

THE HISTORY OF THING, ANIMAL, PERSON AND RELATED CONCEPTS IN MALAY

K. ALEXANDER ADELAAR

1. INTRODUCTION

In this paper I investigate the origin of the Malay words *barang* 'thing', *binatang* 'animal' and *orang* 'person'.¹ I also propose etymologies for some other historically and semantically related words viz. *bər/apa* 'how much/many', *ayam* 'chicken' and *main* 'to play'.

Barang derives from Proto Malayic (PM) **baraʔ*, a 'marker of uncertainty and indefiniteness of object or number' + **ŋ* (a ligature, see below). *Bər/apa* derives from PM **baraʔ* + **apa* 'what?'. I reconstruct Proto Malayo-Polynesian (PMP) **baRa*, a 'marker of uncertainty and indefiniteness of object or number' in place of Dempwolff's (1938:23) PMP **barang* 'goods' and Blust's (1980a:48) Proto Western Malayo-Polynesian (PWMP)² doublets **barang/barəŋ* 'marker of uncertainty, conditionality or hope'.

Binatang must be a relatively recent innovation in Malay. It is phonotactically highly irregular, and it cannot be reconstructed for either PMP or PM. But, on the other hand, it is possible to reconstruct PM **hayam* 'domestic animal (including pig, dog, fowl); plaything'. A pre-PM **q-um-ayam* 'to play' developed into the phonemically reduced and lexicalised Malay *main*.

Dempwolff's (1938:160) PMP **uRaŋ* 'human being' and Blust's (1970:125) PMP **(CrT)uRaŋ* 'in-law' must be related forms. I reanalyse them as PMP **uRaŋ* with a referential derivation **t-uRaŋ*, and I tentatively gloss them 'outsider; affine; friend'.

Finally, my analysis of *barang* sheds some new light on the history of the ligature *ŋ* which among other languages occurs in Philippine languages and in Old Javanese. It was a linker between the parts of a noun phrase, and also between digits and higher order numerals in numeral compounds. The modern West Indonesian languages still reflect **ŋ* in numeral compounds, but they have lost it as a ligature in noun phrases. **ŋ* was lost in Malay, although it must still have been present (in numeral compounds) in PM.

¹ I am grateful to Jim Fox, Chuck Grimes, Laurie Reid and David Zorc for their helpful suggestions and critical reading of earlier versions of this paper. Any errors and omissions left in it are due to my own inadvertency.

² In Blust's classification of Austronesian languages, PWMP is a primary branch of Malayo-Polynesian and it contains all Malayo-Polynesian languages of the Philippines, Malaysia and West Indonesia, the Chamic languages, Malagasy, and (in Micronesia) Chamorro and Palauan. West Indonesia includes here Bali, Lombok and the western half of Sumbawa and Sulawesi together with the Banggai archipelago in the east and Muna-Buton and the Tukangbesi archipelago in the south-east (Blust 1980a:12). PWMP is, however, not defined by strong linguistic criteria, and it is more of a residue category for Malayo-Polynesian languages that do not belong to Blust's Central or Eastern branches.

1.1 *Baraŋ* 'THING'

The Malay word *baraŋ* has a wide variety of meanings. Wilkinson (1959) considers these meanings sufficiently contiguous to group them under the same dictionary entry, and he describes *baraŋ* as:

thing; stuff; wares; goods; something; somehow; more or less; may it happen in some way. In various senses:

(i) *baraŋ Jərman* goods made in Germany

(ii) *B[araŋ] apa* whatever. *B[araŋ] bila*; *B[araŋ]-kala* whenever. *B[araŋ] siapa* whoever. *B[araŋ] suatu* whichever. *B[araŋ] kə-mana* wherever, anywhere

(iii) *Baraŋ-baraŋ* things of all sorts; impedimenta; the usual things. *Bukan b[araŋ]-b[araŋ] oraŋ* not a man of the usual sort; no ordinary man;...

(iv) *Baraŋ dipəliharaŋkan Allah daripada səgala marabahaya* may God protect her from all dangers;...

In other dictionaries the different meanings of *baraŋ* are sometimes allotted to homonyms, cf. Echols and Shadily (1989), which distinguishes three homonyms with basically the following meanings:

*baraŋ*¹ 1. goods, commodity; 2. article, object; 3. (Coll.) s.t. whose direct

mention should be avoided (i.e. marijuana, genitals, etc.); 4. baggage, luggage

*baraŋ*² any; [*baraŋ*] *apa* anything, whatever; [*baraŋ*] *bila* whenever; [*baraŋ*] *di mana* wherever, anywhere

*baraŋ*³ more or less, approximately

It furthermore gives the following derivations:

baraŋ-baraŋ may it happen that, would that

baraŋkali perhaps, maybe

For practical purposes the second arrangement seems more appropriate, but from a semantic viewpoint it is not hard to conceive that the several homonyms distinguished in Echols and Shadily derive from a single etymon.

I would like to show that (1) *baraŋ* with its different meanings derives from a PMP etymon **baRa* (denoting uncertainty and indefiniteness of object or number) + a fossilised ligature **-ŋ*, and that (2) *baraŋ* is historically related to the Malay quantifier *bəŋ/apa* 'how much/many?',³ a lexeme which derives from **baRa* + **apa* 'what? (interrogative pronoun)'.

The analysis of (Old Javanese and Malay) *bara/ŋ* as consisting of **bara* + a ligature **-ŋ* was first proposed by Kern (1918:172-173). On the basis of *baraŋ* and corresponding forms other reconstructions have been made, which I am listing below.

The proposed relationship of *baraŋ* to *bəraŋpa* through PMP **baRa* involves some problems which are dealt with in sections 1.2 to 1.5. For instance, is there other evidence for a ligature **-ŋ* in the history of Malay which would support my explanation of the final nasal in *baraŋ*? Is there additional evidence for a PMP **baRa* denoting uncertainty and/or question? What are the arguments against earlier etymologies proposed for *baraŋ* and *bəraŋpa*?

³ Wilkinson defines *bəraŋpa* as: 'in some quantity; to some extent; how much (if pronounced interrogatively)'.

1.2 EVIDENCE FOR AN ENCLITIC VELAR NASAL

1.2.1 EVIDENCE FOR AN ENCLITIC VELAR NASAL IN MALAY

None of the known varieties of Malay has a separate morpheme *-ŋ*, and this observation includes Classical Malay and the Old Malay of the seventh century inscriptions of South Sumatra. However, it seems that some forms of Malay exhibit a fossilised *-ŋ*. Standard Malay has a relative pronoun *yaŋ* which is generally assumed to have developed from the third person pronoun *ia* and a clitic element *-ŋ*. This *-ŋ* would be a cognate of the ‘linker’ or ‘ligature’ found in languages such as Old Javanese and Tagalog (see below). *Yaŋ* is not likely to be borrowed, as languages that had an influence on Malay do not have such a relative pronoun.

The velar nasal ligature *-ŋ* must be distinguished from a genitive linker. The latter derives from the Proto Austronesian (PAN) genitive marker **ni*, but in the Austronesian daughter languages it may be reduced to an *-N-* which becomes homorganic with the following consonant (Blust 1974b:3ff.). This reduction also occurred in Malay and is witnessed by the third person possessive suffix *-ŋa*, which can be analysed as **ni + *ia* ‘(third person pronoun)’. In Old Malay it is also still shown in *-ŋku*, the allomorph of the first person singular possessive suffix *-ku* when it follows a noun head ending in a vowel, for example:

çatru-ŋku ‘my enemy’ (De Casparis 1956:33 [line5])

huluntuhā-ŋku ‘my senior officials’⁴

cūrikā-ŋku ‘my knife’ (De Casparis 1956:5)

-ŋku can be analysed as **ni + *ku*, and it also occurs still in Brunei Malay (Nothofer 1991:153). In modern Malay this first person singular suffix lost the nasal and became *-ku* throughout.

Examples of **ni > -N-* from Old Javanese:

bapaŋku ‘my father’ < *bapa + -(ŋ)ku*

mpu-ŋku ‘my Lord’ < *mpu + -ŋku*

ŋaraŋku or *ŋaran-ku* ‘my name’ < *ŋaran + -(ŋ)ku*

mitranta ‘your friend’ < *mitra + (n)ta* (Zoetmulder 1983:233)

Modern Javanese has lost all reflexes of **ni* except in *-né*, the third person genitive pronoun suffixed to nouns ending in a vowel.

