Origins of the Philippine Languages

Cecilio Lopez

Origins of the Philippine Languages

CECILIO LOPEZ

1. *Introduction.* In the Philippines there are about 70 languages and in Malayo-Polynesia about 500. To say something about unity and diversity among these many languages and about our evidence for their Malayo-Polynesian source is not an easy task, particularly when it must be done briefly and for readers without the requisite linguistic sophistication. For the purposes of this paper, therefore, the best I can do is to explain some representative phenomena, making the presentation as simple as I possibly can.

In determining similarities and diversities between languages, comparison based on any one of four levels may be used: phonology, morphology, syntax, or vocabulary. The approach using all the four levels is undoubtedly the most reliable, particularly if it fulfills the following conditions: exhaustiveness of coverage, simplicity of exposition, and elegance of form. So far as I know, no comparative study of a group of cognate languages has ever been written which fulfills all these conditions.

The favorite hunting ground of linguists is phonology—which includes vocabulary—because vocabulary items are comparatively easy to collect and are subject to rigorous treatment. In MP languages quite extensive studies have been carried out in phonology, but those in morphology and syn-

tax have been rather scanty.[1]   It is not surprising, therefore, that in this paper my discussion of phonology is more detailed than that of morphology and syntax.   Most of the materials I use will be those of my own collection, supplemented by other sources.[2]

To trace the Philippine languages back to MP and other sources is to get involved in the problems of origin, a subject clouded in the mists of prehistory.   The discovery of origins is not easy, as one knows who has made the attempt.   Because of the pressures of technology and progress, the modern man tends to think in terms of patterns and processes and not concern himself with origins.   Yet since man is the only member of creation who can concern himself with the ideas and events of the past, I shall run the risk of being labeled old-fashioned if thereby I am also more properly human.

The MP world, according to the older school of anthropologists, was peopled by migration(s) from west to east, from the Asian mainland into the Pacific.   One theory asserts that the river regions of western China and the borders of Tibet were the original home of Indonesian (Proto-Malay and Deutero-Malay) culture, and also the home of early Indian culture.   In migrating southward from this homeland, the streams of migration bifurcated, one branching westward into India and the other into Indo-China and Indonesia.   The latter branch then spread northward to the Philippines and Formosa and eastward across the Pacific, where it formed an essential part of the Polynesian population.   These migrations were not to be interpreted in the strict meaning of the word.   They were quite probably slow, long and drawn-out movements, accompanied by some assimilation along the way, with the earlier migrants pushed by the later, by mixture of languages and cultures, or even by imperfect adoption of the language of one group by the other.   According to another theory, there are evidences of waves and periods of migration originating

---

[1] Abbreviations and symbols used in this paper are listed after the text.

[2] I have not attempted detailed documentation; however, the most important sources are listed in the bibliography.

from somewhat different locations in southeastern Asia. In the Philippines, the diversity of the peoples indicates that most of the groups have been occupying the same regions for some time, with limited intercourse with the outside world, interrupted occasionally by sporadic movements.

Like many anthropologists, linguists of the older school point to Indonesia as the original home of the Polynesians, and hold that in a series of eastward migrations the early Indonesians left traces of their languages among the Melanesian-speaking peoples through whose regions they passed.

Some younger linguists, and some ethnologists and archeologists consulting the linguistic evidence, entertain different views. One view holds that the original speakers of MP were the remote ancestors of the present Polynesians, that the first great movement of this sea-faring people was into Indonesia. (This calls to mind Bowring's mention of Father Zuñiga's observation "that the peoples of the Philippines were originally colonized by the inhabitants of America.") Another claims that the place of origin of the MP peoples is rather to be sought between Formosa and Hainan. The assumption is questioned that everything in the Pacific region came out of Southeast Asia (or south China) with nothing going back. New Guinea, Australia, and the larger islands of Melanesia, to be sure, were originally peopled from Southeast Asia, but initial settlement occurred long before the emergence of the MP languages as a distinct family. Once these areas were settled, there is no reason why they could not have become centers of cultural development or a source of movement westward into Indonesia and Southeast Asia. The great diversity of languages in Melanesia is an argument advanced in favor of that area as the springboard of MP languages; this proceeds on the assumption that the whole of a large number of groups of people is not likely to migrate as a collection of distinct groups. On the other hand, the languages of western Indonesia and most, if not all, of the languages of the Philippines, constitute a single group. It is in the framework of these conflicting theories that I will present some linguistic evidence.

2. *Similarities and diversities among Philippine languages.*

2. 1. *Phonology.* On the phonological level, the genetic relationship between languages is determined by comparing the correspondences of sounds and the meaning of the words compared. There are instances where the sound correspondences and the meaning are almost identical; in others deviations may occur in the sound correspondences with the meaning identical or similar; and in others deviations may occur both in sound correspondences and in meaning. Unless other factors are involved, such as borrowing, for instance, such words are cognates, and their proto-forms can be recovered and reconstructed.

I have chosen 15 representative examples for comparison in the Philippine languages. These will be compared with languages in IN (non-Philippine), MN, and PN in section 3. 2. 1.

2. 1. 1. Eng 'eye': Tag Seb Hil Ilk Bkl S-L Png Pmp Ibg Sbl Ivt Nbl Apa Tao Mar *mata* 'eye.'

2. 1. 2. Eng 'thatch (roofing)': Tag Btk P-Neg *atip*, Seb Hil Bkl S-L Mgd Tao *atup*, Ibg *atoq*, Png Sbl Cuy Mar *atĕp*, Pmp Apa *atap* 'thatch (roofing),' Nbl *atip* 'sheath,' Ilt *atap* 'thatch,' *atep* 'roofing.'

2. 1. 3. Eng 'hulled rice': Tag *bigas*, Seb Hil S-L Akl Tao *bugas*, Ilk Bkl *bagas*, Png *bĕlas*, Pmp *abyas*, Ibg *baggaq*, Sbl *buyah*, Apa *baggat*, Btk *fakas*, Cuy *bĕgas*, Nbl *bekas*, P-Neg *buya* 'hulled rice,' Mar *begas* 'boiled rice.'

2. 1. 4. Eng 'tooth': Tag *ngipin*, *ipin*, Seb Hil Bkl S-L Akl *ngipon*, Ilk *ngipĕn*, Png Sbl Mar *ngipĕn*, Pmp *ipan*, Ibg *ngipan*, Ivt *ñipĕn*, Cuy *ipĕn*, Tao *ipun* 'tooth.'

2. 1. 5. Eng 'buy': Tag *bili*, Ibg *balli*, Tao *bü* 'buy,' Png *bili*, Sbl *m-abli* 'expensive,' Pmp *abli* 'payment for merchandise.'

2. 1. 6. Eng 'hair (head)': Tag Seb Hil Bkl *buhok*, Ilk *buqok*, S-L *bohok*, Png *bwĕk*, Pmp *bwak*, Ibg *vuq*, Mgd *buhuk*, *buk*, Ivt *book*, *buk*, Btk *book*, Apa *abok*, Tao *buhuk*, Mar *bok*, Tir *ebuk* 'hair (head).'

2. 1. 7. Eng 'redeem': Tag Seb Hil Bkl S-L Akl Sbl *tubos*, Ilk *subbot*, Png *sĕbĕt*, Pmp *atbus*, Ibg *tabbuq, tavvuq*, Ivt *tuvvut-an*, Bgb *tubbos* 'redeem.'

2. 1. 8. Eng 'new': Tag Mgd *bago*, Seb Hil Akl S-L *bagqo*, Bkl *baqgo*, Png *balo*, Pmp *bayu*, Ibg *bagu*, Sbl *bayo*, Btk *falo*, Ivt *bayu, vago*, Cuy *bagoq*, Nbl *bado* 'new,' Ilk *bago* 'newcomer,' *baro* 'new, unmarried man,' Apa *bago* 'unmarried man,' *baro* 'recent,' *bagbago* 'new,' Ilt *bego* 'fresh,' Tao *bago* 'new, strange,' *baqgu* 'new, young, fresh, recent,' Mar *bago* 'new, raw.'

2. 1. 9. Eng 'sinew, vein': Tag Seb Hil Bkl S-L Tao Mar *ugat*, Ilk *urat*, Png Nbl *ulat*, Pmp Sbl Ivt *uyat*, Ibg *ugaq*, Apa *uhat*, Mgd *ugad* 'sinew, vein,' Ilt *ulat* 'sinew, vein,' *olat* 'elastic, gut.'

2. 1. 10. Eng 'flow off': Tag S-L Png Pmp Mgd Mar *agos*, Ilk *agos, ayos*, Sbl *agoh*, Nbl *ayus* 'flow off,' Ilt *ayut* 'current, stream,' Ivt *ayus* 'river-bed,' Tao *haus* 'the strong current in the Sibutu Passage during the northeast monsoon.'

2. 1. 11. Eng 'ford, wade': Tag Pmp *alog* 'ford, wade,' Ilk Png *alog* 'lowland, small canal with standing water,' Apa *man-alog* 'swim,' Mgd *alug* 'valley,' Mar *alug* 'valley,' *aror* 'raft.'

2. 1. 12. Eng 'night': Tag Apa Mgd *gabi*, Seb Akl *gabiqi*, Hil S-L *gabqi*, Ilk *rabii*, Bkl *banggi*, Png *labi*, Pmp *bengi*, Ibg *gabi, gavi*, Sbl *yabi*, Btk *lafi*, Cuy *gabiq* 'night,' Nbl *k-albi-an*, Tao *ka-abiqi* 'last night,' Mar *gawii* 'day,' *ma-gabi* 'afternoon.'

