenbush's reconstructed Proto-Trukic has six vowels, the same six found in Marck's reconstructed Proto-Micronesian. Woleaian has the same six vowels in its short vowel set, but has two additional vowels in the long vowel set (see note 4).

   All of the Trukic languages/dialects have similar consonantal inventories, differing from such eastern languages as Marshallese and Kosraean in not having three different sets of consonant types — plain, velarised, and labialised. In Trukic, the only distinctively velarised consonants are w and the labiovelars pw and mw.

8. Or, from a surface point of view, four. See Bender (1968) for details.

9. The symbols used in this chart are the standard orthographic representations.

10. See Hsu and Peters (1984) for a description of the development of dictionary-processing by computer at the University of Hawaii.

11. Additional blank spaces have been inserted in this and other examples for legibility. These are not present in the data as entered into the computer, but would be inserted by the DISPLIGN program. See Section 5.1.

12. The question of *s and *S (S' in the computer orthography) in Proto-Micronesian is actually more complicated than Marck realised. See section 7.5.

13. Eventually to be replaced by KIR.

14. Actually, the dots do get picked up by the concordance program, but this does not create any additional complications in the correspondences because the program is designed to generate dots as placeholders when there is no cognate form in a given language.

15. See also POC *RapiRapi in example 3.2.

16. Both this inconsistency and the inconsistent use of X for both cognate and non-cognate forms must be attributed to the inefficiency inherent in such an informally-organised project. Many of the conventions and practices that we describe here are not the outcome of careful deliberation, but are the result of spur-of-the-moment decisions made by the person entering the data into the computer. These and a number of other aspects of the file will have to be revised before any publication of the results of this project.

17. See section 7.5 for one example of this phenomenon.

18. Actually, DISPLIGN inserts more blanks than are shown. This is due to the relative narrowness of the paper used here. DISPLIGN printouts normally appear on wide (15" x 11") paper.

19. This ordering can be changed to some other order specified by the user.

20. The vowel correspondences pose an entirely different — and quite difficult — problem, which will not be discussed here.

21. The periods ('dots') in the $*t_3$ and $*t_4$ correspondences represent gaps where cognates are lacking. The quantities given for these correspondences obviously do not represent full correspondences but occurrences of the given partial correspondence.

22. These are reproduced in slightly altered form in order to fit them onto the page. ALIGN printouts normally appear on wide (15" x 11") paper which allows space for line numbers, longer glosses, and more white space for greater legibility. See the sample concordance page given as Appendix C.

23. Although the three full $*s_2$ correspondences are not augmented by including other correspondences with gaps, one of the three is noteworthy in being a doublet of a regular *s correspondence, because of MRS alternants: PMC $*s_2$ama *outrigger* — TK *tama *outrigger float*, PNP dɛmɛ *outrigger* (3ps), MRS yam *sail with outrigger out of water*, KIR rama *outrigger* (cf. MRS (rey-)tam *the outrigger side of a canoe* in the regular *s correspondence.

24. There were no full correspondences with cognates in all four languages. These correspondences are represented only by overlapping sets of correspondences with gaps.

25. Numbers of occurrences in this table may disagree with those in 7.4. — N here is a total including non-conflicting correspondences with gaps.

26. See Lee and Wang (1984) for examples of such developments in Kosraean.

27. See Harrison (1978) for a fuller discussion of transitivity in Micronesian languages.

28. See Harrison (1982) for a different interpretation of POC *aki(ni).

29. See Harrison (1973) for further discussion of reduplication in Micronesian.

30. See examples in part 1.2, pp.179-180.

31. This has not yet been established.

32. As are such intra-Micronesian borrowings as MOK jimwoa *chicken* < KIR te moa, and various Marshallese terms associated with coconut toddy (Bender 1981).