In some Classical Malay texts dealing with Javanese tales of romance, kinship terms show an additional *-ŋ* instead of the final glottal stop or \emptyset found in other Malay varieties, cf. *bapa/ŋ* ‘father’, *kaka/ŋ* ‘elder sibling’, *cucu/ŋ* ‘grandchild’, *adi/ŋ* ‘younger sibling’. Corresponding forms in Standard Malay are *bapa/?*, *kaka/?*, *adi/?* and *cucu/?* respectively. Comparable formations are found in contemporary Malay dialects which have undergone Javanese influence in the past such as Banjarese of South Kalimantan or the Besemah and Seraway dialects of South Sumatra (Adelaar 1992:119). As these varieties of Malay have been subject to Javanese influence in the past, their ending *-ŋ* in kinship terms must be the result of morphological borrowing from an archaic form of Javanese. This *-ŋ* may reflect an earlier **ni* in the history of Javanese. It could have become part of the preceding noun through backformation when *-ŋku* fell in frequency of usage and its place was taken by *-ku*.

⁴ De Casparis (1956:33 [line 7]) translates this phrase as ‘my empire’.

1.2.2 EVIDENCE FOR AN ENCLITIC VELAR NASAL IN OTHER WEST AUSTRONESIAN LANGUAGES

The linker η occurs rather frequently in Philippine languages, and Tagalog η may stand as an example of these. It has an allomorph *na* occurring after consonants other than a glottal (-ʔ or -h), and it merges with final *n* to η ; η links the different constituents of a noun phrase. Noun phrases consist of a noun and an adjunct, and adjuncts include (a) articles, (b) deictics, (c) interrogatives, (d) quantifiers /indefinites, (e) adjectives and (f) relative clauses (Foley 1976). Examples:

- | | | |
|-----|---|--|
| (a) | <i>a-η bataʔ</i> | 'the child' |
| (b) | <i>ito-η bataʔ</i> | 'this child' |
| (c) | <i>sa ali-η partiʔ</i> | 'to which party?' |
| (d) | <i>marami-η bataʔ</i> | 'many children' |
| | <i>bahagi-η lupaʔ</i> | 'part of the land' (Schachter & Otanes 1972) |
| | <i>ilan mansanas</i> | 'a few apples' (Schachter & Otanes 1972) |
| | (< <i>ilan</i> 'some' + - η) | |
| | <i>apat na piso</i> | 'four pesos' |
| | <i>dalawa-η mansanas</i> | 'two apples' (Schachter & Otanes 1972) |
| (e) | <i>mataba-η maruno-η tao</i> | 'fat wise man' |
| (f) | <i>a-η babae-η nagbabasaʔ</i> | 'the woman reading a newspaper' |
| | <i>na-η diyariyo</i> | |

However, the following examples (from Schachter & Otanes 1972, Chapters 3-4) show that the linker also occurs (a) in compounds consisting of two nouns, (b) in noun phrases consisting of two nouns in apposition, and (c) in numerals between a digit and higher order numbers. In the latter case the linker only occurs if the digit ends in ʔ or *h*, and it is assimilated to the following consonant before *púʔ* 'ten'. Examples:

- | | | |
|-----|--|------------------------|
| (a) | <i>laro-η besból</i> | 'baseball game' |
| (b) | <i>si Pedro-η manigisdáʔ</i> | 'Pedro, the fisherman' |
| (c) | <i>dalawa-m-púʔ</i> | 'twenty' |
| | <i>tatlo-m-púʔ</i> | 'thirty' |
| | <i>isa-η daán</i> | 'one hundred' |

Correspondences with comparable functions and distributions of this Tagalog linker are well represented in the Central Philippine languages.⁵

Another language where η is found is Old Javanese. Old Javanese η is described as a definite article: it makes the following noun definite, unless this noun has already been made so by other elements in the noun phrase to which it belongs (Zoetmulder 1983:9). Examples:

- | | |
|----------------------|--------------|
| η <i>kathá</i> | 'the story' |
| η <i>dánawa</i> | 'the demons' |

η is often cliticised to a preceding word such as the demonstrative pronoun (*iki/iké* 'this', *iku/iko* 'that', *ika/iká* 'yonder'), the emphatic particle *ta*, the connective particle *ni* or the preposition (*r)i* meaning 'in, on, at; by; through; with'. It may also be used to introduce a subject, object or predicate (Zoetmulder 1983:137-142). Examples:

⁵ Zorc (1977:267) even reconstructs Proto Philippine $*\eta a \sim *-\eta$, although the justification of his Proto Philippine language subgroup remains a problem (cf. Reid (1982) and later publications).

<i>ika-ŋ rākṣasa</i>	‘that demon, the demon’
<i>warna ni-ŋ kuda</i>	‘the horse’s colour’
<i>maŋrəŋö ta-ŋ dānawa</i>	‘the demons heard it’
<i>lwah inaranan i-ŋ Sindhusâra</i>	‘a river which got the name S.’
<i>ri-ŋ dlâha</i>	‘in the future’

ŋ is also cliticised to digits when these are used attributively, or when they occur as first constituent in a compound with larger numeral units. If the digit ends in a consonant, the allomorph *-an* applies, as with *pat* ‘four’ and *nəm* ‘six’. Examples:

<i>pat-aŋ tahun</i>	‘four years’
four-aŋ year	
<i>wwalu-ŋ wiji</i> ⁶	‘eight ones’
eight-ŋ item	
<i>rwa-ŋ puluh tuŋgal</i>	‘twenty-one’
two-ŋ ten one	
<i>rwa-ŋ iwu lima-ŋ atus</i>	‘two thousand five hundred’
two-ŋ thousand five-ŋ hundred	

When numerals are used predicatively, they do not have *-ŋ* suffixed. For example:⁷

<i>lima kwèh nika-ŋ ratha</i>	‘there were five charts’
five number that-DEF chart	(lit. the number of charts was five)

The indefinite quantifier *pira* ‘how much, how many; some’ also requires *-ŋ* when it is used attributively. Compare

<i>pira-ŋ warsa kunaj lawas...</i>	‘(now) after some years...’
some-ŋ year now long	

with the following sentence, where *pira* is used predicatively:

<i>pira ta lawas nira hana ŋkâ..</i>	‘when they had been there for some time...’
some EMP long they be there	

Modern Javanese has lost *ŋ* almost completely. It only appears as an enclitic with attributively used digits, or with digits used as the first constituent of a complex numeral, the second constituent of which is *puluh* ‘ten’, *atus* ‘hundred’ or *èwu* ‘thousand’. It also appears with attributive *pira* ‘how much/many’. The allomorph *-aŋ* only applies with *pat* ‘four’. Examples (from low Javanese):

<i>ro, loro</i> ‘two’	<i>ro-ŋ puluh</i> ‘twenty’
<i>pat</i> ‘four’	<i>pat-aŋ puluh</i> ‘forty’
<i>təlu</i> ‘three’	<i>təlu-ŋ atus</i> ‘two hundred’
	<i>təlu-ŋ dina</i> ‘three days’
<i>pira</i> ‘how much?’	<i>pira-ŋ dina?</i> ‘how many days?’

It has become fossilised in the preposition *iŋ* and in the relative markers *si/ŋ* (low register) and *iŋka/ŋ* (high register).

⁶ *Wiji* (lit. ‘seed’) is used as a numeral classifier.

⁷ Abbreviations used in examples are as follows: DEF – definite, EMP – emphatic, O – object, PREP – preposition.

Zoetmulder (1983:33) claims that *iŋka/ŋ* developed from an earlier demonstrative pronoun *ikâ* + a definite article *-ŋ*, and he demonstrates this with the following Old Javanese sentence:

Tan mâŋkang nâga Takşaka, ikaŋ sumahut iŋ wwaŋ atuha nira.
 [mati-ika-ŋ] [ika-ŋ]
 not dead-that snake *Takşaka* that-the bite O-PREP person old his
 'The snake which bit his father did not die.'

The literal translation of this sentence is 'the snake, the-[one]-having-bitten-his-father, did not die'. In this sentence one can see the development from a demonstrative pronoun or definite article⁸ in Old Javanese to a relative marker which it has become in modern Javanese. Zoetmulder (1983:33) likewise assumes that *siŋ* "is still a definite article", which I interpret as meaning that *siŋ* must have developed from the personal article *si* + the definite marker *-ŋ*.

Also among other non-Philippine Western Malayo-Polynesian languages the linker *ŋ* seems to occur only sporadically. The only position where it is found is in numeral compounds. On the basis of evidence from Bario Kelabit (Sarawak), Simalur and Tongan, Blust (1974b:7) reconstructs a PMP numeral linker **ŋ(a)*, reflexes of which are found after digits and before higher order numerals. But the use of the linker in numeral constructions was broader. It occurred:

- (1) between the initial digit and following higher order numerals (i.e. a reflex of MP **puluq* 'ten', **Ratus* 'hundred' or **Ribu* 'thousand'), and
- (2) after attributively used numerals and quantifiers.

