

**THE PHILIPPINE LANGUAGE SCENE —  
THE SOCIOLINGUISTIC CONTRIBUTIONS  
OF HISTORICAL LINGUISTICS**

R. DAVID ZORC

*School of Australian Linguistics  
Darwin Community College*

**0. INTRODUCTORY COMMENTS**

I am pleased to have this opportunity to dedicate an article to Dr. Sibayan and to present some points of sociolinguistic relevance drawn from over fifteen years of work in the historical/comparative area. Many colleagues have commented how tedious and esoteric my work has seemed to be. The former, alas, is true, but the latter is not. Firm statements about the prehistory of the Filipino people and reasonable suggestions about the exact linguistic situation in the Philippines can be put forward as direct fruits of a comparativist's harvest. I trust that these will be of relevance and interest.

**1. LANGUAGES, DIALECTS, OR COMMUNILECTS?**

One popular view, often implicit in the minds of Filipinos, maintains that a language is a widespread and prestigious vehicle of communication (such as English, Chinese, Russian, or Pilipino), while any other kind of speech variety is 'a dialect'.<sup>1</sup> In contrast, sociolinguists determine a language on the basis of mutual intelligibility, whether total or chained (Hockett's L-simplex vs L-complex, respectively). Under this definition, *every speech variety is a DIALECT*, and the combination of all dialects that can communicate directly or indirectly with one another makes up a single LANGUAGE. Further refinements have been made, recognizing the speech of a single individual as an IDIOLECT, while that of a reasonably homogenous social group as a COMMUNILECT.

In terms of the Philippines, there are probably 50 million idiolects (based on a 1984 population estimate) broken up into approximately 500 communilects (based on the number of barrios, sitios, or barangays in non-metropolitan areas), i.e. where people talk in much the same way. While these numbers are very high (and hence not entirely informative), they are a matter of fact. Anyone who has journeyed from town to town within a purportedly common linguistic area (be it Bikol, Panay, or Mountain Province) can attest to the multiple differences in pronunciation, intonation, vocabulary and/or grammar prevailing. In all fairness to the speakers who so choose to identify themselves on the basis of even minor language variations, linguists and laymen alike should accept the communilect as the base-line definition of a Philippine speech variety. We therefore recognize, as do the speakers, a Kalibonhon vs Libakawnon Aklanon, an

<sup>1</sup>This view is compatible in many regards with the concept of communilect discussed below. In practice Filipinos are aware of even the most minute linguistic variations and label them accordingly (even if not always complementarily — statements such as 'They talk like birds' are sometimes heard).

Oas vs Polangui Bikol, an Ilianen vs. Livunganen Manobo, a Marinduque vs Batangas Tagalog, a Tina vs Botolan Sambal, an Amganad vs Kiangnan Ifugao, and so on.

The determination of the precise number of Philippine communilects (as opposed to the estimate of 5000 above) can be made by means of a survey of fifty words with a high probability of replacement.<sup>2</sup> Table 1 herein is derived from principles discussed a decade ago (Zorc 1974a) and virtually separates the Philippines into several thousand linguistic communities. For the purposes of this comparison, it is essential that forms be identical in sound, form, accent, and grammar — any difference whatsoever is important in establishing a communilect, and hence should be counted as minus. While historical linguists and lexicostatisticians are interested in cognates (forms descended from a single ancestral word or etymon), sociolinguists must concern themselves with differences separating speakers. Thus, Tag *maglaró* differs from Sina *mig-lalú* 'play' both in prefixation and phonology (r vs l), and each differs from Aln *ladó*, even if all three descend from an etymon \**ladú*. Tag (*um*)*akyat*, Kpm *mu-kyat*, Abr *apyat*, and Ivt *k(om)ayat* 'climb' again differ from one another even though they are ultimately cognate. Phonological differences (Kin *bedlay* :: Hil *budlay* 'difficult'; Akl *indi* :: Tag *hindi* 'not'; Rom *huyát* :: Akl *hu-lat* 'wait'), accent dissimilarities (Bon *o:tot* :: Png *otót* 'rat'), and semantic mismatches (Tag *do'ón* 'there-far' :: N-S *du'ún* 'there-near') need to be regarded as separators of communilects.