2. 1. 13. Eng 'house': Tag *bahay*, Seb Hil Ilk S-L Cuy *balay*, Pmp Sbl *bale*, Akl *baLay*, Ibg *balay, bale*, Mgd *walay, baay*, Ivt *bahay, vahay*, Tao *baay* 'house,' Bkl *baloy* 'house,' *balay-an* 'hut, lean-to,' Png *balĕy* 'village,' Btk *baqey* 'house,' *fali* 'basket,' Nbl *baley, faley* 'house, village,' P-Neg *bali* 'frame of a house,' Mar *walay* 'house,' *mbalay* 'build.'

2. 1. 14. Eng 'water':

(a) Ilk Png Pmp Ibg Ivt Btk *danum*, Sbl *lanum*, Nbl *cha-num*, P-Neg *lanIm*, Apa *danon*, Ilt *denom* 'water';

(b) Tag Seb Hil Bkl S-L Tao *tubig*, Akl *tubiq* 'water.'

2. 1. 15. Eng 'there is':

(a) Ilk Apa *adda*, Png *wala*, Ibg *uwad*, Bkl *igwa*, Ivt *ara*, Btk *wadaq, waday, waay,* Ilt *wadey,* Mar *aden* 'there is';

(b) Tag Seb Hil *walaq*, Bkl *waraq*, S-L *waraq, waray,* Pmp *alaq*, Ivt *ara-va*, Nbl *gwara*, Mgd *ada*, Tao *waay*, Mar *da* 'none';

(c) Ilk *awan*, Apa *awan, attan,* Ilt *awan, wan* 'none.'

2. 2. *Morphology.* The Philippine languages have perhaps the most complicated morphology in MP. I give below only a few representative examples, excluding the more complex formations. The meanings of the affixes are not always comparable in the different languages, and the glosses in Eng. are only approximate to avoid circumlocution. For types of morphological formation in IN, MN, and PN languages, see section 3. 2. 2.

2. 2. 1. <-an>: Tag *asin-an* 'salt-bed,' *tubu-han* 'sugarcane field,' Seb *asin-an* 'salt-bed,' Ilk *giling-an* 'grinder,' Bkl *inum-an* 'drinking-place,' Png *aral-an* 'place of study,' *an-dĕkĕt* 'a black object,' Pmp *inum-an* 'drinking-water.'

2. 2. 2. <-in>: Tag *inum-in* 'drinking-water,' Seb *tahi-on* 'what is to be sewn,' Bkl *inum-on* 'drinking-water,' Png *in-ka-ugali* 'state of possessing a habit.'

2. 2. 3. <-in>: Tag *s-in-aing* 'boiled rice,' Seb *g-in-aling* 'ground grain,' Ilk *m-in-ata* 'woven sawali with interspaces,' *in-apuy* 'boiled rice.'

2. 2. 4. <mag->: Tag *mag-ama* 'father and child,' Bkl *mag-ama* 'father and child,' Pmp *mag-aral* 'study.'

2. 2. 5. <pa->: Tag *pa-tabaq* 'fertilizer,' Seb *pa-init* 'snack, esp. taken with hot cocoa,' Ilk *pa-suli* 'corner post.'

2. 2. 6. <pag->: Tag *pag-awit* 'singing,' Seb *pag-awit* 'singing,' Ilk *pag-surat* 'what is used for writing,' Bkl *pag-inom* 'drinking.'

2. 2. 7. <pang->: Tag *pam-butas* 'borer,' Seb *panahiq* 'frequent sewing,' Ilk *pamarang* 'front teeth,' Bkl *pang-tanom* 'planting season.'

2. 2. 8. <pagka->: Tag *pagka-tao* 'human nature,' Seb *pagka-bataq* 'childhood,' Bkl *pagka-tawo* 'human nature.'

2. 2. 9. <paki->: Tag *paki-usap* 'request, entreaty,' Ilk *paki-nakcm* 'determination.'

2. 2. 10. <tag->: Tag *tag-lamig* 'cold season,' Seb *tag-sulat* 'author,' Bkl *tag-sadiri* 'owner.'

2. 2. 11. <taga->: Tag *taga-Maynilaq* 'native of Manila,' Seb *taga-Sugbo* 'native of Sebu,' Ilk *taga-Lawag* 'native of Laoag,' Bkl *taga-Maynila* 'native of Manila.'

2. 2. 12. <ma->: Tag *ma-tapang* 'courageous,' Seb *mapintas* 'fierce,' Ilk *ma-pintas* 'beautiful,' Bkl *ma-isog* 'brave,' Png *ma-linis* 'clean.'

2. 2. 13. <maka->: Tag *maka-bayan* 'patriotic,' Seb *maka-hilo* 'poisonous,' *maka-anay* 'enough.'

2. 2. 14. <pala->: Tag *pala-sumbong* 'tattler,' Seb *pala-anak* 'prolific female,' Bkl *para-kaon* 'fond of eating,' Pmp *pala-sumbong* 'tattler.'

2. 2. 15. <um->: Tag *um-alis* 'go away,' *k-um-ain* 'eat,' Seb *mu-kaun* 'eat,' *d-um-ali* 'hurry,' Ilk *um-inom* 'drink,' *d-um-akkel* 'grown big,' Bkl *um-inum* 'drink,' Png *un-loob* 'enter,' Pmp *m-urung* 'shrink, move back,' *t-um-ua* 'grow old.'

2. 2. 16. <mag->: Tag *mag-aral* 'study,' Seb *mag-tanum* 'plant,' Ilk *ag-adal* 'study,' Bkl *mag-harong* 'build a house,' Png *man-aral* 'study,' Pmp *mag-lulan* 'put into.'

2. 2. 17. <-an>: Tag *sulat-an* 'write on, write to,' Seb *(pag) hunong-an* 'stop,' Ilk *surat-an* 'write on, write to,' Bkl *lunad-an* 'ride,' Pmp *putut-an* 'be cut.'

2. 2. 18. <i->: Tag *i-tanim* 'be planted,' Seb *i-palo* 'whip,' Ilk *i-raman* 'include,' Bkl *i-putol* 'cut with,' Png *i-yatol* (*i-atol*) 'keep,' Pmp *i-lutuq* 'be cooked."

2. 2. 19. <mang->: Tag *mang-isdaq* 'go fishing,' *mamaril* 'go hunting,' Ilk *manganop* 'hunt with dogs,' Pmp *mamaril* 'go hunting.'

2. 2. 20. <maka->: Tag *maka-kain* 'be able to eat,' Seb *maka-abut* 'overtake,' Ilk *maka-balin* 'capable of doing,' Bkl maka-kaon 'eat accidentally,' Png *maka-dait* 'be able to sew,' Pmp *maka-sulat* 'be able to write.'

2. 2. 21. <maki->: Tag *maki-sakay* 'request to ride with,' Seb *maki-limos* 'ask for alms,' Ilk *maki-basa* 'read with,' Bkl *maki-inom* 'request a drink,' Pmp *maki-sake* 'request to ride with.'

2. 3. *Syntax.* The principal syntactic relations in Philippine languages are predication, attribution, and serial relation. A study of the examples given in Appendix B reveals the following characteristics.

2. 3. 1. In predication there is a S and a P. The favorite word-order in predication is PS.

2. 3. 2. The S has a marker and so does the P, the latter in some languages only in the word-order SP. In these languages the P-marker drops off in the order PS.

2. 3. 3. The S-marker: Tag Seb Hil *ang*, Ilk *ti*, Bkl S-L *an*, Png *say* — *so*, Pmp *ing*, Ibg *i*, Sbl *hay*, Ivt *nu*, Tao *in*.

2. 3. 4. The P-marker: Tag *ay* — *y*, Ibg Sbl *ay*, Ivt *am*.

2. 3. 5. Attribution is of four types: conjunctive, disjunctive, local, and absolute.

2. 3. 6. The marker for conjunctive attribution: Tag Bkl *na*— *-ng*, Seb Hil S-L *nga*— *-ng*, Ilk *a*, Png *ya* — *-n*, Pmp *-ng*, Ibg *nga*, Sbl *a* — *-n*, Ivt *a*, Tao O. The meaning is similar to opposition, modification, or relative construction in Eng depending on the constituents.

2. 3. 7. The marker for disjunctive attribution: Tag *nang*, Seb *sa*, Hil *sang*, Ilk *ti*, Bkl *kan*, S-L *han*, Png Ibg *na*, Pmp *ning*, Sbl *nin*, Ivt *ni*, Tao *sing*. The meaning is similar to modifier-modified, possessive, or direct object construction in Eng depending on the constituents.

2. 3. 8. The marker for local attribution: Tag Seb Hil Bkl *sa*, Ilk *ti idyay*, S-L *ha* — *na*, Png *ĕd* — *-d* Pmp *keng* — *king*,

Ibg *ta*, Sbl Tao *ha*, Ivt *du*. The meaning is relation to place 'in, on, to, from, etc.'

2. 3. 9. The marker for absolute attribution: Only Ivt has the marker *a*, the rest of the languages O. The absolute attribute modifies the following constituents, and the meaning depends on the constituents.