For instance, the Batak languages maintained *ŋ* in numerals and quantifiers which are used attributively before numeral classifiers or which precede the word for 'ten' in numeral compounds. Compare the following Toba Batak examples:⁹

sa/m-pulu 'ten'
obuk piga-ŋ buluŋ? 'how many hairs on the head?' (lit. head-hair how many leaves)
lasunasa/m-batu 'one piece of onion'
gaol pitu-ŋ sihat 'seven rows of bananas'

Balinese has maintained *-ŋ(a)* in attributively used digits ending in a vowel, and (b) in quantifiers when these are used attributively (i.e. before nouns or classifiers). It has also maintained it in numeral compounds between digits and higher order numerals. Examples from low Balinese (note that final *a* is pronounced as *ə*) are:¹⁰

(a) *lima-ŋumah* 'five houses'
kuda-ŋ dina? 'how many days?'
sela kayu lima-ŋ bæsi? 'five cassava roots' (*sela kayu* 'cassava';
bæsi? = numeral classifier for roots)

(b) *tlu-ŋ asa* 'thirty' ((*d*)*asa* 'ten')
pitu-ŋ asa 'seventy'
tlu-ŋ atus 'three hundred'

⁸ Zoetmulder (1983:31) remarks that the deictic value of *ika-ŋ* is often minimal, so that its function is almost reduced to that of a definite article.

⁹ For the sake of clarity I follow Warneck's (1977) spelling. It is more conservative than the spelling used by Van der Tuuk (1971), and it ignores most of the far-reaching effects of sandhi in Toba Batak.

¹⁰ I am grateful to Ida Ayu Mediani to whom I owe the Balinese examples in this paper.

<i>pitu-ŋ atus</i>	'seven hundred'
<i>ulu-ŋ tali</i>	'eight thousand'

Madurese has maintained *ŋ* in the digits *petto?* 'seven', *ballu?* 'eight' or *saga?* 'nine' when these are used attributively or in a numeral compound with *polo* ('ten') (Penninga & Hendriks 1937, Appendix p.3). Examples:

<i>petto-ŋ ropiya</i>	'seven rupiah'
<i>ballu-ŋ are aggi?</i>	'in eight days' (<i>are</i> 'day'; <i>aggi?</i> 'again')
<i>saga-ŋ polo</i>	'ninety'
<i>saga-ŋ barna</i>	'nine kinds' (<i>barna</i> 'kind')

Malagasy and Maanyan maintain a reflex of **ŋ* in numeral compounds between digits and higher order numerals (Dahl 1951:268ff.). So do the South Sulawesi languages (Mills 1975:230-231). Buginese phrases such as *duaŋŋəssə* 'two days' (< *dua* 'two' + *-ŋ/-* + *əssə* 'day') suggest that in the history of the South Sulawesi languages **ŋ* was also found with attributively used digits (Sirk 1979:104, n.52).

The Way Lima variant of Lampung (South Sumatra) still has a ligature *ŋaN-* in numerals between digits and *puluh*, the word for 'ten', and sometimes before numeral classifiers (Walker 1976:16-17). This *ŋaN-* is possibly a derivation from an earlier **ŋ*. Examples:

<i>xua ŋam-puluh</i>	'twenty'
<i>pa? ŋam-puluh</i>	'forty'
<i>təlu ŋam-biji manu?</i>	'three chickens' (<i>təlu</i> 'three'; <i>biji</i> '(numeral classifier)'; <i>manu?</i> 'chicken')

Nias has a linker *ŋa-* in numeral compounds between digits and higher order units (Sundermann 1913:63).

Sichule (which is closely related to Simalur) has *ŋa* or the assimilated remainders of a proclitic **ŋ* in numeral compounds between digits and higher order units and between attributively used digits and following nouns or numeral classifiers (Kähler 1955:57-59).

Proto Minahasan used a linker **ŋa* in these cases (Sneddon 1978:101-103).

The Malayic languages (including Malay, Minangkabau in Sumatra, Banjarese, Iban, Kendayan and Salako in Borneo), Chamic languages, Acehnese and Sundanese lost *ŋ* everywhere except in a few fossilised cases in Malay (see section 1.2.1) and in Sundanese (see section 1.5).

There is also no evidence for *ŋ* in Ngaju Dayak, in Gayo or in Rejang, but the sources for Rejang are rather incomplete. The loss of a linker in Malayic, Chamic, Acehnese and Sundanese is interesting from a classificatory point of view as it seems to bear out a close relationship between these languages.¹¹

1.2.3 CONCLUSION

The ligature *ŋ(a)* is in some form or other found in a large number of MP languages. Tagalog uses *ŋ/na* in almost any type of adjunct + noun phrase. Old Javanese, where its use is more limited, has it as a definite marker in noun phrases, where it is often cliticised to

¹¹ But cf. Nothofer (1985:297) who considers the relation between Malay and Madurese at least as close as that between Malay and Sundanese.

demonstratives, emphatic particles or prepositions preceding the head. Both Tagalog and Old Javanese also use *ŋ* (*/na*) with digits and quantifiers in larger numeral compounds or noun phrases. South Sulawesi, the South-East Barito languages, modern Javanese, Madurese, Balinese and possibly also Lampung, Nias, Sichule and the Minahassan languages use *ŋ*(*a*) only as an element encliticised to digits in larger compounds and to attributively used digits and quantifiers. Finally, Ngaju Dayak, Gayonese, (Rejang?), Achinese, Sundanese and the Chamic and Malayic languages do not have it as a living morphosyntactic element. This can mean two things: (1) that a PMP linker **ŋ(a)* developed from a linker with numerals and quantifiers to a general linker in phrases consisting of an adjunct + noun, or (2) that a PMP linker **ŋ(a)* occurring in noun phrases and in constructions with numerals and quantifiers lost part or all of its functions in most non-Philippine MP languages. I choose the latter explanation in view of the fact that **ŋ(a)* has the larger range of functions in the Philippine languages, which are morphosyntactically probably among the most conservative in the MP branch, and furthermore that Javanese, a non-Philippine language, has reduced the use of *ŋ* considerably over time. It should be remembered here that Javanese is the only Austronesian language with a well-documented history for over a thousand years. It remains unclear how **ŋ(a)* resulted in an allomorph **na* in Philippine languages.

The fact that Sundanese, Acehnese, and the Chamic and Malayic languages lost *ŋ* as a clitic with digits may be due to the fact that they replaced the PAN numerals **pitu* 'seven', **(w)walu* 'eight' and **siwa* 'nine' with numerals ending in consonants. Such a development might have marginalised the use of *ŋ*.

In view of the gradual loss of the linker **ŋ* in non-Philippine MP languages, I assume that Standard Malay maintained a fossilised **ŋ* in its relative pronoun *ya/ŋ*, and that Malay *bara/ŋ* originated from an original PMP indefinite quantifier **baRa + *-ŋ*. In the early history of Malay, PM **bara* was followed by **ŋ* in noun phrases. This **ŋ* became gradually encliticised and the resulting *bara/ŋ* was in some cases reinterpreted as a noun meaning 'thing', and in other cases as a marker of indefiniteness or uncertainty.

A development similar to **baRa + *-ŋ* must have occurred with Toba Batak *manaj*. This word often occurs in combination with interrogative pronouns and is glossed as follows:

manaj 'or'; *manaj...manaj...* 'either...or...'; *manaj ise* 'somebody, anybody, whoever'; *manaj aha* 'something, whatever'; *manaj tu dia* 'wherever to'; *manaj andigan pe* 'whenever' (Warneck 1977)

(Toba Batak *ise* 'who'; *aha* 'what'; *tu dia* 'whereto'; *andigan* 'when (referring to future)'; *pe* 'also; even'.)

Manaj must be related to Malay *mana*, an interrogative pronoun which, depending on context, means 'where; which; what; how; why' (Wilkinson 1959). It developed from an interrogative pronoun or indefinite marker **mana + *-ŋ*. It acquired **-ŋ* through its occurrence in conjunction with interrogative pronouns or other phrase complements.