Table 1  
DIFFERENTIAL VOCABULARY SEPARATING PHILIPPINE COMMUNILECTS

1. akyat	'to climb (a tree)'	26. isá	'one'
2. alikabók	'dust'	27. itó	'this'
3. anó	'what?'	28. iyón	'that'
4. a:way	'to fight'	29. kailán	'when?'
5. ba:ba?	'chin'	30. kaunti?	'little (bit)'
6. bilís	'fast'	31. kanlu:ran	'west'
7. bu:kas	'tomorrow'	32. kapatíd	'sibling'
8. dagá	'rat'	33. ki:tid	'narrow' (= sikíp)
9. dali:ri'	'finger'	34. kuló	'to boil'
10. da:mi	'many'	35. lalamu:nan	'throat'
11. dibdib	'chest'	36. lamíg	'cold'
12. di:to	'here'	37. la:pit	'near'
13. do'ón	'there'	38. laró	'to play'
14. ga:lit	'anger'	39. líft	'small'
15. ha:gis	'to throw'	40. li:mot	'forget'
16. ha:pon	'afternoon'	41. lu:pa'	'earth/ground'
17. ha:wak	'to hold'	42. luwáng	'wide'
18. hindi?	'not'	43. mabu:ti	'good'
19. hintay	'to wait'	44. magsinunga:ling	'to lie'
20. hi:rap	'difficult'	45. mahi:na'	'weak'
21. hi:wa'	'to cut/slice'	46. masamá	'bad'
22. hiyá	'ashamed'	47. ta:pon	'throw away'
23. i:bon	'bird'	48. uma:ga	'morning'
24. ikli'	'short'	49. upó	'to sit'
25. i:log	'river'	50. walá	'none'

<sup>2</sup>See Dyen, James and Cole (1967) and Zorc (1974a). Conversely, it should be noted that these words have a very low probability of retention, such that very few of these meanings can be assigned an etymon or reconstruction with any reliability at the PAN or even PPH level.

If these criteria are strictly applied, only those speech varieties that score very high (in excess of 45/50) with one another can be regarded as belonging to the same communilect — and if the speakers consider themselves as such. In this way we can have sociological and linguistic confirmation of the Philippine language scene. Because of borrowing, common inheritance, and convergence (e.g., disparate shift of \*p > f, \*d > r, \*r > l, \*e > u, etc.), scores will rarely be 0/50. However, the list has been constructed on the basis of abundant data (Reid 1971, Yap 1977, McFarland 1977) such that it can be stated with confidence that scores will be very low, even between reasonably close genetic relatives.<sup>3</sup>

Although we now have a tool for determining the number of communilects in the Philippines (and adequate data are available in the files of the Summer Institute of Linguistics and many researchers), if we address the question of how many languages there are, numerous problems beset us. Since a language is defined in terms of mutual intelligibility, both the degree and the kind of intelligibility would need to be determined.<sup>4</sup> Some linguists would accept genetic intelligibility — if a Malay says *'Mata ku sakit'* ('My eye hurts'), and if a Filipino understands him (as most would be likely to), then obviously some communication is taking place. But the Malay may rattle on, and the doctor may respond, yet all of the rest of the dialog would well be lost on the Filipino. This is not practical intelligibility. The SIL has conducted extensive tests of intelligibility throughout the Philippines, and is therefore best equipped data- (if not time-)wise to assess the number of dialects and languages in the archipelago. Translators need to know the degree to which the Bible can be understood by speakers/readers in other communities. If too many barriers to understanding exist, a different translation is necessary. Hence, each speech variety is accorded its own dignity. Furthermore, linguistic imperialism is avoided — Warays may understand Cebuano, and Aklanon Hiligaynon, but each deserve their own intimate version of the Word.