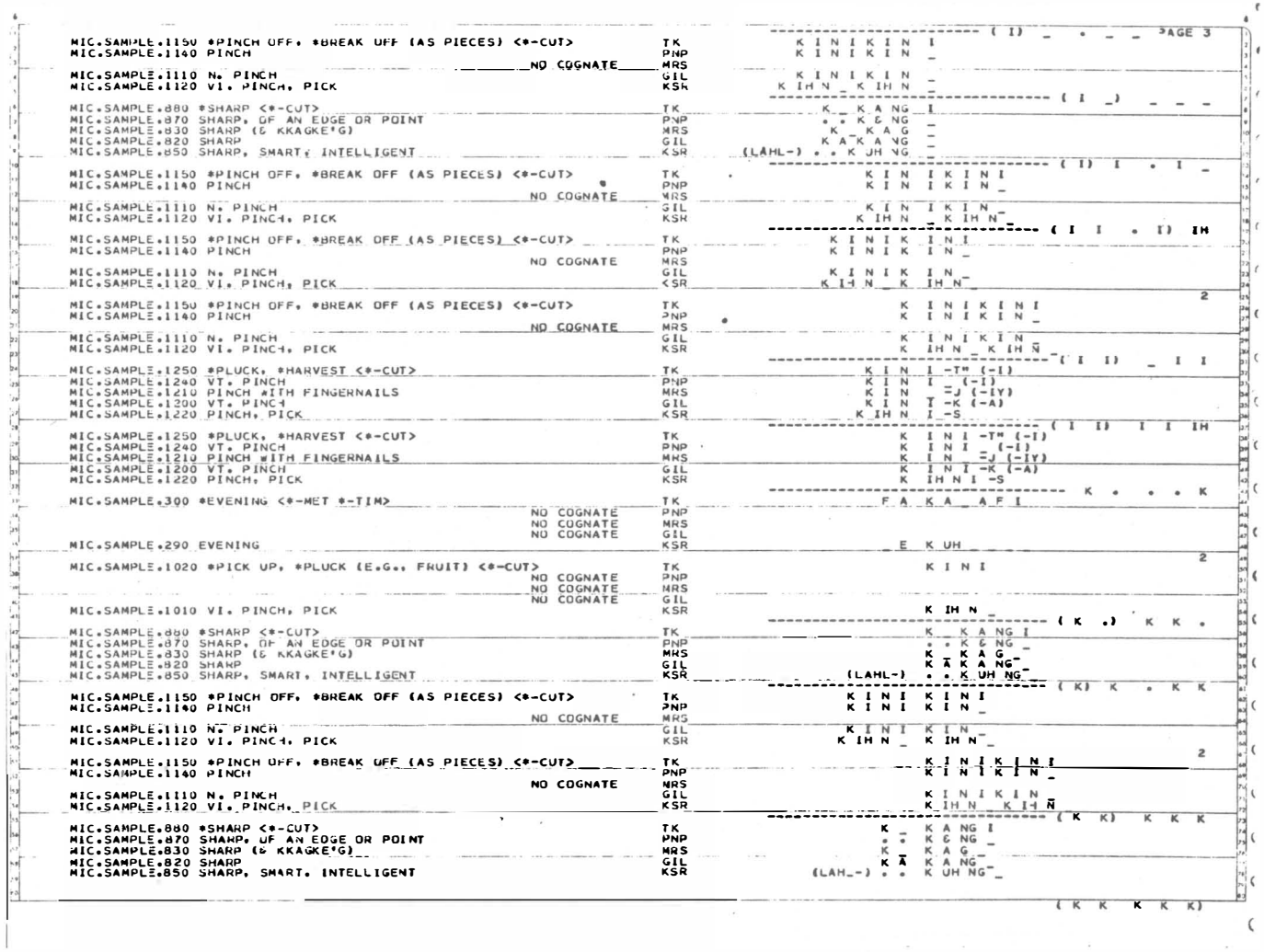
33. Because of the dispersed nature of the Micronesianist group in the past few years, the most recent additions to the file have not been as thoroughly checked — either for the existence of cognates or for accuracy of data entry — as were earlier data. It is probable that many of the gaps in the data will be filled when all languages have been carefully checked.