1.3 EVIDENCE FOR A PMP **baRa* DENOTING INDEFINITENESS AND/OR QUESTION

Blust (1980a:48) considers Bare'e *bara* 'particle of uncertainty of knowledge, perhaps' as a reflex of either **baraŋ* or **barəŋ* 'marker of uncertainty, conditionality or hope'. Although this is phonologically sound (as Bare'e lost original final consonants), this form can equally

well be related to a PMP **baRa* denoting uncertainty and indefiniteness of object or number. Other reflexes are:

Tagalog *baga* ‘interrog[ative] adv[erb], (particle used in questions). Var[jiants:] *ba* (northern Tag[alog], *ga* (southern Tag[alog])’

Karo Batak *bara ŋe* ‘maybe’¹²

Buginese *səbara tau* ‘any person’ (Sirk 1979:104, n.53; *tau* ‘human’)

Madurese (*sa-*)*bara* ‘(a) hundred million’

Old Javanese *bara-bara*, *ba-bara* = *baraj*; *baraj* ‘anything which, whatever, just any(-body); just as it comes, just anyhow, indiscriminately, without further thought, without sufficient cause; at any time, continually’

modern Javanese *bara-bara* ‘fortunately that, it would be good if’ *bara-a*, *m-bara* ‘maybe’; *pira-ŋ bara*, *sə-bara* ‘when, as soon as’

Iban *bara?* ‘given to frivolity: *b[ara?] bəmain kita? tu?* ‘you do nothing but play about’; *b[ara?] dara* ‘a frivolous woman’

The Tagalog correspondence became a question marker. The change from a marker of indefiniteness and uncertainty into a question marker is not difficult to conceive, especially since *ba* and *ga* are more frequently used in information questions than in yes-no questions (Schachter & Otnes 1972:424). A complicating factor for Tagalog is that it also has another correspondence, *bala*, an indefinite pronoun which often occurs with the linker *ŋ* suffixed to it, as in *sa balaŋ manalo* ‘to whoever will win’ (Kern 1918:172); Kern proposes *bala* as a regular cognate form of Malay *bara/ŋ*. The semantics of *bala* would allow for such a cognate relation, but the sound correspondences do not, since PMP **R* regularly becomes Tagalog *g* (as in *baga*), and not *l*, except in Malay loanwords. This leaves three possibilities for the origin of *bala*. Firstly, the agreement between Tagalog *bala* and Malay *bara/ŋ*, etc. is due to chance resemblance. Secondly, *bala* is borrowed from an early form of Malay in which **bara?* had not yet acquired the now fossilised suffix **-ŋ*. Thirdly, *bala* is borrowed from Malay and it reflects Malay *bara/ŋ*, but the final nasal was interpreted as a linker and became disconnected from the root through backformation. With the historical data at hand I find it impossible to decide between these three possibilities. In spite of this, however, it is clear that *baga* (whether in full or reduced form) is the inherited Tagalog reflex of PMP **baRa*, and not *bala*.

In Madurese *sa-bara*, the semantic shift went from ‘indefinite number’ to ‘a number that is difficult to grasp because of its largeness’, such as a hundred million. A shift from ‘indefinite number’, ‘night’, ‘obscurity’ or ‘mist’ to ‘thousand’ or higher numbers is not uncommon, (cf. Malagasy *alina* I ‘10,000’, *alina* II ‘night, obscurity’; Dairi Batak *gəlap* I ‘obscure’, *gəlap* II ‘10,000 or higher numbers’).

In Iban, the meaning elements of uncertainty and indefiniteness changed to frivolity.

I assume that pre-Malay **bara*, Tagalog *baga*, Bare’e, Karo Batak and Buginese *bara*, and Iban *bara?* are reflexes of the same MP protoform which I reconstruct as **baRa*, a marker of uncertainty and indefiniteness of object or number. Old and modern Javanese *bara* is probably borrowed from Malay, since both Javanese *b* for PMP **b* and Javanese *r* for PMP **R* (instead of expected *w* and *ø* respectively) indicate borrowing.

¹² *ŋe* is an emphatic marker which is cliticised to the preceding word and which is realised as [ŋ] if this word ends in a vowel. As a result, *bara ŋe* is pronounced [baraŋ]. This apparently confused Neumann (1951) who wrote for this form: ‘*bara ŋe* = *baraj* I?’ and for *baraj*: I ‘maybe, possibly; II good(s), belongings, possessions’.

The assignment of Malay *bər/apa* to a now lost pre-Malay **bara* (< PMP **baRa*) + the interrogative pronoun *apa* ‘what?’ is phonologically sound. In Malay phonological history a word of more than two syllables tended to become contracted if it contained adjacent vowels or vowels that were separated from each other by a glottal. Furthermore, in most present-day variants of Malay, antepenultimate vowels have as a rule been neutralised to ə (for examples, see section 2). An allegedly frequently used phrase like **bara *apa* would therefore have become *bər/apa* in allegro speech, for example:

Proto Malayic **bara + *apa* > pre-Malay **barāpa* > **barapa* > *bərapa*

That Proto Malayic **bara* and **apa* did not yet occur as a compound is evident from the fact that this compound has only a limited distribution in the Malayic language subgroup; cf. Banjarese *sa’apa*, Salako *saṅape*, Kendayan *saṅahe*, Iban *məsak* (occurring along with *bərapa*, which is borrowed), Minangkabau *barā* (< **bara - a(ha)*), Jakartanese *bərapè*, South Sumatran Malay *bəxapə*, Urak Lawoi’ *brapa* (which is a loan).¹³

1.4 EVIDENCE AGAINST EARLIER RECONSTRUCTIONS ON THE BASIS OF MALAY *baraŋ* AND ITS CORRESPONDENCES

The following reconstructions have been proposed on the basis of Malay *baraŋ* in the sense of ‘goods, commodity; article, object; thing; baggage, luggage’:

PMP **ba[r]aŋ* ‘goods’ on the basis of Malay *baraŋ*, cf. Toba Batak *baraŋ* ‘goods’, Javanese *warəŋ* [sic]¹⁴ ‘goods’ and Ngaju Dayak *baraŋ* ‘goods’ (Dempwolff 1938:23)

Proto Malayo-Javanic [PMJ] **baraŋ*, cf. Sundanese, modern Javanese *baraŋ* and Madurese *bháráŋ* ‘thing, stuff, goods’ (Nothofer 1975:137)

Moreover, Mills (1975:625) on the basis of corresponding forms in South Sulawesi languages reconstructs

Proto South Sulawesi [PSS] **baraŋ* ‘goods, merchandise’.

On the basis of *baraŋ* in the sense of ‘any’, ‘more or less’, ‘perhaps’ and ‘would that’, Blust (1980a:48) reconstructs the following doublets:

PWMP (a) **baraŋ*, (b) **barəŋ* ‘marker of uncertainty, conditionality or hope’, cf. (a) Kankanay *bálaŋ* ‘provided, it is to be seen if, we will see if, it is to be hoped that (used only in tales)’, Hanunóo *báraŋ* ‘perhaps, expressing uncertainty or suspicion’, Karo Batak ‘perhaps’; (b) Ilokano *báreŋ* ‘let us hope, hoping; if perhaps, maybe, haply, possibly’, Kayan *barəŋ* ‘apologetically – “it is not as if”; interrogatively – “is it not?”’

In a note Blust adds:

Isneg *báraŋ* ‘a conjunction: if, perhaps, haply, let us hope’, Iban *baraŋ* ‘according to, if, any, depending on, etc.’, Malay *baraŋ-kali* ‘perhaps’, *baraŋ siapa* ‘whoever’, Bare’e *bara* ‘particle of uncertainty of knowledge, perhaps’ can be assigned to either set.

¹³ My conclusion that Iban *bərapa* and Urak Lawoi’ *brapa* are Malay borrowings is based on the fact that these languages have no interrogative pronoun *apa* nor any derivation thereof (both languages have *nama* instead).

¹⁴ In Javanese *warəŋ* actually means ‘rat poison’. It seems that Dempwolff unconsciously applied an expected sound law to Javanese *baraŋ*, which has more or less the same meaning as Malay *baraŋ*.

I reject the protoforms proposed above on a number of grounds which may not be conclusive in themselves but which reinforce each other to the point of making it unlikely that there ever was a P(W)MP, PMJ or PSS **baraŋ*.

Firstly, these etyma reflect PAN **r*. This protophoneme has been the subject of some controversy among Austronesianists. Some reject it and find that it is based on the evidence of loanwords (e.g. Wolff 1974), whereas others believe that it is a firmly established protophoneme (e.g. Blust (1980a:20ff.) and numerous other Blust publications). But no matter what stand one takes on this matter, there is general agreement about the fact that PAN **r* was reconstructed in a large number of etyma which turn out to be false (in contradistinction to etyma containing **R*).