2. 3. 10. The marker for serial relation: Tag *at — t*, Seb *ug — -g*, Hil *kag*, Ilk *ken*, Bkl *saka — asin*, S-L *ngan*, Png *tan*, Pmp *at — ampong*, Ibg *anni*, Sbl *tan*, Ivt *kani — kanu*, Tao *iban*. These markers join two or more words or constructions which stand in coordinate relation one to the other(s).

2. 3. 11. Predication, attribution, and serial relation, together with the active-passive dichotomy (which I have no time to discuss, but will illustrate by examples), constitute the core, the central nervous system, as it were, of the syntax of the Philippine languages.

2. 4. *Discussion.*

2. 4. 1. *Qualitative.* If the words given in section 2. 1., above, are compared, the following characteristics of sounds in the different languages emerge.

All languages have *mata* for 'eye' (1).

Different V's occur: in Tag *i*, in Seb Hil Bkl S-L Akl Mgd Tao *u* (*o*), in Png Sbl Cuy Mar *ĕ*, in Pmp *a* (2), and Ibg also *a* (3).

Where other languages have *-p*, this is reduced to *-q* in Ibg (2). This reduction of final C's occurs often in Ibg (3, 6, 7, 9).

C's vary in Tag Ilk Seb Hil.Bkl S-L Akl Cuy Tao Mar *g*, Png *l*, Pmp Sbl P-Neg *y*, Btk Nbl *k* (3). The C is also *g* in Ibg Apa, but it is long C symbolized here by two identical C's in succession. Long C's are found in the languages of northern Luzon, including Ilk (7, 15a). Where the other languages have the sequence bV-, that is, CV-, in Pmp bV- interchanges positions to Vb-; this is a characteristic of Pmp (5, 7). Sbl has *h* for *s* of other languages because this dialect (Iba, Sambales) has

no *s-* sound. Where *s* occurs in Sbl, it is very likely a borrowing (7). The *f-* sound occurs in Btk (3, 8, 13); Btk Nbl *k* is to be explained away as the unvoiced counterpart of *g.*

Some languages have *ng-* where Ivt has *ñ,* a sound found only in this language so far. The difference between Tag *ngipin* and *ipin* is dialectic (4).

Like Ibg, Bgb also has long C's (7). Sbl *a* for the expected *ĕ* with the CV- interchanged to VC- is a probable borrowing from Pmp (5). The sequence -V¹*l*V¹- in other languages, that is, *l* occuring between two identical V's in mid-position, drops the *l* in Tao, and the two V's brought together as a result of this loss coalesce into one long V symbolized here by two identical V's in succession. This phenomenon is characteristic of Tao.

Where *h* occurs between two similar V's, it may either be reduced to glottal stop, as in Ilk, or lost, as in Ivt Btk, and the V's are fused into a single V, as in Ibg Mgd Ivt Mar. The *v*-sound occurs not only in Btk, alongside the *f*-sound, but also in Ibg Ivt (7). After the loss of *h* in Png Pmp and the expected vocalic change in Png *ĕ,* Pmp *a* (as in 2), accelerated articulation reduced the disyllabic into monosyllabic word, Png *-uĕ-* becoming *-wĕ-* Pmp *-ua-* becoming *-wa-* (6). Apa Tir show CV- interchange to VC-, like Pmp. This interchange of positions occurs not only between C and V but also between C's in different distribution. For instance, where Tag and other languages have *t-* . . . *-s,* Ilk Png have *s-* . . . *-t* (7). Intermediate between this interchange of C's in terminal positions and that of CV- to VC- in adjacent positions, as in Pmp (5, 6, 7), there are two other types of C interchange, one between two adjacent C's in medial position, as in Seb Hil S-L Akl Mgd Tao *adlaw* versus Ilk Bkl Apa *aldaw,* Pmp *aldo* (-*aw* becoming a simple V *o*) 'sun,' and the other between two succeeding C's separated by a single V, as in Seb Hil Bkl S-L Tao *gatus* 'hundred,' Pmp *gatus* 'hundred thousand' versus Ilk *gasut* 'hundred.' These different types of interchange of position are found in all the Philippine languages.

Phonological (and morphological) variants with cognate meanings are illustrated in 8-12. Where Seb Hil S-L Akl have

the sequences -*gq*-, Bkl Tao have -*qg*- (8). This contrast is very noticeable between the Bisayan languages on the one side and Bkl on the other. The variations *g-l-r-y-*(*h*) may be further pursued as probable witnesses to the theory that they could have been innovations of more than one *r*-sound in the proto-language. Bkl *banggi*, Pmp *bengi* (12) are reflexes of a doublet in the proto-language (see section 3. 2. 1. item 12b). Bkl *banggi* has been erroneously identified as reflex of the cognate words in the other languages for 'night.'

The reduction of a diphthong to a simple V is not an exclusive characteristic of Pmp, but is shared by Ibg Sbl Btk Ilt P-Neg (13). A type of *l*-sound has been found to occur only in Akl. A homorganic nasal in initial position in verb-like words is a morphological characteristic of Mar (and perhaps of other languages in Mindanao). (*Cf.* Fij *mbongi* 'night,' in section 3. 2. 1., item 12b.)

There are found in Philippine languages phonological doublets where one occurs only in a group of geographically contiguous languages and the other in another group, as the doublets for 'water' (14). Another example of a doublet is for 'blood,' one group occurring in Ilk *dara*, Png *dala*, Pmp P-Neg *daya*, Ibg *daga*, Ivt *rayaq*, Nbl *chala*, and the other in Tag Seb Hil Bkl S-L Akl Cuy *dugoq*, Mgd Tao *duguq*, Mar *rogo*. (*Cf.* the C variation in 3.)

Some phonological cognates have opposite meanings (15a, b), which might have been the cause of the creation of new words to accommodate the opposite meaning which shifted to the other cognate (15c). Velarized *g*- (symbolized by *gw*-) is characteristic of Nbl. (*Cf.* section 3. 2. 1. items 9, 15 Cha.)

I have not exhausted all the sound correspondences, only those which are unmistakable.

Since the words for the 15 meanings are undoubtedly cognates, their proto-forms can be reconstructed. The correct procedure would be to reconstruct the Proto-Philippine forms, test these with the other IN languages for possible revision, and label them Proto-IN, then test the Proto-IN with the MN languages again for possible revisions, and label them Proto-

MN, and finally repeat the procedure for PN languages and for Proto-PN. The final reconstructions would be labeled PMP. Another procedure would be to reconstruct a proto-language for each of the groups, IN, MN, and PN, and compare them among themselves for reconstructing PMP. I will not go into this laborious process now, but simply label the reconstructed Proto-Philippine forms PMP since they are applicable—with certain slight revisions—to all groups in MP. However, I adopted the latter procedure towards the end of section 3. 2. to avoid repetition of details which are tabulated in Appendix A.

It must be remembered that a reconstructed form is a postulated shape of the proto-language recovered in the light of correspondences established in the comparison of the "daughter" languages, merely as a point of reference. A reconstructed form is unpronounceable; it has no objective reality. These PMP forms are: 1 *matah* 'eye,' 2 *hatep* 'thatch (roofing),' 3 *beras* 'hulled rice,' 4 (*h*)*ipen* 'tooth,' 5 *belih* 'buy,' 6 *buhok* 'hair (body),' 7 *tebus* 'redeem,' 8 *b/ah/aRuh* 'new,' 9 *huRat* 'sinew, vein,' 10 *haRus* 'flow off,' 11 *haluR* 'ford, wade,' 12 (a) *Rabih* 'night,' (b) *beNih* 'night,' 13 *balay* 'house,' 14 (a) (*dD*)*anum* 'water,' (b) (*t*)*ubiR* 'water,' 15 *waDah* 'there is.'

To the PMP sources of the Philippine languages may be added some words of Skt origin which have been naturalized since the prehistoric period and for which PMP forms have been reconstructed. To cite only a few reflexes in selected languages (bold-face type indicates a long vowel): Tag *baroq*, Ibg *barwasi*, Mal *badyu* 'upper garment,' Skt *barasi* 'article of clothing,' PMP *bazuh* 'upper garment,' Ilk *bagi*, Ng-D *bagi* 'portion, share,' Skt *bhagi-* 'to divide, share,' PMP *bagih* 'portion, share,' Seb Apa *basa*, Sbl *baha* 'read,' Mal *bahasa* 'language,' Skt *bhasha* 'speech, language,' PMP *bacah* 'read,' Hil *gadyaq* 'elephant,' Bkl *gadya* 'dog,' Skt *gaja* 'elephant,' PMP *gazah* 'elephant.' These PMP naturalized Skt words are not reflected in MN and PN. In historic times, as early as the seventh century and thereafter, Chinese traders introduced numerous Chinese words, particularly terms for kinship and for cookery. Islamic culture, around the 15th

century, introduced, directly or indirectly, such terms as Tag Seb *alam*, S-L *aram* 'know (a fact),' Hil *alam* 'wisdom,' Ilk *alam* 'habit,' Bkl *alam* 'prudence,' *aram* 'know (a fact),' Pmp *alam* 'generosity,' Akl *ma-alam*, Cuy *ma-ĕlam* 'learned,' Arabic *'alamin* 'the worlds, the universe,' Hag Seb Hil Bkl S-L Png Pmp Akl Sbl Cuy Apa Ilt *salamat* 'thanks,' Tao *salaamat* 'peace, safety, word for greeting,' Arabic *salam* 'greeting,' *salamat* 'peace, safety, sometimes used as a word for greeting,' to mention only these two items. The overlays from Spanish and Eng are recent adaptations and are too numerous to list.