34. In fairness to Pawley, it ought to be said that he has since revised his Eastern Oceanic subgroup a couple of times (1977, 1979), so that it no longer includes either the South-East Solomonic or the nuclear Micronesian languages.

```
F A K A _ A F I              46   AD    P A . . R. A P I    EVENING
F A K A _ A F I              43   AN    . . . . R. A B I    EVENING
_ A _ F A R A                25   AN    . . . B A R. A      SHOULDER
F A _ O _ U                  66   AN    B A Q E R. U (Q)    NEW
F A _ O _ U                  60   CRL   (-) F E' _ _ _ _    NEW
F A K A _ A F I              36   CRL   (LEE-) F A' _ _ _ _ FF _      EVENING
F A K A _ A F I              37   CRL   (LEE-) F A' _ _ _ A' F _      IN THE EVENING
_ A _ F A R A                18   CRL   _ A (Y) U' F A R A (-)    SHOULDER (& AIFARA)
_ A _ F A R A                20   CRL   _ A I F A R A (-L)   HIS SHOULDER (& AFAR)
K I N I                      96   CRL   GH I L I (-)    PICK UP, GATHER
K I N I > K I N I -T (-I)   122   CRL   GH I L I _ (-I)    VT. TO PICK UP, HARVEST
K A K A NG I                 86   CRL   K _ K A' NG _    SHARP
M A M A T A                 143   CRL   M _ M A S A (-TA')    WAKE UP
_ A _ F A R A                19   CRLX  _ A I F A R _        SHOULDER (& AFARAL)
M A M A T A                 142   CRLX  M _ M A S _    AWAKE
F A K A _ A F I              47   EO    . . . . R. A V I    EVENING
F A _ O _ U                  70   EO    V A Q A R. U    NEW
M A M A T A                 148   FIJ   . . M A T A    IO
F A K A _ A F I              44   FIJ   (MATA-) _ A K _ _ A V I    EVENING
F A _ O _ U                  72   FIJ   V O _ _ _ U    NEW
F A K A _ A F I              49   FIJ   Y A K _ _ A V I    EVENING
K A K A NG I                 75   GIL   K A K A NG _    SHARP
K I N I > K I N I -T (-I)   113   GIL   K I N I -K (-A)    VT. PINCH
K I N I > K I N I K I N I   104   GIL   K I N I K I N _    N. PINCH
F A _ O _ U                  52   GILL  B O _ O _ U    NEW (<TUV FOOU)
M A M A T A                 132   GILX  . . . . . . .    (MAMATA = INTUITION, SAGACIOUS)
K A K A NG I                 78   KSR   (LAH-) . . K UH NG _    SHARP, SMART, INTELLIGENT
F A K A _ A F I              29   KSR   _ E K UH _ _ _ _ EVENING
K I N I                      94   KSR   K IH N _    VI. PINCH, PICK
K I N I > K I N I K I N I   105   KSR   K IH N _ K IH N _    VI. PINCH, PICK
K I N I > K I N I -T (-I)   115   KSR   K IH N I -S    PINCH, PICK
_ A _ F A R A                 2   .MC   _ A _ F A R A    *SHOULDER <*-BOD>
F A _ O _ U                  51   .MC   F A _ O _ U    *NEW (QB 122)
F A K A _ A F I              28   .MC   F A K A _ A F I _ *EVENING <*-MET *-TIM>
K A K A NG I                 74   .MC   K A K A NG I    *SHARP <*-CUT>
K I N I                      93   .MC   K I N I    VI. *PINCH, *PICK <*-CUT>
K I N I > K I N I -T (-I)   112   ..MC  K I N I -T (-I)    VT. *PINCH, *PICK <*-CUT>
K I N I > K I N I K I N I   103   ..MC  K I N I K I N I    VI. *PINCH, *PICK <*-CUT>
M A M A T A                 131   .MC   M A M A T A    *A*WAKE <*-PHY>
K A K A NG I                 79   MOK   (I) NG _ K OA NG _    SHARP
_ A _ F A R A                 6   MOK   _ A _ P _ R OA    SHOULDER (3PS)
K I N I > K I N I -T (-I)   116   MOK   K I N I _    VT. PINCH WITH FINGERNAILS
K I N I > K I N I K I N I   106   MOK   K I N I K I N _    PINCH WITH FINGERNAILS
_ A _ F A R A                 4   MRS   H A _ Y E R A (-)    SHOULDER (COMBINING FORM FOR PER. POSS. SUFFIXES) (& HAYERAY)
K A K A NG I                 76   MRS   K _ K A G _    SHARP (& KKAGKE'G)
K I N I > K I N I -T (-I)   114   MRS   K I N _ -J (-IY)    PINCH WITH FINGERNAILS
M A M A T A                 133   MRS   M _ M E J _    KEEP AWAKE, STAY UP
_ A _ F A R A                 3   MRSX  H A _ Y E R A (Y)    SHOULDER
K A K A NG I                 77   MRSX  K _ K A G _ (-KE'G)    SHARP (POSTPOSITIONAL FORM) (& KKAG)
M A M A T A                 136   MRT   (A-) M _ M A S A' (A'TA')    VT. WAKE HIM
F A K A _ A F I              31   MRT   (LE-) F A' _ _ _ A' F _    EVENING
F A K A _ A F I              32   MRT   (LEE-) F A' _ _ _ A' F _ _ IN THE EVENING
_ A _ F A R A                11   MRT   _ A (W)U' F A R A (-N)    SHOULDER (3PS)
F A _ O _ U                  55   MRT   F E' _ _ _ _    NEW
K A K A NG I                 82   MRT   K _ K A NG _    SHARP
K I N I > K I N I -T (-I)   119   MRT   K I N I _ (-Y)    BREAK PIECES OFF, CUT
K I N I > K I N I K I N I   109   MRT   K I N I K I N _    ACTION OF CUTTING OR BREAKING OFF PIECES
F A _ O _ U                  54   MRTX  (SE'E') F E'    NEW
_ A _ F A R A                12   MRTX  _ A (W)U' F A R _    SHOULDER
M A M A T A                 135   MRTX  M _ M A S _    VI. GUARD
```