Another factor is that the alleged reflexes of P(W)MP/PMJ/ PSS **baraŋ* have remained remarkably similar in meaning, and that some of them even reflect the nominal as well as the adverbial meanings which have been reconstructed. Both meanings are found in Malay, and all languages showing reflexes of P(W)MP/PMJ/PSS **baraŋ* have been under considerable direct or indirect Malay influence.

Finally, the phonological shapes taken by the alleged reflexes of **baraŋ* generally favour a borrowing argument. This is particularly clear in the case of Javanese *baraŋ*, Madurese *bharaŋ* and Kankanay *bálaŋ*. Javanese *b* (in positions other than following a schwa) and Madurese *bh* are usually loan phonemes (Nothofer 1975:130). Kankanay *bálaŋ* (which only occurs in tales, cf. Vanoverbergh 1933) reflects neither PAN **r* nor PAN **R* or **l*. These three protophonemes generally merged as Kankanay *l*, but Reid (1973) points out that in certain positions they merged as *ʔ*, *w* or a voiced velar fricative. In the environment **a_ a* the expected reflex is a voiced velar fricative, and therefore *bálaŋ* must be a loan.

Ilokano and Kayan both have *barəŋ*, which obliged Blust to reconstruct a PWMP doublet **barəŋ*. The schwa in Ilokano and Kayan *barəŋ* would at first sight suggest that these forms cannot be borrowed from Malay. But both languages sometimes change the last syllable *a* of loanwords into a schwa. Compare the following loanwords in Ilokano (Reid 1982:209-210):

<i>pandək</i> 'short'	< Tagalog <i>pandak</i> 'id.'
<i>siŋkəd</i> 'ratify, confirm'	< Tagalog <i>siŋkad</i> 'ratify, confirm'
<i>timbəŋ</i> 'scales'	< Malay, Tagalog <i>timbangan</i> 'id.'

Compare the following loanwords in Kayan:

<i>adət</i> 'customs, law; behaviour'	< Malay <i>adat</i> 'id.' (< Arabic)
<i>harəp</i> 'hope, wish'	< Malay <i>harap</i> 'id.' (< [Old] Javanese)
<i>səluən</i> 'trousers'	< Malay <i>səluar</i> 'id.' (< Persian)
<i>akən</i> 'cunning'	< Malay <i>akal</i> 'id.' (< Arabic)
<i>tuən, tuan</i> 'European'	< Malay <i>tuan</i> 'id.'

So there is no phonological obstacle to considering Ilokano and Kayan *barəŋ* as loanwords from Malay. Moreover, in the case of Kayan *barəŋ* (with the above meaning) Southwell (1990) also has a synonym *barəŋ* 'equipment, things', which he labels as a Malay loanword, and which occurs along with a more original synonymous term *davən*.

1.5 EVIDENCE AGAINST BRANDES' ETYMOLOGY FOR MALAY *bərapa*

As far as I know, no reconstruction has been proposed on the basis of *bərapa*. *Bərapa* replaced an earlier PAN **pija* which is still well represented all over the Austronesian language family.

Brandes (1884) analysed *bərapa* as a form of *apa* 'what?' prefixed with the intransitive verbal marker *bər-*, and later scholars have not yet questioned his etymology. But there is no strong functional or semantic argument for this analysis. Besides, Brandes' argument would fail to explain quantifiers on the basis of *bar-* or *bər-* in languages which do not have an intransitive verbal prefix *bar-/bər-*, such as Sundanese and Madurese.

Sundanese has (*sa*)*baraha* 'how much/many?', which consists of an unidentified (*sa-*)*bar/-* + *aha* 'what?' Sundanese intransitive verbs are generally marked with an infix *-um-*, and there are also a few intransitive verbs with *a-*, *ma-* and *ba-*, but not with *bar-*.¹⁵

Madurese has *barampa* (low register) 'how much/many; how large?', It also has *bar-iyā* (low register) and *bar-into* (middle register) 'such' and *sa-bar-iyā* (low register), *sa-bar-into* (middle register) 'as big as (this)'.¹⁶ But it does not have a prefix *bar-*, nor does it have **ampa*. Other Madurese forms reflecting **baRa* are *bharəŋ* 'something, good(s), object, luggage; (the thing) that/which', *sa-bharəŋ* 'each', *bha-barəŋ* 'all', *sa-bharəŋ-an* 'average, usual; no matter what', and *sa-bara* (section 1.3). *Bharəŋ* is likely to be borrowed from Malay (section 1.4). *Sa-bara* on the other hand reflects MP **baRa* as an indicator of indefinite number (see section 1.3).

Bər/apa, *bar/aha*, *bar/ampa*, *bar/iyā* and *bar/into* have replaced PAN **pija* 'how many'. It is likely that these innovations originated through language contact between Malay, Sundanese and Madurese.

This is most clearly the case with Madurese *bar/ampa*, which may be an early borrowing from Malay.¹⁷ Madurese has no interrogative pronoun **ampa* or a derivation thereof, whereas Malay has the following range: *apa* 'what?'; *si/apa* 'who?'; *kən/apa* 'why? how?'; *məŋ/apa* 'why?'; *bət/apa* 'to what extent?'

Sundanese (*sa*)*bar/aha* is derived from **aha*, which is still found in *s/aha* 'who?', *kumaha* 'how?', *n/aha* 'why?' *ir/aha* 'when?' The development of (*sa*)*bar/aha* in analogy to *bər/apa* may be the result of Malay influence, but it could also an independent development.

2. ON THE HISTORY OF MALAY WORDS FOR ANIMAL

A remarkable fact about the Proto Austronesian lexicon is that it seems to have lacked a general term for 'animal'. The *Comparative Austronesian dictionary* (Tryon ed., 1994) provides data from eighty languages from all main branches of the Austronesian family. The general term for 'animal' in these languages is given under gloss no. 03.110.

¹⁵ Interestingly enough, what Sundanese does have is a verbal prefix or proclitic *barəŋ-* denoting indefiniteness of object. Compare:

<i>hakan</i> 'eat'	<i>barəŋ-hakan</i> 'eat things, eat (in general)'
<i>gawe</i> 'work'	<i>barəŋ-gawe</i> 'do some work, work a bit'
<i>bili</i> 'buy'	<i>barəŋ-bili</i> 'do some buying'

¹⁶ *-iyā* and *-into* are deictic elements which only occur in combination with *bar-*, *da²-* or *ka²-*.

¹⁷ In a more recent borrowing one would expect *bh* instead of *b*.

From the wide variety of terms listed under this gloss it appears that:

- there is no set of related terms which serves as firm evidence for the reconstruction of a Proto Austronesian etymon;
- many languages use paraphrases such as 'living creature', 'animate thing', 'hunting object', or they use the same term as for 'game' or 'meat';
- some languages use words the primary meaning of which is 'dog', 'bird', 'pig' or some other specific animal;
- some of the terms given actually refer to 'domestic animal' or 'cattle' rather than to 'animal' in general;
- many languages use loanwords; as far as Indonesian languages are concerned, these loanwords are often derived from either Malay *binataŋ*, Sanskrit *sattva* or Arabic *haywān* (all meaning 'animal').

An exception to the heterogeneity of general terms for 'animal' seems to be the South Sulawesi languages, which generally reflect a protoform which could be reconstructed as Proto South Sulawesi **olo?-olo?* for this notion.

As all languages which have *binataŋ* or a corresponding form have undergone Malay influence, I assume that they borrowed this form from Malay.¹⁸

Another problem is the origin of Malay *binataŋ*, which cannot be an inherited word either. Inherited trisyllabic Malay words always have a schwa in the first syllable, except when the following consonant is a semivowel or *h* (Adelaar 1992:10). The shape of *binataŋ* furthermore suggests that it contains the infix **-in-*, a PAN passive marker or nominalising infix which was already lost in PM (Adelaar 1992:193). Other Malayic languages do not necessarily have a form corresponding to *binataŋ*, cf. Iban *jəlu* '(wild) animal, game', Salako and Kendayan *laok* 'wild animal; meat, side-dish'. When they do have a corresponding form, its shape suggests borrowing because of its *i* (instead of schwa) in the first syllable, cf. Jakartanese, Seraway (Sumatra) *binataŋ*.¹⁹ The origin of Malay *binataŋ* remains unclear. Brandes (1884:175, fn.3) related it to Ibanag *batan* 'trap-net' and to Sundanese *pamataŋ* 'someone hunting deer on horseback with a spear, a sword or a lasso' (derived from *wataŋ* which is glossed 'lance' in Eringa 1984). While not impossible, the connection remains vague.