Worthy of mention are adaptations from non-Philippine languages of IN, from Mal, for instance, like the compounds *dalamhatiq* and *tanghaliq*. PMP *Dalem* 'interior, depth,' Tag *lalim*, Png *dalĕm*, Pmp *lalam*, Ivt *a-rahĕm* 'interior, depth,' Seb *lalum*, Hil *ha-lalum* 'deep,' Ilk *dalem* 'interior, depth, liver,' Bkl *rarum* 'interior, depth,' *dalum* 'deep,' S-L *dalum* 'dark, occult,' *ha-larum* 'deep,' Ibg *dalam*, Sbl *lalum*, Mgd *dalem* 'interior,' Mar *dalem* 'interior, depth,' Mal *dalam* 'interior, depth.' PMP *hatay* 'liver, mind, mood,' Tag Seb Hil S-L Akl Ivt Cuy *atay*, Bkl *atay, katoy*, Png Nbl *altĕy*, Pmp *ate*, Ibg *attay*, Mgd *atay, hatay*, Sbl Apa *agtay*, Btk *atey*, Ilt *agse*, Tao *atay, hatey, haqti* 'liver,' Mar *atay* 'liver, heart,' Mal *hati* 'liver.'

In *dalamhatiq*, lit. 'in the interior, in the depth, of the liver,' the two constituents *dalam* and *hatiq* do not occur in independent position, but structurally as a four-syllable word. The shift in meaning may be explained by the old cultural concept that courage resides in the liver, not in the heart (as understood by the modern man). Therefore, when something is 'in the interior of the liver' this organ is upset, thus causing 'extreme sorrow.' Final *q* is a Tag innovation, perhaps by analogy. (See *tanghaliq*, below.) Mal equivalent of '(extreme) sorrow' is *dukatjita*, in Bahasa Indonesia *duka*, but the compound occurs, as in *membatja dalamhati* 'to read to oneself.'

PMP *teNah* 'the middle': Seb S-L Akl *tungaq*, Cuy *tĕngaq* 'center,' Ibg *tangnga*, Mgd *tungaq* 'noon,' Mal *tĕngah* 'middle.' PMP *waRi* 'day, sun': Tao *haarii* 'day' in the compound *haarii raaya* 'holiday, feast day,' Mal *hari* 'day.' The reflexes for 'day, sun' in the majority of Philippine languages are of the

doublet PMP *ha(n)Daw*: Tag *araw*, Seb Hil S-L Akl Mgd *adlaw*, Bkl *aldaw* 'day, sun,' Ilk *aldaw* 'day' ('sun' is *init* in Ilk). In Mal 'sun' is *matahari*, lit. 'eye of the day.' The Mal model, then, of Tag *tanghaliq* is Mal *tĕngah hari* (lit. 'middle of the day') 'noon.' The *-q* in *tanghaliq* is analogous to that in *dalamhatiq*. (Without going into details, it may be mentioned here that in many adaptations from Malay, Mal *r* becomes Tag *l*.) The disyllabic reflexes of PMP *tĕNah* (there is none in Tag) becomes monosyllabic *tang-* in Tag. The two constituents do not occur independently, and the compound becomes a trisyllabic word.

Philippine languages, without doubt, must also have borrowed internally from one another (see comments on borrowings in Sbl earlier in this section, 2. 4. 1.). This is a large topic, but I will mention only one outstanding example in Tag. In *kalulwang bengeq*, which refers to 'souls making the rounds the night of All Saints' Day begging alms before the doors of Heaven are closed on them,' *bengeq*, which occurs only in this unique compound, is adapted from Pmp (see section 2. 1. 12.).

2. 4. 2. *Quantitative.* A mere inspection of the words compared in section 2. 1. reveals the similarities of the Philippine languages among themselves, as well as their diversities. But this inspection does not show the degree of relationship among these languages, not even between any two of them. With the methods of comparative linguistics, the sound changes in the contemporary languages that developed from the parent language, including the retentions, can be singled out, and the shape of the words in the parent language can be recovered and restored, as I have done in section 2. 1. But these methods do not permit the fixing of even the approximate date of separation of any two languages from the parent language. Greater accuracy than this is sought by anthropologists, historical linguists, and archeologists, who are interested in knowing the dates when linguistic changes took place.

Over a decade ago, a paper dealing with American Indian languages attempted to provide the more precise kind of dating that is needed. This new branch of linguistics is called lexi-

costatistics and has since acquired a considerable bibliography. Lexicostatistics, which is to historical linguistics as the Carbon-14 technique is to archeology, makes several assumptions: first, that some parts of the vocabulary of a language are less subject to change than other parts; second, that the rate of retention of meanings in the basic core of relatively stable vocabulary is constant through time; third, that the rate of loss of basic vocabularly meanings is approximately the same in all languages; fourth, that if the percentage of cognates within the core vocabulary is known for any pair of languages, the length of time that has elapsed since the two languages began to diverge from a single parent language can be computed.

In lexicostatistics words for certain non-cultural items (like sun, water, two, blood) are collected for the languages to be compared. The original test list consisted of 165 items, ultimately reduced to 100 preferred, or diagnostic, items. The word for an item in Language A is compared with the word in Language B for the same item, to see if the words are probable cognates. The number of probable cognates in the entire list is then converted to a percentage of cognates for use in the time-depth formula.

Time depth is computed by the formula $t = \log C / 2 \log r$. T stands for time depth in millenia, C stands for percentage of cognates, r for the "constant," that is, the percentage of cognates assumed to remain after a thousand years of diverging (usually .805), and log means "logarithm of" (so that log C means the logarithm of the percentage of probable cognates; 2 log r means twice the logarithm of the constant). The value of t, the time elapsed since the two languages began to diverge, may be changed to years by multiplying by 1,000. (I will not go into the formula which provides for the computation of the range of error, except to mention that there is such a formula.)

Lexiscostatistical data suggest the chronological order of development of languages and dialects. They also imply the location and cultural contact of ancient languages, those languages presumably being relatively homogeneous until the time when they began to diverge. However, it does not follow that

lexicostatistics can determine the language spoken by people responsible for artifacts found in a given place.

Lexicostatistics has to date been applied only on a limited scale. Through its application to Philippine and neighboring languages, the following corrections have been made of earlier subgroupings. Sangirese, spoken on islands between Celebes and the Philippines, had been classified as a Philippine language; yet while its highest percentage of cognates (39.9) is with Bisayan, this is not much different from the Sangirese percentage with Buginese in South Celebes (37.3). Similarly, Tontemboan in northern Celebes had also been regarded as a Philippine language, but its highest cognate percentage (29.8) is with Sangirese. Chamorro of Guam had also been regarded as belonging to the Philippine group, but its cognate percentage does not favor any Philippine language; its highest percentage (27.1) is with Maanjan of Borneo, followed by 23.9 per cent with West Futuna in PN. Chamorro's highest percentage with a Philippine language is 23.5 (with Agutaynon). Palau had also been considered a Philippine language, but its highest score (26.5 per cent) is with a Polynesian language in West Futuna, and its next highest (26.2) is with Sangirese. The highest cognate percentage of Palau with a Philippine language is 24.5 (with Buhid of Mindoro).

In the Philippines proper there are two distinct groups: a Central Philippine Group, including Tagalog, and a Northern Luzon Group, including Iloko, which converge in the neighborhood of 40 per cent. These two groups do not exhaust all the languages in the Philippines. Ilongot, for instance, whose highest score (28.4 per cent) is with Tagalog, may be presumed to belong to neither group.

Lexicostatistics is not in universal favor. The vocabulary items in the diagnostic list are considered too few to yield convincing results. It has been suggested that comparison of morphology and syntax should reinforce that of vocabulary.

Further, some meanings in the list are ambiguous. For instance, when a friend abroad asked me for some equivalents in Philippine languages, I ran into difficulties. To cite a few

cases, 'bark (of trees)' and 'skin' are both *balat* in Tag, a word which also means 'leather' (not in the list). 'Not' is matched by two meanings in Tag, the negative *hindiq* (*Hindi siya dumating* 'He/she did not arrive') and the prohibitive *huwag* (*Huwag kang pumasok* 'Don't enter'). 'Five' and 'hand' are matched by only one meaning in Png Ibg Btk Apa Mgd Tao Mar *lima*, Nbl *dima*. (In some MP languages where meanings are available in counting only up to 'two' or 'three,' 'five' will serve no purpose at all.) The meaning of 'he' in the majority of Philippine languages includes 'she' (not in the list) because there is no distinction of gender. And in Ilk the pronoun, third person sg in the unemphatic position of S is O (*Agsursurat* is a complete predication 'He/she is writing.')

According to a linguist who had been publishing extensively on statistical comparisons of different languages of the world long before the first paper on lexicostatistics was published, the techniques used in lexicostatistics have no mathematical basis and the studies of the past ten years have been illusory. Although I have not published my observations on the subject, my reaction to lexicostatistics, expressed in conversations and correspondence with other linguists, has been negative. Without summarily discounting its usefulness, I feel that lexicostatistics needs refinement if it is to find universal application.