Appendix B: finderlist sample

```
FINDERLIST FOR MIC.SAMPLE                                  00000420          COMPILED 01/13/83 I

      1      AWAKE
      2                M_MATWA <TK>::   .
      3                MĀMATA <MC>::   .

      4      BREAK
      5                KINIKINI <TK>::  PINCH OFF, BREAK OFF (AS PIECES).

      6      EVENING
      7                FAKA_AFI <MC>::   .
      8                FAKA_AFI <TK>::   .

      9      HARVEST
     10                KINI-T*(-I) <TK>::  PLUCK, HARVEST.

     11      NEW
     12                FA_O_U <MC>::   NEW (QB 122).
     13                FA_O_U <TK>::   NEW (QB 122).

     14      PICK
     15                KINI <MC>::  VI. PINCH, PICK.
     16                KINI <TK>::  PICK UP, PLUCK (E.G., FRUIT).
     17                KINI-T(-I) <MC>::  VT. PINCH, PICK.
     18                KINIKINI <MC>::  VI. PINCH, PICK.

     19      PINCH
     20                KINI <MC>::  VI. PINCH, PICK.
     21                KINI-T(-I) <MC>::  VT. PINCH, PICK.
     22                KINIKINI <MC>::  VI. PINCH, PICK.
     23                KINIKINI <TK>::  PINCH OFF, BREAK OFF (AS PIECES).

     24      PLUCK
     25                KINI <TK>::  PICK UP, PLUCK (E.G., FRUIT).
     26                KINI-T*(-I) <TK>::  PLUCK, HARVEST.

     27      SHARP
     28                K_KANGI <TK>::   .
     29                KĀKANGI <MC>::   .

     30      SHOULDER
     31                _A_FAHA <MC>::   .
     32                _A_FARA <TK>::   SHOULDER (& AUFARA).
     33                _AUFARA <TK>::   SHOULDER (& AFARA).

     34      WAKE
     35                MAMATA <MC>::  AWAKE.

END OF LISTING.
35 LINES. 11 MAIN ENTRIES. 0 SUBENTRIES.
24 RECORDS IN SORTED FILE.  11 DIFFERENT KEYWORDS.
LENGTH OF LONGEST PARAGRAPH SET: 54
```



BIBLIOGRAPHY

ABO, Takaji, Byron W. BENDER, Alfred CAPELLE, and Tony DEBRUM

1976 *Marshallese - English dictionary*. Honolulu: University Press of Hawaii.