In general, then, there is little evidence for a protoform on whatever level on the basis of *binataŋ* and its correspondences, and there is also little evidence for a general and exclusive term for 'animal' in Proto Austronesian.²⁰

But we are on firmer ground when looking for the history of the term for domestic animal. Dempwolff reconstructed PMP **'ajam* 'be domesticated' which was allegedly homonymous to his PMP etymon **'ajam* 'to play'. With the increase of data since he compiled his PMP lexicon, we now know that these homonyms were probably one etymon covering a configuration of related meanings, and that this etymon should be reconstructed (in Dyen's orthography) as **qayam* 'domestic animal (including pig, dog, fowl); plaything'.

¹⁸ A remarkable corresponding form is Cham *pinataŋ*, which is glossed 'machine' in Moussay (1971).

¹⁹ A problem in evaluating forms corresponding to Malay *binataŋ* in other Malayic isolects is that the available dictionaries do not specify what the status, degree of adaptation or frequency of occurrence of such forms is.

²⁰ Dr D.J. Prentice (pers.comm.) first drew my attention to the lack of a term referring to animals as a category in Austronesian languages.

In Malay the reflex of **qayam* underwent a semantic narrowing-down, and we find *ayam* 'chicken'. But this is not the meaning which should be reconstructed for PM **hayam*, as other Malayic languages have different meanings. Compare corresponding forms like Kedayan *pa-hayam-an* 'livestock', Iban *ayam* 'plaything, toy, pet'; *uduk ayam* 'pet dog'; *ŋ-ayam* 'play'. These lead to the conclusion that Proto Malayic still had the original Proto Western Malayo Polynesian meaning configuration of 'tame' (or 'domesticated') and 'playing'. From a Proto Malayic **hayam* the meaning narrowed down to one particular domesticated animal in Malay, to 'domesticated animals (in general)' in Kedayan, and to 'plaything' and 'pet animal' in Iban. In Malay it replaced PM **manuk* 'chicken', reflexes of which are still found in Kedayan and Iban.

But it seems that developments concerning PM **hayam* have not stopped there. Malay has a word *main* for 'to play', which has corresponding forms in other Malayic languages and in Achehnese and Cham.²¹ *Main* must be a reflex of a pre-PM verbal derivation²² **q-um-ayam* 'to play', and there are a number of less common but regular Malayic changes to explain its present shape. These are:

(i) Vowel contraction and assimilation of adjacent vocoids

As I already mentioned in the last part of section 1.3, in the history of Malayic languages the tendency to vowel contraction is most commonly observed in words of more than two syllables. It was already operative in Proto Malayic and it is still so in present-day Malayic languages. It is one of the ways in which tri- or tetrasyllabic words could become disyllabic, which is phonotactically the preferred structure of a Malayic root. The contraction affected adjacent vowels, or vowels which were separated by a glottal. Examples:

PMP **Rahut* 'split wood' + **-an* > Malay *rotan* 'rattan'

PMP **ma-iRaq* 'red' > Malay *merah*

PM **kələ(h?)əmə* 'obscure; to set' **hari* 'day' + **-an* > **kələm (h)ari-an* > Malay *kə(l)maren, kə(l)marin* 'yesterday'

PM **k(a,ə) + *iri* 'left' > Malay *kiri*

PMP **ma-kaʔən* 'eat' > PM **makan*

PM **huluʔ tuʔət* 'kneecap' > Malay *lutut*

P(W)MP **(dD)əhuk (-dD)əhuk* 'sit' > PM **duduk*

Another, less common, form of vowel reduction is the assimilation of the sequence **-Vya-* to *-Vi-* in Minangkabau and Seraway, two Sumatran Malayic languages. This reduction has also occurred in disyllabic roots. Examples:

PM **bayar* 'pay' > Minangkabau *bai^ə23* Seraway *baix* 'id.'

**layar* 'sail' > Minangkabau *lai^ə23* 'id.'

**lAmpuyəŋ* 'ginger plant' > Minangkabau *lampuiəŋ* 'id.'

**bayas* 'k.o. palm tree' > Seraway *bais* 'id.'

21 The sound changes outlined in what follows were not shared by Cham or Achehnese. Therefore, as a consequence of my etymology for Malay *main*, Cham *mu'in* and Achehnese *meu'en* must be loanwords from Malay.

22 Or at least a derivation from an earlier stage than Proto Malayic, since there is no evidence that Proto Malayic still had productive infixes (cf. Adelaar 1992:193-194).

23 The endings of the Minangkabau forms *bai^ə* and *lai^ə* are the result of velarisation of **i* through a following velar fricative which was subsequently lost (so, Proto Malayic **bayar* > **baix* > **bai^əx* > *bai^ə*, and Proto Malayic **layar* > **laix* > **lai^əx* > *lai^ə*).

The same **-aya- > -ai-* assimilation has taken place in Malay *lain* 'other', which derives from an original *lai* (still found in Old Malay) + *-an* (Adelaar 1988:71).

(ii) A constraint against **-ip* and **-im* endings

In PM there seems to have been a constraint against last syllables containing **i* + a final labial (i.e. a constraint against **-ip* and **-im* endings). Several factors indicate this:

- (a) as a rule, Malayic roots ending in *-ip* and *-im* are borrowed (mostly from Arabic, e.g. *hakim* 'judge'; *iklim* 'climate'; *musim* 'season'; *nasip* 'fate'; *tabip* 'physician');
- (b) the few words ending in *-ip* or *-im* that are not readily identifiable as loanwords do not have sufficient correspondences within the Malayic subgroup to yield evidence for a PM etymon;
- (c) the best attested (and one of the very few) higher order etyma ending in **-ip* is reflected with vowel metathesis in the Malayic languages, for example:

PAN **quDip* 'live' > PM **hidup* 'id.'

(iii) Apocope of the first syllable after affixation of **-um-*

The PAN affix **um-* (before initial vowels) or **-um-* (after initial consonants) only appears in fossilised forms in Malayic languages, and it probably had already ceased to be a living affix in PM (Adelaar 1992:193-194). When **-um-* was infixated after initial **q* or a labial, the resulting syllable became reduced to *m/-* in Malay, as is shown in PAN **-um-* + **qinum* 'drink' > Malay *m/inum* 'id.' and in PAN **-um-* + **paCəy* 'death' > Malay *m/ati* 'dead'. In fact, in cases where an initial labial applies, this reduction is a rather widespread phenomenon among Austronesian languages. In the case of initial **q*, several factors account for this reduction:

- loss of initial **q* (PAN **q-* > PM **h-* > Malay (usually) \emptyset , (sometimes) *h-*;
- antepenultimate vowel neutralisation;
- a tendency towards disyllabicity;
- particularly in trisyllabic roots, initial *(h)ə-* sequences are disfavoured, and **(h)ə-* sequences as a rule were lost (Adelaar 1992:52-53).

These changes and tendencies must have led to the present shape of *main*. If we accept the possibility that contraction was limited not only to adjacent vowels but also to adjacent vowels plus semivowels (as it sometimes was in Minangkabau and in Seraway disyllabic roots, and as it may have been in Malay *lain*), Malay *main* can be derived from it through the following stages:

I	WPMP <i>*-um-</i> + <i>*qayam</i>	---->	<i>*q-um-ayam</i>
II	assimilation of adjacent vocoids: <i>*-ya- > *-yi-</i>	---->	<i>*qumayim</i>
III	contraction of <i>*-yi-</i> to <i>*-i-</i>	---->	<i>*qumaim</i>
IV	PMP <i>*-im > PM *-in</i>	---->	<i>*qumain</i>
V	<i>*q > *h</i> or \emptyset ; antepenultimate vowel neutralisation	---->	<i>*(h)əmain</i>
VI	loss of initial <i>*(h)ə-</i> sequence	---->	<i>main</i>

The relative order of stages IV, V and VI could also have been different, but they must have taken place after stages I and II.

A semantic contingency of 'domestic animal', 'pig', 'dog', 'chicken' and 'to play' is exemplified in the reflexes of **qayam* of many other Austronesian languages (Dempwolf)

(1938:13) and other sources).²⁴ Compare in this respect also Malay *anjij* 'dog' which has a phonologically regular cognate *epɛkɪ* 'domestic pig' in Salako.

3. HUMAN BEING

The general term for 'human being' reconstructed for Proto Austronesian is **Cau*. This term has reflexes in languages of Taiwan, the Philippines, Borneo,²⁵ Sulawesi, the Molucca Islands and Oceania. Other languages have a reflex of PMP **qaRtaq*²⁶ for this notion. These languages are found in Simalur (off Sumatra's west coast), in eastern Indonesia (eastern Sulawesi, central and southern Molucca Islands, Lesser Sunda Islands), in the Negrito languages of the Philippines and in some Oceanic languages (Blust 1972b:166ff.).