If linguists could agree on a criterion for determining mutual intelligibility (and the SIL tests and scores are both accurate and sound in this regard), and factors such as bilingualism and sesquilingualism (when someone understands but cannot speak another language) could be controlled, then we would be well on our way to knowing how many dialects and languages there are in the Philippines. The exact answer could be known within this

<sup>3</sup>Preliminary counts made for Tag, Ceb, and Ilk indicate that Ilk has 31 unique forms, Tag 23, and Ceb 11 — the latter is due to Ceb's strong influence in the central and southern Philippines resulting in copious loans. These unique forms dictate that no other communilect could share a score higher than 19 with Ilk, 27 with Tag, or 39 with Ceb, except a communilect that was indeed Ilk, Tag, or Ceb respectively. Note, however, that in fact scores are substantially lower than the highest figures cited: Tagalog scores 8 with Sinauna, 7 with Kapampangan, 4 with Botolan, 3 with Bikol; Cebuano scores 25 with Hiligaynon (clearly inflated by borrowing), 18 with Samar-Leyte, 12 with Surigao; Ilokano scores 11 with Itneg, 8 with Kankanay dialects collectively, and 7 with Luba.

<sup>4</sup>For a more detailed discussion of the problems and issues involved see Zorc (1977:165-70).

decade, depending on research interests of M.A. or Ph.D. scholars and access to SIL files.<sup>5</sup> In the meantime, genetic linguistics can provide a solution.

## 2. THE HISTORICAL PICTURE – MAJOR PHILIPPINE SUBGROUPS

The number of languages in the Philippines has long been debated and estimated by linguists and laymen.<sup>6</sup> The question is, of course, a complex one, and the most reasonable solution that can be put forward is based on the concept of a linguistic subgroup as established in the discipline of historical/comparative linguistics. There are, at most, twenty-seven major linguistic groups (see Table 2) that can be described as 'Philippine' on the basis of geographic or genetic criteria. Note, however, that three are spoken in Celebes (Sulawesi, Indonesia), but can be proven to be immediately related to Southern Philippine languages. Another two families are spoken in Borneo (B1-2) and share features with Philippine and Sabahan languages,<sup>7</sup> while three (U1-3) are spoken in the Pacific and await definitive classification. One subgroup, Sama (I1), is clearly intrusive to the Philippines within the last millenium, and is genetically of an 'Indonesian type'. Within the political boundaries of the Philippines, then, there are at most eighteen language groups (N1-9+S1-9) which could share an immediate genetic ancestor, which have in all likelihood developed *in situ* over at least three thousand years, and which can not be attributed to multiple migrations as popular history suggests. As linguistic research progresses, these groups will probably be collapsed, but the current state of knowledge and debate dictates some prudence, so that the maximum number (18 Philippine + 1 Indonesian intrusive) represents a core of agreement among Philippinologists, amidst otherwise widespread disagreement as to the collapsability of these to ten<sup>8</sup> or two<sup>9</sup> or even one.<sup>10</sup>

<sup>5</sup>The Summer Institute of Linguistics has always been most generous and open with its data. However, since years of labor and research are involved, it would be most appropriate if an SIL team member drew up a comprehensive Philippine matrix of intelligibility test scores, possibly as part of his/her studies for a degree.

<sup>6</sup>Blumentritt (1901) recognized 194 native groups made available in the literature of his time. But he well knew that many of these were repetitious and/or inaccurate in several ways. Conklin (1952) was concerned with linguistic criteria and outlined 75 main groups broken up into a total of 156 members.

<sup>7</sup>Blust (1974) has proposed that these languages, while 'of the Philippine type', are more intimately related to the North Sarawak subgroup based on the sharing of the innovation involving strengthened reflexes of PAN \*b, \*d, \*f, \*g. By including them here, I do not take issue with his subgrouping hypothesis; I merely wish to indicate the counter hypothesis of scholars such as Charles (1974) or Prentice (1970). The innovations shared could be the result of borrowing (see 2.2 below).

<sup>8</sup>Ruhlen (in progress), following Reid for the most part, splits Western Malayo-Polynesian into 11 branches, ten of which concern us here: (1) Chamorro, (2) Palauan, (3) Yapese, (4) Northern Philippines, (5) Southern Philippines, (6) Meso-Philippine, (7) South Mindanao, (8) Celebes, (9) Borneo, (10) Sama-Bajaw.

<sup>9</sup>That is, a Proto Northern Philippine and a Proto Southern Philippine as the ancestors of most Ph languages respectively, excluding Sama.

<sup>10</sup>That is, a common Proto Philippine ancestor from which all Ph languages except Sama descended.

Table 2.