3. *Similarities and diversities between the Philippine languages (as a group) and other groups of MP.*

3. 1. *Divisions of MP.* The MP languages occupy an area stretching from Madagascar in the west to Easter Island in the east, and from Formosa, Cham, and Hawaii in the north to Indonesia, New Zealand, and Polynesia in the south. The major groups and subgroups are the following (Capell 1962).

3. 1. 1. *IN.*

(1) *Western* — Sumatra, Java, Bali, South Borneo, the Malay Peninsula, Cham (in Cambodia), some less known languages on the Asian mainland, Li (Dai) on the island of Hainan, and Malagasy (in Madagascar).

(2) *Eastern* — Sumbawa, Timor, western New Guinea, Rotti, southeast Celebes, Alor (with reliable information still lacking).

(3) *Northern* — Philippines, the IN languages of Formosa, Palau, and Chamorro (in Guam) (which are in MC territory), northern and the entire west of Celebes, Illanun (in north Borneo), Sangir, Bantik, Bentenan, and others less known between the Philippines and Celebes.

To the west of Sumatra the languages of Nias, Batu (Sichule), Siberut, Mentawai, and Enggaño are supposed to stand rather apart from the western group. The languages in central Celebes form a particular subgroup. The languages of the Negrito groups surviving in the Philippines are assumed to be IN and do not constitute a separate Negrito family. However, it may be more correct to infer that after later IN immigrations had pushed them back to the mountain regions, where they broke up into small nomadic groups and lost communication with one another, the Negritos eventually lost their language(s). Today the Filipino Negrito speaks whatever Philippine language happens to prevail among the more general Filipino population of his district.

3. 1. 2. *MN*. Here a veritable population explosion has aroused considerable interest among investigators. Seven, or possibly eight, subgroups may be distinguished.

(1) *West New Guinea*—Along the coast of West Irian and sporadically along the north coast.

(2) *Eastern New Guinea* — In the area along the eastern border of West Irian; on the north coast of New Guinea; round the coast of the Papuan border; the MN languages of Papua along the south coast and offshore islands; in the Admiralty Islands; New Ireland and the adjacent islands; and in New Britain, where a definite Philippine element is noticeable, particularly the use of the infix *in* to form nouns from verbs, as *mat* 'die,' *m-in-at* 'death.'

(3) *Solomon Islands*.
(4) *Santa Cruz and Reef Islands*.

(5) *The New Hebrides.*

(6) *New Caledonia.*

(7) *Fiji Island.*

(8) *Rotuma Island,* 200 miles north of Fiji.

3. 1. 3. *PN.*

(1) *Western* — Tonga, Futuna, Uvea, Niue, Ellice and Tokelau Islands.

(2) *Eastern* — Hawaii, Marquesas, Cook Islands, Tahiti, Mangareva, Tuamotu, Easter Island, and the geographically misplaced Maori in New Zealand.

(3) *The Outliers* — These are within the geographic territory of MN and still await solution.

To the IN, MN, and PN groups is sometimes added a fourth with many fewer members than any of the three, and whose linguistic position has not been fully defined, namely the MC, which has been classified by some as a subgroup of MN.   Not all languages within the geographic boundary of MC belong to the group; for instance, Palau and Chamorro are definitely IN.   The best known of the "true" MC languages are Truk, Ponape, the Marshall Islands in the American Trust Territory, and Gilbertese in the British territory.   Some of these languages are influenced by both MN and PN.

3. 2. *Comparisons.*

3. 2. 1. *Phonology.* If we compare the cognates of the 15 items in section 2. 1. with the other MP languages, we obtain the following correspondences.

(1) PMP *matah* 'eye': IN — Jav Mal T-B, B-T Cha *mata*, Ng-D *mate*, Hov *masu* 'eye'; MN — Fij *mata-*, Sa'a *maa-* 'eye.'

(2) PMP *hatep* 'thatch (roofing)'; IN — Jav *atĕp,* Mal *atap,* Cha *atof* 'thatch (roofing)'; MN — Sa'a *s-ao* 'thatch (roofing).'

(3) PMP *beRas* 'hulled rice': IN — Jav *wos* = *beras*, Mal *bĕras,* Ng-D *behas,* Cha *pugas* 'hulled rice.'

(4) PMP (*h*)*ipen* 'tooth':   IN — T-B *ipon*, Hov *nifi*, Cha *nifin*, Tsuihwan *nipin* 'tooth'; MN — Sa'a *niho-* 'tooth'; PN — Ton Fut Sam *nifo* 'tooth.'

(5) PMP *belih* 'buy':   IN — Mal *běli*, T-B *boli* 'buy'; MN — Fij *voli-a*, Sa'a *holi* 'buy.'

(6) PMP *buhek* 'hair (head)':   IN — T-B *obuk* 'hair (head),' Jav *woq* 'beard,' PN — Ton *fuk-a*, Sam *fuqa* 'cut the hair.'

(7) PMP *tebus* 'redeem':   IN — Jav Mal *těbus*, Ng-D *tewus*, T-B *tobus* 'redress.'

(8) *PMP baRuh, b/ah/aRuh* 'new':   IN — Jav *wau = mau* 'recently,' Mal *baru*, T-B *im-baru*, Ng-D *bahu-a*, Hov *vo* 'new'; MN — Sa'a *hagalu* 'new.'

(9) PMP *huRat* 'sinew, vein':   IN — Jav Mal T-B *urat*, Ng-D *uhat*, Hov *uzatra*, Cha *g-ugad*, Favorlang *oggoch* 'sinew, vein'; MN — Fij *ua*, Sa'a *ule-ule* 'sinew, vein'; PN — Ton *uo-ua*, Fut Sam *ua* 'sinew, vein.'

(10) PMP *haRus* 'flow off':   IN — Jav *arus* 'current of water,' *asoh* 'flow off.'

(11) PMP *haluR* 'ford, wade':   IN — Jav *alur* 'marsh,' Mal *alur* 'water-course,' Ng-D *l-aloh-an* 'navigable water,' Hov *alu* 'pour out,' *aluz-ina* 'irrigation.'

(12) (a) PMP *Rabih* 'night':   IN — T-B *robi* 'for a long time'; MN — Fij *yak-avi*, Sa'a *seu-lehi* 'evening'; PN — Ton *ofi-ofi*, Fut Sam *afi-afi* 'evening.'

(b) PMP *beNih* 'night':   IN — Jav *běngi = wěngi*, T-B *bo-r-ngi-n*, Cha *puenge, poynge*, Lilisha *monggi* 'night'; MN — Fij *mbongi* (cf. Mar *mbalay* in section 2. 1. 13.), Sa'a *pongi* 'night'; PN — Ton Fut Sam *po* 'night,' Ton Sam *pongi-pongi* 'morning, twilight,' Fut Sam *pongi-a* 'benighted.'

(13) PMP *balay*, 'house':   IN — Jav *bale* 'bench,' Mal Ng-D *balay* 'hall,' T-B *bale* 'hut'; MN — Fij *vale* 'house,' Sa'a *hale* 'lean-to'; PN — Ton Fut Sam *fale* 'house.'

(14) (a) PMP *(dD)anum* 'water': IN — T-B *ranum*, Ng-D *danum*, Cha *hanum* 'water'; MN — Fij *ndranu* 'water'; PN — Ton *anu* 'wade through,' *anu-anu* 'bathe,' Fut *lanu* 'rinse,' *ma-anu* 'bathe,' Sam *lanu* 'wash off,' *lanum-ia* 'washed.'

(b) PMP *tubiR* 'water': IN — Mal *tubir* 'bank, shore.'

(15) PMP *waDah* 'there is': IN — Jav *ora* 'none,' Mal *ada* 'there is,' T-B *so-ada* 'none,' Cha *gwaha* 'there is' (*cf.* Nbl *gwara*, also with *gw-*, in section 2. 1. 15b.); PN — Fij *wara-i* 'none.'

Pending further investigation, I consider Ilk *awan*, Apa *awan*, *attan*, Ilt *awan*, *wan* 'none' of Philippine provenience.

The correspondences between PIN (=PMP) on the one side and PMN and PPN on the other can be summarized as follows. *First*: PMN and PPN behave in a similar manner towards PIN:

(a) PIN *a*, *i*, *u* are retained, PIN *e* becomes *o*.

(b) PIN *b* and *p*, *t* and *T*, *d* and *D*, *s*, *z*, *c* and *j*, and *k* and *g* fall together, unified.

(c) PIN *mb* and *mp*, *nt* and *nT*, *nd* and *nD*, *ñs*, *ñz*, *ñc* and *ñj*, and *Nk* and *Ng* also fall together.

(d) All PIN final C's disappeared in WB's, but reappear before suffixes. (See section 3. 2. 2.)

Second, the PN languages differ further from MN in the following:

(a) by still more inclusive unification of PIN *nt* and *nT* with *t* and *T*, of *Nk* and *Ng* with *g* and *k*, and *l* and *r* with *d* and *D*;

(b) by loss of PIN *R* and *y*.

See Appendix A for detailed MP correspondences.

