A PRELIMINARY GLOTTOCHRONOLOGY FOR NORTHERN LUZON[1]

ROBERT B. FOX, WILLIS E. SIBLEY, AND FRED EGGAN

Introduction: Linguists have long used similarities in vocabulary to order languages for preliminary purposes,[2] but the bulk of their attention has been properly directed towards the complex problems of form and structure and the establishment of historical relationships. Recently, however, their attention has been directed back to basic vocabulary comparisons through the researches of Morris Swadesh[3] and Robert B. Lees.[4]

Swadesh discovery of a relatively constant rate in vocabulary change was unexpected to him, and to other linguists, as well, since it was generally believed that some languages change much more slowly than others. Involved also in the discovery was the process of "stimulus diffusion" described by Kroeber. To quote Swadesh:

> The fact that fundamental vocabulary changes at a constant rate was discovered accidentally, but carbon dating was the specific stimulus for the research which brought it forth. For it was the achievements of radiocarbon dating which led the author four years ago [1948] to undertake studies of rate of vocabulary change. His only expectation at that time was to find an approximate maximum rate of change.[5]

In the meantime he had set out to order the dialects of Salish in Washington, Oregon, and British Columbia, on the basis of comparative vocabularies collected long ago by Boas. He worked out simple statistical constants (based on percentages of vocabulary

---

[1] A contribution from the Philippine Study Program recently established in Chicago with the aid of the Carnegie Corporation. Fred Eggan conceived the project and did the preliminary analysis; Fox and Sibley developed it to its present state and did all the basic comparisons and computations, the conclusions have been worked out jointly.

[2] The famous Powell classification of North American Indian languages was worked out essentially on this basis.

[3] Morris Swadesh, "Lexico-statistic Dating of Prehistoric Ethnic Contacts," *Proceedings* of the American Philosophical Society, Vol. 96, No. 4 1952, pp. 452-463; and "Salish Internal Relationships." *International Journal of American Linguistics*, Vol. 16, No. 4, 1950, pp. 157-167, are his basic papers.

[4] Robert B. Lees, "The Basis of Glottochronology," *Language*, Vol. 29, No. 2, 1953. pp. 113-127.

[5] Swadesh, M. "Lexico-statistic Dating of Prehistoric Ethic Contacts," pp. 453-54.

in common and using rates of change established for English) by
which the relative relationships of the dialects of Salish to one
another could be stated. The results agreed well with his observa-
tions on similarities in linguistic structure, and on the geograph-
ical distribution of the dialects.[6] This encouraged further research
on the rate of change, additional comparisons of various languages
and dialects, and studies of the statistical basis for the constants.

During the past months we have applied these techniques to
the interpretation of the complex and confused linguistic situation
in Northern Luzon. All Philippine languages belong to the Malayo-
Polynesian or Austronesian stock and to the Indonesian or the
western group. Basic similarities within the Philippines have led
Scheerer[7] and others to assume a "Philippine Group" of languages,
to which may possibly be added some of the North Borneo lang-
uages.[8] But beyond an enumeration and listing of major languages
and dialects in the Philippines, no systematic comparisons have
been made by which to classify them into groups and subgroups
on the basis of modern linguistic criteria.

In Northern Luzon Beyer[9] has recognized an "Iloko group," to
which Conklin[10] has recently added some possible members. Dyen's
studies in the central Philippines and Pittman's in Mindanao should
give us an ordering of these areas in the near future. Within the
"Iloko group" a number of attempts have been made to establish
further divisions, either on a phonetic basis or in terms of general
impressions. Thus Beyer early suggested an east-west division of
/f/ and /v/ in groups on the eastern side of the Cordillera Cen-
tral; and Barton placed pagan Ifugao, Bontok and Kankanay to-
gether with Christian Iloko in one group, and pagan Nabaloy and
Kalinga with Christian Pangasinan in another, noting that "tɬ
differences between languages of the same group is comparab

---

[6] Swadesh, M. "Salish Internal Relationships," pp. 157-67.

[7] Scheerer, Otto, "The Batán Dialect as a Member of the Philippine
Group of Languages," Publications of the Bureau of Science, Division of
Ethnology. Vol. X, Pt. I, Manila, 1908.

[8] Rutter, Owen, The Pagans of North Borneo, London, Hutchinson & Co.,
1929.

[9] Beyer, H. Otley, Population of the Philippine Islands in 1916, Manila,
1917.

[10] Conklin, H., Outline Gazetteer of Native Philippine Ethnic and Lin-
guistic Groups. Hectographed. 1952.

to that between Spanish and Italian, while languages of different groups compare about as Spanish and French."[11]

Before summarizing and discussing our own results, it will be useful to indicate our procedures and minor modifications of the Swadesh techniques. The details of the method and its statistical justifications can be found in the papers of Swadesh and Lees cited above.

Method and Procedure: This study had its origin in an interest in defining the position of Sagada in relation to Lepanto (Northern Kankanay) on the one hand, and the Bontok region on the other.   The availability of Father Vanoverbergh's comparative vocabularies[12] for a number of Northern Luzon languages led to preliminary experiments using his basic list, which included some 85 words from the Swadesh "test vocabulary."   Since these two series correlated very well, we went ahead with more detailed and comprehensive comparisons.