Other terms for 'human being' which are not limited to well-defined linguistic subgroups are those reflecting PMP **qulun* and **uRaŋ*. Reflexes of **qulun* meaning 'human being' are found in Borneo (including Malagasy) and in the various forms of Lampung; **qulun* has reflexes meaning 'slave' or 'servant' in the Chamic languages, in some Bornean languages, and in languages of Java and Sumatra. Reflexes of **uRaŋ* meaning 'human being' are found in the Chamic and Malayic languages and in Acehnese, Javanese, Sundanese and Madurese.

There is a complication involved in the meaning that should be assigned to **qaRtaq* and **qulun*. Reflexes of these etyma mean either 'human being' or 'slave' (or, in some Philippino languages, 'Negrito, black person' for **qaRtaq*).

For **qaRtaq*, Blust tries to reconcile these notions through a reconstructed meaning 'outsiders, alien people' (Blust 1972b:169).²⁷ By giving different meaning connotations to **Cau* ('real people; us; our own kind') and **qaRtaq* ('outsiders, alien people') he also accounts for the otherwise awkward fact that there are two Proto Austronesian etyma with reflexes having the meaning 'human being'. But where it is easy to see how a word can change its meaning from 'outsiders; alien people' to 'slave', it is much less easy to see how it could change this meaning into 'human being'. Some of the Philippino Negritos use a reflex of **qaRtaq* to refer to themselves, and Reid (1994) finds it unlikely that they would use as an endonym a term which originally referred to 'outsider'.

A more likely explanation is that a post-PMP **qaRtaq* originally meant 'human being' and that via slave trade and subordination this term became reinterpreted as 'slave' or 'subordinate' by the slave-trading or subordinating people. Parallel semantic developments must have taken place in the developments of the terms for 'slave' and 'Slavic person' in Western European languages, and in the uses of the term *kanaka*²⁸ in the Pacific. One of the implications of this explanation is that reflexes of **qaRtaq* meaning 'slave' or 'Negrito' must be borrowings. This would allow for a more accurate insight into the spread of inherited reflexes. Applying this explanation to reflexes of **qulun* would single out Lampung (South

²⁴ Compare also semantic shifts of PAN **manuk* 'fowl' in the daughter languages.

²⁵ But here, as far as I know, only in the Tamanic languages which are closely related to the South Sulawesi languages (Adelaar 1994).

²⁶ Compare Blust's PAN **qa(R)(Ct)A* 'outsiders, alien people' (Blust 1972b:169) as modified by Reid (1994). This etymon has no reflexes in Taiwanese languages.

²⁷ Blust (1972b) seems to suggest a reconstruction of the same meaning for **qulun* (for which Dempwolff (1938:162) gives 'human being', in which he observes 'a parallel duality of meaning').

Sumatra) and Bornean languages (+ Madagascar), as languages in which **qulun* is inherited, as these are the only languages where it means ‘human being’.

Blust’s gloss ‘outsider; alien people’ for **qaRta* would certainly befit PM **uraŋ*, which is a reflex of PMP **uRaŋ* ‘human being’ and which I initially glossed ‘human being’ (in my 1985 thesis). I did this on the basis of the usual meaning of its modern Malay reflex and on the basis of its general meaning given in dictionaries of Malayic languages and dialects. However, field experience with Salako and a further investigation of the data provided by the Malay and Iban dictionaries induce me to expand the meaning of PM **uraŋ* to ‘human being; outsider’.

The Salako term *urəkŋ*, although generally meaning ‘human being’ is never used for one’s own relatives or close friends. In fact, calling relatives or close friends *urəkŋ* would insult them and alienate them: one would call them *urəkŋ* only in order to imply a break of ties. The term *urəkŋ* is used to refer to outsiders (possible enemies, headhunters), or to an ethnic entity as in *urəkŋ Saribas* ‘the Ibans’, *urəkŋ Laut* ‘the Malays’, *urəkŋ Salako* ‘the Salakos’.

For Iban *uraŋ*²⁹ Richards (1981) gives two meanings: 1. ‘person, people, someone, anyone’; 2. ‘someone’s, another’s, other people’s’. Along with the second meaning go phrases like *apay uraŋ* [‘father of people’ =] ‘father of a family’, *utay uraŋ* [‘things of people’ =] ‘other people’s belongings’, etc.

The connotation of ‘outsider’ was lost in Malay *oraŋ*, although it is still attested in phrases like *nəgəri oraŋ* [‘countries of people’ =] ‘abroad, foreign countries’, *istəri oraŋ* [‘the wife of people’ =] ‘someone else’s wife’ and *baraŋ oraŋ* [‘things of people’ =] ‘other people’s belongings’. Further study of Malayic languages and of Classical Malay texts may yield additional evidence for an original meaning of ‘outsider’ for PM **uraŋ*.

There is corroborating evidence for this gloss outside the Malayic subgroup. The connotation of ‘outsider’ is more explicit in the Jarai and Moken reflexes of PMP **uRaŋ*.³⁰ Jarai *arəŋ* is glossed as ‘undefined person’ (Lafont 1968), and in running text it is often translated as ‘someone else’ or ‘other people’ (cf. French *autrui*, Lafont 1963:39 and passim). Moken *olaj* is glossed ‘another, [? he, etc.]’ (Lewis 1960:90).

The Jarai and Moken evidence would indicate that the connotation ‘outsider’ can also be attributed to an ancestral form **uRaŋ* in a protolanguage of a higher order than PM. As a matter of fact, this seems to be possible, although there remain some reflexes of PMP **uRaŋ* showing semantic developments which are not directly clear, and which need further investigation. Dempwolff labelled his PMP **uRaŋ* ‘human being’, but present-day Malayo-Polynesian languages often have corresponding forms meaning ‘affine’, ‘friend’ or ‘cross-sibling’. Compare:

Cebuano	<i>ugáŋan</i> ‘parent of one’s spouse’
Macassarese	<i>uraŋ</i> ‘companion’; <i>si-uraŋ</i> ‘with’
Yamdena (Tanimbar)	<i>ure</i> ‘cross-sibling + parallel cousin’
Kei	<i>uran</i> ‘cross-sibling, cousin’

²⁸ *Kanaka*, the Hawaiian word for ‘person’, acquired the meaning ‘seasonal labourer’ in Tok Pisin, where it is now used in a pejorative sense to refer to an uneducated person from the bush.

²⁹ Spelled ‘orang’ in Richards’ orthography.

³⁰ Jarai is spoken in Pleiku province, Vietnam; Moken is spoken in the Mergui Archipelago, South Myanmar.

Sikka (East Flores)	<i>wra</i> 'affine of the same sex of a woman (father's sister's daughter, mother's brother's daughter, husband's sister, brother's wife)'
Tana Ai (dialect of S. Kanere, East Flores)	<i>ura</i> 'sister, brother's wife, mother's brother's daughter, father's sister's daughter (woman speaking); cross-cousin, affine of the same sex (woman speaking)'
Tetun (Timor)	<i>oan</i> 'child, offspring'
Toba Batak	<i>uraj</i> 'offspring'
Tondano	<i>uraj</i> 'offspring'

These meanings must be related, and Dempwolff's gloss 'human being' for PMP **uRaŋ* is therefore somewhat misleading. The latter must be a semantic specification that took place at a relatively recent date in a few West Indonesian languages (Malay, Cham, Javanese, Madurese). Dempwolff's PMP **uRaŋ* should rather be glossed 'outsider; affine; friend'.

Cebuano *ugáŋan* refers to an affinal relationship, and Macassarese *uraj* to friendship.

The unspecified meaning which Tanimbar *ure* and Kei *uran* have in common is that of 'cross-sibling'. In the Tanimbar and Kei context, cross-sibling terms emphasise the fact that female siblings will eventually become members of a different clan, which is ideally one particular wife-taking clan in a chain of clans in a circular connubium system. So they will eventually become outsiders, members of an affinal clan. The meaning of Tana Ai *ura* and Sikka *wra* (basically, 'female affine of a woman') seems to be a later development thereof.