3. 2. 2. *Morphology* (Lopez 1939). The Philippine languages share a common structure of word-formaton, that is, they are all agglutinative. By separating the formative elements, the affixes, WB's are left which occur as productive free

forms in utterances. WB's are generally disyllabic of the shape CV.CVC and CVC.CVC. Polysyllabic and monosyllabic words also occur, the latter mostly as syntactic markers. The stock of words is expanded in the following manner (limiting examples to Tag in IN; English glosses are not always exact):

(1) by compounding, *hampas/lupaq* 'vagabond';

(2) by doubling, *taun/taon* 'every year';

(3) by reduplication, *i/isa* 'only one';

(4) by affixation, *ma/buti* 'good,' *k/um/ain* 'ate,' *tawag/in* 'call';

(5) by complex formation, *mag/hampas/lupaq* 'be a vaga-bond,' *t/in/aun/taon* 'did, done every year,' *i/isa/isa/hin* 'will be reckoned one by one,' *nag/ma/ma/buti* 'ingratiating oneself.'

This list of expansions of the stock of words is not exhaustive, but will serve as basis of comparison with the other MP languages.

The stock of words in Jav Mal T-B Ng-D and Hov are similarly expanded.

In the MN languages, Fij is also agglutinative and WB's are generally disyllabic. There are no medial C-clusters, CVC. CVC words, or closed syllables. The stock of words is expanded by:

(1) compounding, *mata/ndravu* (lit. 'eye of the hearth') 'fire-place';

(2) doubling, *lawa/lawa* 'spider's web' (*lawa* 'net');

(3) affixation, *i/zuki* 'stick for digging,' *zukit/a* 'excavate';

(4) complex formation, *ma/uri/mu* (lit. 'your being alive') 'be blessed.'

Sa'a is likewise agglutinative, and WB's are generally disyllabic. Like Fij there are no medial C-clusters. The stock of words is expanded by:

(1) compounding, *mate/hulu* (lit. 'ripe-decayed') 'ultimate ripening';

(2) doubling, *hono/hono* 'closed';
(3) reduplication, *ho/hono* 'closed';

(4) affixation, *mal/hono* 'separated,' *honos/i* 'bet against somebody';

(5) complex formation, *hea/heat/i* 'arms.'

In the PN languages, Ton Fut Sam are likewise agglutinative, and most WB's are disyllabic. C- and V-clusters do not occur, and neither do closed syllables. Expansions are by:

(1) compounding, Ton Fut *maka/afi* 'fire-stone,' Sam *mata/liqi* (lit. 'small eyes') 'Pleiades';

(2) doubling, Ton *efi/afi*, Fut Sam *afi/afi*, 'evening';

(3) reduplication, Ton Fut *ki/kila*, Sam *qi/qila* 'glitter';
(4) affixation, Ton Fut *ma/kila* 'be shiny,' Sam *ta/ili* 'to laugh';

(5) complex formation, Ton *fe/haqit/aki* 'join with one another,' Fut *kafa/ma/ofa* 'to mount,' Sam *faqa/puput/a* 'blow in.'

3. 2. 3. *Syntax.* Syntactic comparisons are limited to Philippine languages (see Appendix B).

4. *Summary.* In this paper I have presented evidence of the similarities and diversities of the Philippine languages among themselves, evidence which points to an ultimate unity. The discussion above, which was detailed in matters of phonology, less so in morphology, and in syntax limited to Philippine languages, reflects quite well the extent and nature of work done in the MP family of languages. I gave random examples of external and internal adaptations in the Philippine languages as additional sources for PMP. The results of the comparisons of the Philippine languages as a group are compared with the other groups of MP to test correspondences in these groups. The division of the MP languages into groups and

subgroups is the result of the most recent studies on the subject.[3]

5. *Conclusion.* There are two conflicting theories of migration in the Pacific. From my exposition, on purely empirical evidence alone at the moment, it is clear that I favor the theory of the origins of the Philippine languages somewhere in Asia, in the neighborhood of IN. The opposite theory which holds that the dispersion was eastward, is attractive — because it is novel — as are the arguments presented in its favor. But its validity needs further proof. Let me elaborate.

5. 1. From the presence of words of Skt origin which were reconstructed in PMP and which have no reflexes in MN and PN several inferences can be drawn. These words must have been adapted in prehistoric times and thus support the theory which asserts that the home of the original IN culture was also the home of early Indian culture. That these words are not reflected in MN and PN would indicate that their bearers did not migrate beyond the confines of IN. If migration were eastward, it must have been recent and must have occurred at the time when the Skt words were already present in IN. It should be recalled that Skt is older than both Greek and Latin, with which it is related in the Indo-European family of languages.

5. 2. Many PMP sounds fell together in MN and PN. There are no studies which prove the contrary, namely, that the IN languages added sounds to MN and PN, if, indeed, the migration came from the east.

5. 3. Lexicostatistics fixes only the date when any two sister languages split from a common parent language; it does not fix the date of origin of the parent language. Lexicostatistics must go beyond this if its use as basis of the presumed dispersal from MN is not to remain speculative.

5. 4. Comparative studies of IN languages and those of mainland Southeast Asia with which relationship was not even suspected before have revealed some promising and en-

---

[3] However, the bibliography is complete only to 1963, when the paper was originally written.

couraging results. I do not know of any comparable studies made on the relationship between PN languages and the American Indian languages, say, of Central or South America. The pidginized languages of MN must be the result of intense cultural interpenetration of some duration between an earlier and more primitive culture and a later and more advanced one.

5. 5. To fix the original homeland of MP speakers beyond doubt — and to do the same for the Philippine languages — much work must be done. Archeologists, for artifacts and other evidences of material culture; systematic botanists, particularly for cultivated plants; systematic zoologists, particularly for domesticated animals; anthropologists in the broad sense, for family and spiritual life and social structure; paleontologists and geographers; and even meteorologists — all these specialists must join forces with linguists and undertake systematic studies, first on a small scale, but gradually expanding until the whole area is covered. The work may be slow and tedious, but the conclusions will be far less speculative than those we have today.

REFERENCES CITED

CAPELL, ARTHUR
   1962 Oceanic linguistics today.   Current Anthropology 3:371-428.

LOPEZ, CECILIO
   1939 Studies on Dempwolff's Vergleichende Lautlehre des austro-
   nesischen Wortschatzes.   Publications of the Institute of
   National Language, Bulletin 1:41, 50, 54-55.

Selected Sources

General

BOWRING, SIR JOHN
   A visit to the Philippine Islands.   London, Smith, Elder and Co.,
   1859.

CODRINGTON, R. H.
   The Melanesian languages.   Oxford, Clarendon Press, 1885.   With
   vocabularies, grammatical sketches and comparative grammar of
   some languages.

COLE, FAY-COOPER
   The peoples of Malaysia.   New York, D. Van Nostrand Co., Inc.,
   1945.

COSTENOBLE, H.
   Die Chamoro Sprache. s'Gravenhage, M. Nijhoff, 1940.

ECHOLS, JOHN M. and HASSAN SHADILY
   An Indonesian-English dictionary. Ithaca, Cornell University Press,
   1961.

GRACE, GEORGE W.
   The position of the Polynesian languages within the Austronesian
   (Malayo-Polynesian) language family.   Indiana University Publi-
   cations in Anthropology and Linguistics, Memoir 16, 1959.

KROEBER, A. L.
   Peoples of the Philippines.   New York, American Museum of Na-
   tural History Handbook Series 8, 1943.

MONIER-WILLIAMS, SIR
   A Sanskrit-English dictionary.   Oxford, Clarendon Press, 1899.

RAY, SIDNEY HERBERT
   A comparative study of the Melanesian Island languages.   Cam-
   bridge, Cambridge University Press, 1926.   With vocabularies and
   grammatical studies.

VAN RONKEL, PH. S.
   Maleis Woordenboek. Batavia, G. B. van Goor Zonen's Uitgeversmij,
   1943.

Divisions of Malayo-Polynesian and Lexicostatistics

CHRÉTIEN, C. DOUGLAS
The mathematical models of glottochronology. Language 38:11-37, 1962.
A classification of twenty-one Philippine languages. The Philippine Journal of Science 91:485-506, 1962.

DYEN, ISIDORE
Language distribution and migration theory. Language 32:611-626, 1956.

The lexicostatistical classification of the Malayo-Polynesian languages. Language 38:38-46, 1962.

THOMAS, DAVID AND ALAN HEALEY
Some Philippine language subgroupings: a lexicostatistical study. Anthropological Linguistics 4:22-33, 1962.

WURM, S. A.
The changing linguistic picture of New Guinea. Oceania 31:121-136, 1960.

Phonology

BENEDICT, PAUL K.
Thai, Kadai and Indonesian. American Anthropologist 44:576-601, 1942.

BRANDES, J. L. A.
Bijdrage tot de Vergelijkende Klankleer der Westerche Afdeeling van de Maleisch-Polynesische Taalfamilie. Utrecht, P. W. van de Weijer, 1884.

CONANT, CARLOS EVERETT
F and V in Philippine languages. Bureau of Science Division of Ethnology Publications 5:135-141, 1908.
The RGH law of Philippine languages. Journal of the American Oriental Society 31:70-85, 1910.
The pepet law of Philippine languages. Reprinted in Filipiniana, 2, Linguistics 126-168, 1938.
Indonesian I in Philippine languages. Journal of the American Oriental Society 36:181-196, 1916.

DEMPWOLFF, OTTO
Vergleichendes Lautlehre des austronesischen Wortschatzes. Beihefte zur Zeitschrift für Eingeborenen Sprachen, nos. 15, 17, 29. Berlin, 1934-1938.