BENDER, Byron W.

1968 Marshallese phonology. *Oceanic Linguistics* 7/1:16-35.

1971 Micronesian languages. In Thomas A. Sebeok, ed. *Current trends in linguistics,* vol.8, Part 1: *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 and cultures: in honour of Bruce Biggs,* 209-228. Auckland: The Linguistic Society of New Zealand.

BENDER Byron W., ed.

1984 *Studies in Micronesian linguistics*. *PL*, C-80.

BINGHAM, Hiram

1908 *A Gilbertese-English dictionary*. Boston: American Board of Commissioners for Foreign Missions.

BLUST, Robert

1976 A third palatal reflex in Polynesian languages. *Journal of the Polynesian Society* 85:339-358.

1978 *The Proto-Oceanic palatals*. Wellington: The Polynesian Society.

1984 Malaita - Micronesian: An eastern Oceanic subgroup? *Journal of the Polynesian Society* 93/2:99-140.

CORDY, Ross

1982 A summary of archaeological work in Micronesia since 1977. *Bulletin of the Indo-Pacific Prehistory Association* 3:118-128.

DYEN, Isidore

1949 On the history of the Trukese vowels. *Language* 25:420-436.

1965 *A lexicostatistical classification of the Austronesian languages*. *IJAL* Memoir No. 19. Indiana University Publications in Anthropology and Linguistics.

ELBERT, Samuel H.

1972 *Puluwat dictionary*. *PL*, C-24.

1974 *Puluwat grammar*. *PL*, B-29.

GOODENOUGH, Ward H.

1961 Migrations implied by relationships of New Britain dialects to Central Pacific languages. *Journal of the Polynesian Society* 70:112-136.

GOODENOUGH, Ward H.

1963 The long or double consonants of Trukese. In the *Proceedings of the Ninth Pacific Science Congress*, 1957, vol.3:77-86.

GOODENOUGH, Ward H. and Hiroshi SUGITA

1980 *Trukese-English dictionary*. Philadelphia: American Philosophical Society Memoirs 141.

GRACE, George W., Janet BLACK, Peter C. LINCOLN, and James THARP

n.d. Proto-Oceanic lexical reconstructions. Computer printout (1978). Honolulu: Department of Linguistics, University of Hawaii.

HARRISON, Sheldon P.

1973 Reduplication in Micronesian languages. *Oceanic Linguistics* 12/1-2: 407-454.

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

1978 Transitive marking in Micronesian languages. In S.A. Wurm and Lois Carrington, eds *Second International Conference on Austronesian Linguistics: proceedings*, 1067-1127. *PL*, C-61.

1982 Proto-Oceanic *aki(ni) and the Proto-Oceanic periphrastic causative. In Amran Halim, Lois Carrington, and S.A. Wurm, eds *Papers from the Third International Conference on Austronesian Linguistics*, vol.1: *Currents in Oceanic*, 178-230. *PL*, C-74.

HARRISON, Sheldon P., and Salich ALBERT

1977 *Mokilese-English dictionary*. Honolulu: University Press of Hawaii.

HSU, Robert W. and Ann M. PETERS

1984 Computers and Micronesian dictionaries: a chronicle of systems-design fieldwork among the lexicographers. In Bender, ed. 1984:1-36.

JACKSON, Frederick H.

1984 Reflexes of Proto-Oceanic in the Trukic languages of Micronesia. In Bender, ed. 1984:259-280.

JACKSON, Frederick et al.

f/c *Saipan Carolinian dictionary*.

JENSEN, John Thayer

1977 *Yapese-English dictionary*. Honolulu: University Press of Hawaii.

LEE, Kee-dong

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

1976 *Kusaiean-English dictionary*. Honolulu: University Press of Hawaii.

LEE, Kee-dong and Judith W. WANG

1984 Kosraean reflexes of Proto-Oceanic phonemes. In Bender, ed. 1984: 403-442.

LEVY, Richard

1979 Languages of the Southeast Solomon Islands and the reconstruction of Proto-Eastern Oceanic. In Naylor, ed. 1979:213-226.

LINCOLN, Peter C.