Seventeen languages from Central and Northern Luzon (see the accompanying Map for their distributions) were selected for comparisons of basic vocabulary.   Geographical and cultural criteria, in addition to availability of lexical lists, were primary factors in the selection of the languages to be compared.   We are particularly grateful for Father Vanoverbergh's vocabularies; without them the study would be fragmentary indeed.   The linguistic terms used in the comparison were obtained from the following sources:

Language (Specific Area)		Source
Tagalog		Vanoverbergh (1937)
		Conklin (1951)
Pampangan	(Apalit)	Fox (1949)
Sambal	(Botolan)	Fox (1947)
Pangasinan		Austria Macaraeg (1898)
		Cosgaya (1865)
Ibaloy	(Itogon)	Vanoverbergh (1937)
		Scheerer (1908)
Kankanay	(Bauco)	Vanoverbergh (1933, 1937)
Bontok	(Bontok)	Vanoverbergh (1937)
		Scheerer (1908)
Sagada		Eggan (1950)
		Vanoverbergh (1937)
Ifugao	(Kiangan)	Scheerer (1908)

[11] Barton, R. F., *The Half-way Sun.*   New York.   1930, pp. 29-30.
[12] Vanoverbergh, Morice, *Some Undescribed Languages of Northern Luzon*, Nijmegan, 1937.

Kalinga-1	(Lubuagan)	Vanoverbergh (1937)
Kalinga-2	(Balbalasan)	Richardson (1935)
Iloko		Vanoverbergh (1937)
		Scheerer (1908)
Tinguian	(Peñarrubia)	Eggan (1935)
Isneg	(Kabugao)	Vanoverbergh (1937)
		Fox (1946)
Ibanag		Vanoverbergh (1937)
		Scheerer (1908)
Gaddang	(Bayombong)	Fox (1946)
Ilongot	(Kadiyakan)	Vanoverbergh (1937)
	(Dibutunan)	Fox (1948)

A basic vocabulary of 197 items was finally compiled by the authors; of these 106 items appear also in the Swadesh list. The wordlist is reproduced below with the items included in the Swadesh "test vocabulary" indicated by asterisks. Some items in the Swadesh list do not occur in the Philippine environment or have ambiguous reference; others were not available in a sufficient number of wordlists to be useful for our purposes. Our preliminary experiments indicated that this did not cause a serious distortion, however.

*afraid (fear), *all, anus, *ashes, arrive, *back (of person), bad, bear (a child), beautiful, *belly, body, born, breast (of woman), bring (carry in hand), broken jars, *bury, cheap, chest, *child, *cloud, coat (jacket), *cold (chilly), companion, crooked, cry (weep), deer, *die, difficult, *dig, diligent, dirty, *dog, don't, don't like (dislike), *drink, *dry, dwell, *ear, *earth, easy, *eat, eel, *egg, *eight, eleven, expensive, *eye, face, *far, *farther, fence, finger, *fire, first, *fish, *five, *flower, follow, *foot, *forest (woods), forget, *four, friend, girl (maiden), *give, go, *good, *guts, *hair (head), *hand, *he (she, it), *head, *heart, *heavy, *here, high (tall), *hot (warm of weather), house, *how? (manner), how many?, *I, knee, *know (a fact), *leaf, *left (side), *leg, *liver, *laugh, like, *large (big), light (in weight), male, *man (person), *many, mat (sleeping), moon, *mother, *mountain, *mouth, my, *neck, *night, *nine, *no, *nose, now, *near, old man, old woman, *one, once, *other (different), our (excl.), our (incl.), penis, quick, *red, remember, rice (husked), rice field (wet), *river, *road (way), roof (of house), *root, run, *salt, *sand, *say, *sea, second, *seven, shameful (shame), *short, *sibling, *six, *sky, *sleep, *small, *snake, snap (as a rope), space (under the house), *spouse, *star, *stone, *straight, strong, *sun, sweet, *sun, sweet, *ten testicles, the (article of the common noun), *their, there (near addressee), *there (far), *they, *three, there is, there is none ((none), *this, tired, tomorrow, *tongue, *tooth, to them, to us (excl.) to you (pl.) ,town, tree (wood), *twenty, twice, *two, urinate, vulva, *walk, *water, *we (excl.), we (incl.), we (you and I), weak, *wet, *what?, *where?, *who?, why?, widow (widower), wild boar, *woman, *word (speech), *year, yes, yesterday, *you (sing.) you (pl.), young man, your (sg.)

Our procedure was to compare each of the 136 possible pairing of languages (e.g., Iloko-Isneg, Bontok-Sagada, Isneg-Ibanag, etc.) in terms of the available vocabulary items. Large tabular sheets were used for this analysis, the English equivalents being listed vertically and the language pairs horizontally. For each English entry, the Philippine language equivalents were compared for all 136 language pairs in order to discover if the terms in each pair were cognate or non-cognate (cognate terms also included those terms 'borrowed" from one language to another during the period of assumed separation).