KAHLO, GERHARD
Kleines vergleichendes Malayo-Polynesisches Wörterbuch. Leipzig, Otto Harvassowitz. 1941.

Lopez, Cecilio
   A comparative Philippine word-list. Ms.

Thomas, Dorothy M.
   Proto-Malayo-Polynesian reflexes in Rade, Jarai, and Chru. Studies in Linguistics 17:59-75, 1963.

Morphology

Kahler, Hans
   Untersuchungen zur Morphologie polynesischer Dialekte. Afrika und Ubersee 36:145-162, 1952; 37:35-48, 1952; 37:119-142, 1953; 38:73-88, 1954.

Lopez, Cecilio
   Studies on comparative Philippine morphology. Ms.

Syntax

Lopez, Cecilio
   Studies on comparative Philippine syntax. Ms. (Guggenheim project).

Appendix A

MALAYO-POLYNESIAN CORRESPONDENCES

[After Dempwolff as revised by Dyen, and with slight revisions by Lopez for purposes of this paper. Errors are not to be attributed to the first two.]

PMP	Tagalog	Toba-Batak	Javanese	Malay 3	Malay 2	Malay 1o	Malay 1c
a	a	a	a	ě	a	a	a
i	i	i	i, ě, e	ě	i, e	i	e
u	u, o	u	u, o, o	ě	u, o	u	o
e	i, (a¹, u¹)	o	ě	ě	ě	(?)	a
b	b	b-,-b-, -p	(b-, -b-), -b (w-, -w-)	b-		-b-	-p
d	d-, -r-, -d	d-,-d-, -t	d	d-		-d-	-t
D	d-, -l-, -d	d-, -d-,-r	d-, -d-, -d	d-		-d-	-r
z	d-, -r-	j	j			j	
g	g	g-, -g-,-k	g-, -g-, -q	g-		-g-	-q
j	-l-, -d	-g-,-k	r			-d-	-t
R	g	(r)	ø			ø	
hₓ, -'-, -'	h (-ø)	ø	ø			ø	
'-, -h-, -h	q	ø	ø-, -h-, -h			h	
y-, -y-, -ay	y-, -y-, -ay	(?)-,-ø-,-e	y-, -y-, -e	(?)-		-y-	-ay
-uy	oy	i	i			i	
k	k	h-, -h-, -k	k-, -k-, -q	k-		-k-	-q
c	s	s	c			c	
l	l	l	l			l	
r	l	r	r			r	
m	m	m	m			m	
n	n	n	n			n	
ñ, -ñ-	n	n	ñ			ñ	
N	ŋ	ŋ	ŋ			ŋ	
p	p	p	p			p	
T-, -T-	t	t	ţ			t	

¹By assimilation to a preceding or following vowel.

PMP	Tagalog	Toba-Batak	Javanese	Malay 3	Malay 2	Malay 1o	Malay 1c
s	s	s	s			s	
w-,-w-,-w	w-,-w-,-aw	∅ᴛ, -∅-, -o	w-, -w-, -o	h-		-w-	-aw
-mb-	-mb-	-mb-	-mb-			-mb-	
-nd-	-nd-	-nd-	-nd-			-nd-	
-nD-	-nd-	-nd-	-ṇḍ-			-nd-	
-nz-	-nd-	-nj-	-nj-			-nj-	
-Ng-	-ng-	-ng-	-ng-			-ng-	
-ñj-	(?)	-ng-	(?)			(?)	
-ŋk-	-ŋk-	-kk-	-ŋk-			-ŋk-	
-ñc-	-ns-	-ts-	-ñc-			-ñc-	
-mp-	-mp-	-pp-	-mp-			-mp-	
-nt-	-nt-	-tt-	-nt-			-nt-	
-nT-	-nt-	-tt-	-ṇṭ-			-nt-	
-ns-	-ns-	(?)	-ŋs-			(?)	

PMP	Ngaju-Dayak New Level 3	2	1o	1c	Ngaju-Dayak Old Level 2	1o	1c	Hova 3	2	1o	1c
a	X	a	ɔ	a	e			a	á	a, i	a
i	a	i, e	i	i, e				i	i	i	i
u	a	u, o	u	u, o				u	ú	u	u
e	X	a	(?)	a	e, e		e, e	e	é	(?)	i

PMP	Ngaju-Dayak			Old Level	Hova Before suffix	Cpds	Before X	Final
b	b-	-w-	-p		v-, -v-	-f+		
d	d-	-d-	-t	-r-	r-, -r-		-trX	
D	d-	-d-	-r	-r-	r-, -r-	-r+	-trX	
z-, -z-		j			z-, -z-			
g	g-	-g-	-k		h-, -h-	-h+	-kX	
-j-, -j		-r-	(?)	-r	-r-	-r+	-trX	
R		r		h	z-, -z-	-z+		-∅

PMP	Ngaju-Dayak							Hova				
	New level				Old level				Before suffix	Cpds	Before ǎ	Final
	3	2	1o	1c	2	1o	1c					
h-, -'-, -'			ø					ø				
'-, -h-, -h			h		ø			ø				
y-, -y-, =ay	(?)		-y-	-ay			-ey	-z-	-az+			-i
-uy				-oy								-u
k			k					h-, -h-	-h+	-k+	-kǎ	
c-, -c-			c			s		ts-, -ts-				
l			l					l-¹, -l-¹			-nǎ	
r			r		(?)	-h-	-h	r-, -r-	-r+		-trǎ	
m			m					m-, -m-	-m+	-m+	-nǎ	
n			ñ					n-, -n-	-n+	-n+	-nǎ	
ñ-, -ñ-			n					n-, -n-	-n+			
N			n					(?), -n-	-n+		-nǎ	
p			p					f-, -f-	-f+		-ø	
t			t					t-², -t-²	-t+²	-tr+	-trǎ	
T-, -T-			t					t-², -t-²				
s			s		h-	-(?)	-h	s-, -s-	-s+			-ø
w-, -w-, -aw	w-		-w-	-aw				v-, -v-				-u

PMP	Ngaju-Dayak	Normal		Reduced	
		1	2		
-mb-	nb	-mb-	-m-b	b-	-b-
-nd-	nd	-ndr-		tr-	-tr-
-nD-	nd	-ndr-	-n-dr-	tr-	-tr-
-ñz-	ñj	-ndz-	-n-z-	dz-	
-Nj-	ng	-ŋg-	-ŋ-g-	k-	-k-
-Ng-	nd	-ndr-			
(-NR-)				dz-	
-Nk-	ŋk	-ŋk-	-ŋ-k-	k-	-k-
-nc-	ñc	-nts-			
(-nl-)			-n-d-		
-mp-	mp	-mp-	-m-p-	p-	-p-

¹Becomes d before *i (asterisk indicates reconstruction).
²Becomes ts before *i.

PMP	Ngaju-Dayak	Normal 1	Normal 2	Reduced
-nt-	nt	-nt-	-n-t-	
-nT-	nt	-nt-		
-ns-	(?)	(?)	-n-ts	ts-

PMP	PMN[1]	Fiji		Sa'a	
a	a	a		a e[5] e[6]	
i	i	i		i	
u	u	u		u	
e	o	o		o	
-aw					
-ay	e	e		e	
-uy	i	i			
p, b	b	v		h	
t, T	t	t		ø	
d, D	d	r, -t+		r	
s, z, c, j	z	z		t, s[7]	
k, g	g	k		(q)	
m	m	m		m	
n, ṇ	ŋ	ŋ		ŋ	
N	n	n		n	
w	w	w[2]		w[2]	
y	j	z[3]		s[3]	
h	h	ø[4]		ø	
q	q				
l	l	l		l	
r	r	r		r	
mp, mb	mb	mb-	-mb-	p-[8]	-p-
nt, nT	nt	nd-	-nd-	dr-	-dr-

[1] Final consonants lost. [2] But disappears in the neighborhood of *u. [3] But disappears after *i. [4] But becomes y initially before *a. [5] If *i, *u was in the following syllable. [6] If *i, *u was in the preceding syllable. [7] Only before *i and *u. [8] Occasionally Sa'a q instead.

PMP	PMN	Fiji		Sa'a	
nd, ṇD	nd	ndr-	-ndr-	dr-	-dr-
ñs, ñz, ñc, ñj	nd	s-	-s-	dr-[9]	-dr-
Nk, Ng	ŋg	ŋg-	-ŋg-	k-	-k-

[9]tš before *1 and *u.

PMP	PPN[1]	Tonga	Futuna	Samoa
a	a	a, e[3]	a, e[3]	a, e[3]
i, -uy	i	i	i	i
u	u	u	u	u
e	o, e	o, e	o, e	o, e
-aw	o	o	o	o
-ay	e	e	e	e

PMP	PPN[1]	Tonga		Futuna		Samoa	
p, b	f	f		f		f	
t, T	t	t		t		t	
d, D, l, r	l	l		l		l	
s, z, c, j	s	h		s		s	
k, g	k	k		k		k	
m	m	m		m		m	
n, ñ	n	n		n		n	
N	N	ŋ		ŋ		ŋ	
w	w[2]	v		v		v	
y, R, h	∅						
q	q			q[4]		∅	
-mp-, -mb-	p	p-	-p-	p-	-p-	p-	-p-
-nt-, -nT-	t		-t-		-t-		-t-
-nd-	q			q-	-q-		
-nD-	q	∅-	-∅-	q-	-q-	∅-	-∅-
-ñs-, -ñz-, (-ñc-), -ñj-	h	h-	-h-	∅-	-∅-	∅-	-∅-
-Nk-, (-Ng-)	k		-k-				-q-

[1]Final consonants disappeared. [2]Disappeared in the neighborhood of *u.
[3]In the neighborhood of *i and *u. [4]Disappeared unaccountably in many words.