1981 Micronesia. In Stephen A. Wurm and Shirô Hattori, eds *Language atlas of the Pacific area*, Part I: *New Guinea area, Oceania, Australia*. Canberra: The Australian Academy of the Humanities, in collaboration with the Japan Academy; *PL*, C-66.

LÜTKE, Frederic

1835-1836 *Voyage autour du monde, exécuté par ordre de Sa Majesté l'empereur Nicolas Ier sur la corvette Le Seniavine, dans les années 1826, 1828 et 1829*. 3 vols. Paris: Atlas.

MARCK, Jeffrey

1977 Preliminary nuclear Micronesian sound correspondences. Paper presented to the Linguistic Society of America Summer Meeting, Honolulu.

MATTHEWS, William K.

1950 Characteristics of Micronesian. *Lingua* 2:419-437.

McCAULEY, David P.

1966 *Lessons in Kusaien*. Hoolehua, Molokai: UH Peace Corps Training Center.

NAYLOR, Paz B., ed.

1979 *Austronesian studies: papers from the Second Eastern Conference on Austronesian Languages*. Ann Arbor: University of Michigan Center for South and Southeast Asian Studies.

ODA, Sachiko

1977 The syntax of Pulo Annian: a nuclear Micronesian language. Ph.D. dissertation, University of Hawaii.

PAWLEY, Andrew

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

1973 Some problems in Proto-Oceanic grammar. *Oceanic Linguistics* 12/1-2: 103-188.

1977 On redefining 'Eastern Oceanic'. MS.

1979 *New evidence on the position of Rotuman*. University of Auckland Anthropology Department Working Papers. No. 56.

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

PAWLEY, Andrew and Lawrence A. REID

1979 The evolution of transitive constructions in Austronesian. In Naylor, ed. 1979:103-130.

QUACKENBUSH, Edward M.

1968 From Sonsorol to Truk: a dialect chain. Ph.D. dissertation, University of Michigan.

QUACKENBUSH, Hiroko C.

1966 *Lessons in Pullapese*. Honolulu: University of Hawaii Peace Corps Training Center.

1970 Studies in the phonology of some Trukic dialects. Ph.D. dissertation, University of Michigan.

REHG, Kenneth L.

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

1982 Ponapean possessive paradigms: the smart solution, the dumb solution, and the Ponapean solution. Paper presented to the University of Hawaii Austronesian Circle, Honolulu, April 22, 1982.

REHG, Kenneth L., and Damian G. SOHL

1979 *Ponapean-English dictionary*. Honolulu: University Press of Hawaii.

SABATIER, Fr. E., MSC, translated by Sister Oliva, FNDSC.

1971 *Gilbertese-English dictionary*. Noumea: South Pacific Commission Publications Bureau.

SMYTHE, W. E. (ed. by A. HEALEY)

1970 Melanesian, Micronesian, and Indonesian features in languages of the Admiralty Islands. In S.A. Wurm and D.C. Laycock, eds *Pacific linguistic studies in honour of Arthur Capell*, 1209-1234. *PL*, C-13.

SOHN, Ho-min

1973 Relative clause formation in Micronesian languages. *Oceanic Linguistics* 12/1-2:353-392.

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

SOHN, Ho-min and Byron W. BENDER

1973 *A Ulithian grammar*. *PL*, C-27.

SOHN Ho-min and Anthony F. TAWERILMANG

1976 *Woleaian-English dictionary*. Honolulu: University Press of Hawaii.

SUGITA, Hiroshi

1973 Comparison of verb-object relationships in Micronesian languages. *Oceanic Linguistics* 12/1-2:393-406.

TRUSSEL, Stephen

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

n.d. PMC *t and *k in Gilbertese. MS.

WANG, William S-Y.

1979 Language change — a lexical perspective. *Annual Review of Anthropology* 8:353-371.

Bender, B.W. and Wang, J.W. "The status of Proto-Micronesian". In Pawley, A. and Carrington, L. editors, *Austronesian linguistics at the 15th Pacific Science Congress.*
C-88:53-92. Pacific Linguistics, The Australian National University, 1985. DOI:*10.15144/PL-C88.53*