MAJOR PHILIPPINE SUBGROUPS	EXTRA-PHILIPPINE SUBGROUPS
N1 Iwatanic = Bashiic	C1 Minahasan
N2 North Cordilleran	C2 Mongondowic
1 South (Gaddang-Yogad)	C3 Gorontalic
2 North (Ibanag-Atta)	B1 Kadazan-Dusunic
3 Central	B2 Murutic
4 Agta	I1 Sama
N3 Dumagat = East Cordilleran	U1 Chamorro
1 Negrito	U2 Palau
2 Paranan	U3 Yapese
3 Central/Casiguran	
4 South/Umirey	
N4 Ilokano	
N5 Central Cordilleran	
1 South (Isinai)	
2 North (Itneg)	
3 East (Kalinga)	
4 Nuclear (Balangaw)	
5 Ifugaø	
6 Bontok-Kankanay	
N6 Ilongot	
N7 South Cordilleran	
N8 Southern Luzon	
1 Sambalic	
2 Sinauna	
3 Kapampangan	
N9 North Mangyan	
S1 South Mangyan	
S2 Palawanic	
1 North	
2 South	
S3 Kalamianic	
S4 Central Philippine	
1 Tagalog	
2 Bikol	
3 Bisayan	
4 Mansakan	
S5 Subanon	
S6 Manobo	
1 North	
2 Inland	
3 South	
S7 Danao	
S8 Bilic	
S9 Sangiric	

R. DAVID ZORC

2.1. THE ESTABLISHMENT OF PHILIPPINE SUBGROUPS

The major subgroups presented in Table 2 are the product of intensive research, most of which has taken place in the last decade:

- Bashiic = Ivatanic (Yamada 1965, 1966, 1973, and personal communications)
- North Cordilleran (Tharp 1974)
- Dumagat (Headland, personal communication)
- Central Cordilleran (Reid 1974)
- South Cordilleran (Reid and Zorc, personal communications)
- Mangyan (Zorc 1974b; Barbian 1977;<sup>11</sup> Pennoyer 1979)
- Palawanic (Thiessen 1981)
- Central Philippine (Zorc 1977, especially Bisayan; Bikol – MacFarland 1974; Mansakan/Northeast Mindanao – Gallman 1979)
- Manobo (Elkins 1974)
- Danao (Allison 1979; Fleischman 1981)
- Sangiric (Sneddon 1978)
- Mongondow (Charles, personal communications)
- Gorontalo (Little, personal communications)
- Sama (Pallesen 1977)
- Dusunic and Murutic (Prentice 1970; contrasted with Blust 1974)

An overview of Philippine linguistic reconstruction is contained in the work of Charles (1974) and Paz (1981), and the establishment of a single Philippine subgroup has been taken up by Chrétien (1962), Thomas and Healey (1962), Llamzon and Martin (1974), and Walton (1977). However, Reid (1981) has challenged these conclusions and the long-standing assumption that all Philippine languages form a single Austronesian node. Indeed, the rift between Northern and Southern Philippine languages is a big one, and some groups such as Bashiic and Bilic do not easily lend themselves to inclusion within NPh or SPh. A reasonably thorough coverage of both the facts and the disagreements linguists confront can be found in McFarland (1980) and in Ruhlen (in progress).

But these disputes should not be mis-construed. They indicate a vigorous interest on the part of many scholars in Philippine linguistics, and reflect a commitment to diverse theories drawn from an ever-growing corpus of data. Intuitive judgments have been put aside in favor of lexicostatistical, genetic, and grammatical criteria, and an overview of the relationships of most speech varieties has become possible, particularly in language blocs.

Hence, the groups presented in Table 2 represent the agreements of most scholars as to the maximum number of subgroups. It is in organizing these into higher-order subgroups (macro-subgrouping) that Philippinologists are in disagreement. For example, Tharp proposes that N2-4 form a single

<sup>11</sup>Barbian's study contains much valuable lexical data, but is unfortunately unsound from the genetic viewpoint. He defends a single Mangyan family (that includes Datagnon, which is demonstrably a West Bisayan dialect – Zorc 1972 and 1977) and therefore proves little more in his lexical counts than that Mangyan languages are genetically Philippine. Were Ilokano and Tagalog included in his study, by his own criteria and scoring, they would also be 'Mangyan'.