A. Predication: S and P markers (see section 2.3.3., 2.3.4.).

	(a)	(b)
Tag:	Ang batay bumabasa.	Bumabasa ang bataq.
Seb:	Ang bata nagábasa.	Nagábasa ang bataq.
Hil:	Ang bata nagabasa.	Nagabasa ang bataq.
Ilk:	Ti ubing agbasbasa.	Agbasbasa ti ubing.
Bkl:	An aki nagbabasa.	Nagbabasa an aki.
S-L:	An bata nagbabasaq.	Nagbabasa an bataq.
Png:	Say ugaw mánbabasa.	Mánbabasa so ugaw.
Pmp:	Ing anák bábasaya.	Bábasa ing anák.
Ibg:	I abbing mabbibbig.	Mabbibbig i abbing.
Sbl:	Hay anak ampamaha.	Ampamaha hay anak.
Ivt:	Maylir nu motdĕh.	Maylir nu motdĕh.
Tao:	In bata nagbabatsa.	Nagbabatsa in bataq.

But compare:

Tag:	Akoy sumusulat.	Sumusulat ako.
Ibg:	Soq ay matturaq.	Matturaq ako.
Sbl:	Hiko ay ampanulat.	· Ampanulat ako.
Ivt:	Yakĕn am maytulás.	Maytulás ako.

B. Attribution (see section 2.3.6.—2.3.9.)

	1. *Conjunctive*		2. *Disjunctive*
	(a)	(b)	
Tag:	bahay na bato	batong bahay	libro nang maestro
Seb:	balay nga bato	bato nga balay	libro sa maestro
Hil:	balay nga bato	bato nga balay	tunlonqan sang maestro
Ilk:	balay a bato	bato a balay	libro ti maestro
Bkl:	harong gapoq	gapong harong	libro kan maestro
S-L:	balay na bato	bato nga balay	libro han maestro
Png:	abung ya bato	baton abung	libro na maestro
Pmp:	bale batu	batung bale	libro ning maestro
Ibg:	balay nga batu	batu nga balay	libru nam maestro
Sbl:	bali a bato	balin bato	libro nin maestro
Ivt:	vahay a vato	vato a vahay	libro ni maestro
Tao:	bai batu	bai batu	buk sing guroq

3. *Local*

	(a)	(b)
Tag:	hangin sa bukid	galing sa Maynilaq
Seb:	hangin sa daruhan	gikan sa Manila
Hil:	hangin sa uma	naghalin sa Manila
Ilk:	angin ti taltalon	naggapu idyay Manila
Bkl:	duros sa uma	hali sa Manila
S-L:	hangin ha uma	tikang na Manila
Png:	dagĕm ĕd uma	nanlapurad Manila
Pmp:	angin keng taldawa	ibat king Manila
Ibg:	paddag ta koman	naggafu ta Manila
Sbl:	lĕpĕt ha buwat	ibat ha Ibali
Ivt:	salawsaw du takĕy	nawara du Manila
Tao:	hangin ha gimba	dain ha Manila

4. *Absolute*

Tag:	Ibig niyang magqaral nang kastilaq.
Seb:	Guston siyang magtuqon ug kinatsila.
Hil:	Luyag niya nagtuqon sang katsila.
Ilk:	Kayatna ti agqadal ti kastila.
Bkl:	Gusto niyang magqadal nin kastila.
S-L:	Karuyag niya magqaram hin kinatsila.
Png:	Labay toy manạral na kastila.
Pmp:	Bisa yang magaral kastila.
Ibg:	Gusto da i maddiddiamu tu gassila.
Sbl:	Labay nan magqaral nin kastila.
Ivt:	Makĕy a machinanaw su español.
Tao:	Mabayaq siya mangadyi kastilaq.

C. Serial relation (see section 2.3.10.)

	(a)	(b)
Tag:	si Unggoy at si Pagong	batat matandaq
Seb:	ang Unggoy ug ang Bao	batag tigulang
Hil:	si Amo kag si Bao	bata kag tigulang
Ilk:	ni Sunggu ken ni Pagqong	ubing ken baket
Bkl:	si Kabalang saka si Baoqo	aki asin lakay
S-L:	an Amoy ngan an Bako	bata ngan hin lagas
Png:	si Bakĕs tan si Bakukol	ogaw tan masi
Pmp:	i Matchin at i Pau	anak ampong matwa
Ibg:	si Ayong anni Dagga	abbing anna baku
Sbl:	hi bakĕh tan hi Pagqong	anak tan matontawo
Ivt:	sa Pachiĕng kani Irang	mutdĕh kanu malkĕm
Tao:	hi Amo iban hi Baoqo	bataq iban maas

D. Active-passive construction (see section 2.3.11.)

	1. *Active*	2. *Direct passive*
Tag:	Sumusulat **siya** nang nobela.	Sinulat **niya** ang nobela.
Seb:	Nagasulat **siya** sang estoria.	Sinulat **niya** ang estoria.
Hil:	Nagásulát siya sang nobela.	Ginsulat **niya** ang nobela.
Ilk:	Agsursurat ti nobela.	Sinuratna ti nobela.
Bkl:	Nagsurat **siya** nin nobela.	Sinurat **niya** an nobela.
S-L:	Nagsusurat **hiya** hin libro.	Ginsurát **niya** an libro.
Png:	Mansusulat na nobela.	Say nobela insulat **to.**
Pmp:	Súsulat **yang** nobela.	Sinulat **yang** nobela.
Ibg:	Natutturaq **yaya** ta nobela.	I nobela ay **niturraq na.**
Sbl:	Ampanulat **ya** nin nobela.	Yay nobela ay hinulat **na.**
Ivt:	Taytu **siya** a **maytulas** su nobela.	Nu nobela aya am **pina-tulas na.**
Tao:	Nagsusulat **siya** kitab.	Sinulat **niya** in kitab.

	3. *Instrumental passive*	4. *Local passive*
Tag:	Isinulat **niya** ang aking lapis.	Sinulatan **niya** ang kan-yang tatay.
Seb:	Gisulát **niya** ang akong lapis.	Gisulatán **niya** ang iyang amahan.
Hil:	Ginsulát **niya** ang akon lapis.	Ginsulatán **niya** ang iyang tatay.
Ilk:	Insuratna ti lapisko.	Sinuratanna ni tatangna.
Bkl:	Isinurat **niya** an sakong lapis.	Sinuratan **niya** an sa iyang amaq.
S-L:	Iginsurát **niya** an akon lapis.	Ginsuratán **niya** an iya tata.
Png:	Inpansulat **toy** lapis ko.	Nansulat ĕd ama to.
Pmp:	Pinyulat **ne** ing lapis ku.	Silatanan **ne** i tata na.
Ibg:	Natturaq **yaya** ta lapis ko.	Tinurattan **na** i amoq na.
Sbl:	Ipinanulat **nay** lapis ko.	Hinulatan **na** hi tatay na.
Ivt:	Pinaytulas **na** u lapis ko.	Naytulas **siya** di ama na.
Tao:	Nagsulat **siya** iban sin pinsil ko.	Sinulatan **niya** in amaq niya.

ABBREVIATIONS AND SYMBOLS

Akl	Aklan	MP	Malayo-Polynesia(n)
Apa	Apayao	Nbl	Nabaloi
Bgb	Bagobo	Ng-D	Ngadyu-Dayak
Bkl	Bikol	PIN	Proto-Indonesia(n)
B-T	Botel-Tobago	PMN	Proto-Melanesia(n)
Btk	Bontok	PMP	Proto-Malayo-
Cha	Chamorro		Polynesian
Cuy	Cuyunon	Pmp	Pampangan
Eng	English	PN	Polynesia (n)
Fij	Fiji	P-Neg	Pinatubo-Negrito
Fut	Futuna	Png	Pangasinan
Hil	Hiligaynon	PPN	Proto-Polynesia(n)
Hov	Hova	Sa'a	Sa'a
Ibg	Ibanag	Sam	Samoa
Ilk	Iloko	Sbl	Sambales
Ilt	Ilongot	Seb	Sebu
IN	Indonesia(n)	Skt	Sanskrit
Ivt	Ivatan	S-L	Samar-Leyte
Jav	Javanese	Tag	Tagalog
Mal	Malay	Tao	Taosug
Mar	Maranao	T-B	Toba-Batak
MC	Micronesia(n)	Tir	Tirurai
Mgd	Magindanao	Ton	Tonga
MN	Melanesia(n)		

ch	in Nbl, an affricate	—	alternates with
ĕ	the mid central un-rounded vowel	=	equals
I	in P-Neg. high mid central vowel	< >	enclose a class
		( )	enclose provisional reconstruction in PMP
L	in Akl, a lateral-velar	/ /	enclose formative elements in PMP
ñ	in Ivt, palatal nasal		
q	stands for glottal stop	-	separates affix(es) in 2.1, 2.2, 3.2
o	zero	/	separates affix(es) and compounds in 3.2.2
C	consonant		
V	vowel		
WB	word base		