The Tetun, Toba Batak and Tondano meaning of 'offspring' must also have developed from PMP **uRaŋ*, but here the semantic development is much less transparent (particularly considering the fact that Toba Batak *uraj* has a relatively long history in common with Karo Batak *t/uraj* 'cross-sibling', see below). The Toba Batak term for 'son' is *anak*, which reflects PMP **anak* 'offspring'. The fact that PMP **anak* became marked for male gender in Toba Batak would suggest that another term for 'offspring', such as Toba Batak *uraj*, was originally also marked for (female) gender. However, the present Toba Batak term for 'daughter' is *boru*, which speaks against such a speculation. The semantic developments of Tetun *oan* and of Toba Batak and Tondano *uraj* require further study.

The configuration of meanings 'outsider', 'affine' and 'friend' is also represented by the reflexes of Blust's (1970:125) PMP etymon **tuRaŋ* 'in-law'.³¹ Compare:

Tagalog	<i>ma-núgaŋ</i> 'child-in-law'
Bikol	<i>tugaŋ</i> 'sibling; in-law'
Western Bukidnon Manobo	<i>e-nugaŋ</i> 'parent-in-law'
Proto Oceanic	<i>*turaŋ</i> 'companion, friend, neighbor, various kinsmen'

The following reflexes can be added to Blust's material:

Ilokano	<i>katugaŋan</i> 'parent of one's spouse'
Bontok	<i>katogáŋan</i> 'id.'
Isneg	<i>túxaŋ, manúxaŋ</i> 'son-in-law, daughter-in-law'; <i>katuxáŋan</i> 'father-in-law, mother-in-law'

³¹ As Blust's PMP **(Ct)uRaŋ* only goes back as far as PMP (which merged PAN **C* and **t* to **t*) and furthermore PMP **T* turns out to be an erroneous protophoneme (Dahl 1981:23-25), I will use a more convenient PMP representation **tuRaŋ*.

Sangirese	<i>tuhaj</i> 'older sibling'; <i>manuhaj</i> 'child-in-law'
Karo Batak, Dairi Batak, Alas Batak (= the northern Batak linguistic subgroup)	<i>turaj</i> '(term of address to cross-sibling)'
Buru	<i>tuha</i> 'to accompany; with' (Grimes 1991:269)

It appears that reflexes of PMP **tuRaŋ* refer to affinal relationships in Philippine languages,³² to relations of friendship or to kin relationships in Oceania, and to a cross-sibling relationship in Northern Batak. As the Batak peoples have social organisations which are in some crucial ways similar to that of the Tanimbarese and Keiese peoples (notably with a prescribed circular connubium system), the Karo Batak concept of 'cross-sibling' must be closely associated with that of 'affine'.

The meaning of Sangirese *manuhaj* indicates an affinal relationship (the meaning of Sangirese *tuhaj*, on the other hand, does not).

The Buru reflex has become a verb the meaning of which is still associated with that of 'companion'.

It is quite likely that this PMP **tuRaŋ* is related to PMP **uRaŋ*. The range of meanings of their reflexes is covered by the same meaning configuration (although in both cases there are reflexes with problematic semantic developments viz. Tetun *oan*, Toba Batak and Tondano *uraj*, Sangirese *tuhaj*). As for the initial **t* in **tuRaŋ*, Blust (1979:228) points out that PMP had a referentiality-marking prefix **t-* which is still found – mainly in fossilised form – in kinship terms in the Malayo-Polynesian daughter languages. The referential value of kinship terms reflecting **t-* is often lost. I assume that Blust's **tuRaŋ* was a referential form **t-uRaŋ* used for outsiders who became accepted to one's kin group through marriage or friendship, and that it was derived from a PMP **uRaŋ* which I tentatively gloss 'outsider; affine; friend'. The referential connotation was clearly lost in Oceanic and in the Northern Batak subgroup.

POSTSCRIPT

Terms for 'person' or 'human being' are also often used for '(house)post', 'pole' and 'mast'. Brandes (1884:120) first showed a relation between these notions by pointing out that parallel semantic configurations were found in the following cognate sets:

Tagalog, Bisaya *haligi*, Buli *arihi*, Ngaju Dayak *jih*, Buginese *aliri* 'pole, post', Malay *diri* 'stand; oneself' (cf. PMP **haDiRi* 'stand; person; self');

Malay *tiaŋ* 'pole, post', High Javanese *tiaŋ* 'person';

Malay *orag* 'person', Ngaju Dayak *owaŋ* 'post, pole'.

However, Ngaju Dayak *owaŋ* does not reflect PMP **R* (which should have become *h* in Ngaju Dayak) and its meaning is actually 'piece of wood cut out of a tree or out of other wood' (Hardeland 1859). A relation with Malay *orag* is very doubtful.

Another use of the term for 'human' is as a first person pronoun. Compare Sundanese *urag* '(1st p.pl.incl.)'; Javanese (polite language) *ŋ/hulun*, Achehnese *lōn* 'I' (cf. Achehnese

³² Except for Western Bukidnon Manobo *e-nugaj*, the Philippine reflexes agree in having a *ka-an* derivation for 'parent-in-law' and a *maN-* derivation for 'child-in-law'. In agreement with the semantics of their affixes, the *ka-an* derivations literally mean 'someone with the nature of (**tuRaŋ*), and the *maN-* derivations literally mean 'become/take on (**tuRaŋ*)'.

ulōn 'servant'; High Javanese *tiyaŋ* 'person; I', High Balinese *tiyaŋ* 'I'; Salako, Kendayan *diri*? '(1st p.pl.incl.)' (cf. Malay *diri* 'self' and *bər-diri* 'to stand' < PMP **DiRi* 'self; person; to stand'). The use of Javanese *ꦤꦸꦭꦺꦤ* and Achehnese *lōn* for the first person singular is a consequence of the habit of using terms for 'slave' or 'servant' to refer to oneself in polite language (cf. Malay *sahaya* and High Javanese *kawula* 'slave; I [= your slave]')).

LIST OF LANGUAGE SOURCES

Unless otherwise mentioned, the following language sources were used in this paper:

- | | |
|---|--|
| Acehnese: Kreemer (1931) | Ngaju Dayak: Hardeland (1858, 1859) |
| Alas Batak: Osra et al. (1985) | Nias: Sundermann (1913) |
| Balinese: I. A. Mediani (pers.comm.) | Old Javanese: Zoetmulder (1982, 1983) |
| Banjarese Malay: Abdul Jebar Hapip (1977) | Proto Austronesian: as indicated in text |
| Bare'e: Adriani (1928) | Proto Malayic (PM): Adelaar (1992) |
| Bontok: Reid (1976) | Proto Malayo-Javanic (PMJ): Nothofer (1975) |
| Buginese: Sirk (1979), Mills (1975) | Proto Malayo-Polynesian (PMP): as indicated in text |
| Buru: Grimes (1991) | Proto Western Malayo-Polynesian (PWMP): as indicated in text |
| Cam: Moussay (1971) | Proto Minahasan: Sneddon (1978) |
| Cebuano: Wolff (1972) | Proto Philippines: Zorc (1977) |
| Dairi Batak: Tindi Radja Manik (1977) | Proto South Sulawesi (PSS): Mills (1975, 1981) |
| Gayo: Soravia (1984) | Rejang: McGinn (1982), Jaspán (1984) |
| Ilokano: Constantino (1971) | Rhade: Tharp and Y-Bham Buon-Ya (1980) |
| Isneg: Vanoverbergh (1972) | Salako: Adelaar (unpubl. fieldnotes) |
| Jakarta Malay: Abdul Chaer Mad'ie (1976) | Sangirese: Steller and Aebersold (1959) |
| Jarai: Lafont (1968) | Seraway ("Middle Malay"): Helfrich (1904) |
| Javanese: Gericke and Roorda (1901) | Sichule: Kähler (1955) |
| Kankanay: Vanoverbergh (1933) | Sikka: Arndt (1933) |
| Karo Batak: Neumann (1951) | South Sulawesi: Mills (1975, 1981) |
| Kayan: Southwell (1990) | Sundanese: Kats and Soeridiradja (1933) |
| Kei: Barraud (1979) | Tagalog: Schachter and Otnes (1972), Foley (1976), Panganiban (1980) |
| Kendayan: Adelaar (unpubl. fieldnotes) | Tana Ai (East Flores): Lewis (1988) |
| Lampung: Walker (1976) | Tetun: Morris (1984) |
| Maanyan: Dahl (1951) | Toba Batak: Van der Tuuk (1971), Warneck (1977) |
| Madurese: Penninga and Hendriks (1937) | Tondano: Wantalangi et al. (1985) |
| Malagasy: Dahl (1951) | Urak Lawoi' Malay: Hogan (1988) |
| Malay: Wilkinson (1959) | Yamdena (Tanimbar): Drabbe (1932b) |
| Minangkabau: Moussay (1981), Van der Toorn (1891) | |