Word comparisons were based on established sound correspondences, metathesis, assimilation, reduction and reduplication. The transcriptions in the wordlists were based upon phonetic principles; hence it was necessary to partially phonemicise the seventeen languages compared. This was accomplished rather easily for the consonants, but the phonemic pattern for vowels is not completely certain. We do not believe, however, that this has introduced any serious distortions.

With these operations completed, the comparisons were reduced to a table of percentages of cognate forms in common (see Table II). Here for each comparison, we indicated the number of items compared. This is important in evaluating the reliability of the results, since the probable error will vary with the number of items compared. Thus, the Gaddang comparisons are less reliable than others since the Gaddang list was considerably smaller.

For tentative dating of the time of separation of two languages, or dialects, we have utilized the statistical constants and formulae developed by Swadesh and Lees. Their calculations indicate that on the average a language will retain 81% of its basic vocabulary after 1,000 years. The changes which take place in languages are assumed to be independent; hence two languages separated for 1,000 years will have approximately 66% of their basic vocabulary in common (81% of 81%).

Utilizing this constant, and using 1940 for the "ethnological present" we have calculated the approximate time values for the assumed separation of the Northern Luzon languages one from another. We have not calculated the limits of error involved, since we hope to perfect our use of the method with larger samples, but Lees' calculations suggest a "probable error" of around 10% at a

time depth of 1,000 years. Table I gives the results of our comparisons expressed in terms of separation dates.

The authors of this paper are fully aware of the general and specific qualifications which need to be made before these figures can be utilized as actual dates. But in the relative absence of archeological chronology in Northern Luzon we think they are of the greatest value as working hypotheses for giving us a preliminary indication of our time scale for this important region. For those who prefer to utilize the percentages of vocabulary in common, it is also possible to draw far-reaching conclusions, as Swadesh's analysis of the Salish dialects indicates.

Preliminary Conclusions: We have time to indicate only our most general conclusions, since the study is still in progress.

(1) Our comparisons indicate that the great majority of the languages of Northern Luzon, with the possible exception of Ilongot, belong together in a single group. For this grouping we propose the name "Northern Luzon Type" in place of the earlier "Iloko Type." This latter term implies an ancestral status for Ilokano with reference to other related Northern Luzon languages, which is not justified by the evidence so far.

(2) Within the Northern Luzon Type, there are a number of subgroupings. Tentatively we propose the following groupings and languages for this type:

Northern Luzon Type

A. Northern Division
   (1) Iloko
   (2) Tinguian
   (3) Isneg
   (4) Ibanag
   (5) Gaddang

B. Central Division
   (1) Kankanay
   (2) Bontok
   (3) Kalinga
   (4) Ifugao

C. Southern Division
   (1) Ibaloy
   (2) Pangasinan

D. Southeastern Division
   (1) Ilongot

The majority of these languages share from 40-80% of their vocabulary, when compared one with another. The relationship of Ilongot to all other groups ranges below 40%, with the exception of Gaddang and Pangasinan, but in its structure Ilongot belongs with the Northern Luzon Type, as against the Central Luzon languages. Pangasinan likewise has its closest relationships with Ibaloy and Tinguian, a situation intelligible in terms of its earlier distribution before the Ilokano expansion, but it also has moderately close relationships to Sambal and Tagalog. Pampangan, on the other hand, is quite distant from the majority of the Mountain Province languages, but relatively close to Tagalog, and to a lesser extent, to Sambal.

(3) Within Northern Luzon these tables allow us to formulate more precise working hypotheses in our study of individual groups. That Bontok, for example, shares about 80% of its basic vocabulary with Kankanay, with an indicated time separation of less than 600 years, suggests strongly that these two groups were one linguistically around 1400 A.D. and presumably shared in a common culture. The Sagada relationships, somewhat erratic in character, are probably to be explained in terms of a relatively late settlement of the region by families from several areas. The relationship of Ilokano and Tinguian, postulated by Cole as being relatively close, is confirmed by 73% common vocabulary and an indicated date of separation somewhere around 1200 A. D.

(4) When parallel, or superior, information is available for other regions of the Philippines, the nature of the relationship of northern to central and southern languages can be more closely stated. Ultimately, it may be possible to get a working conception of the time of separation from the Asiatic mainland, utilizing Benedict's studies of Indochinese languages[13].

---

[13] Benedict, Paul, "Thai Kadai and Indonesian: A New Realignment in Southeast Asia," *American Anthropologist*, Volume 44, 1942.

BIBLIOGRAPHY

1. Cole, Fay-Cooper, *The Tinguian,* Chicago: Field Museum of Natural History, 1922.

2. Conklin, Harold, *Lexical List For Philippine Languages,* 1951 MS.

3. Cosgaya, Lorenzo F., *Diccionario Pangasinan-Español,* Manila: Collegio de Santo Tomas, 1865.

4. Dyen, Isidore, "The Tagalog Reflexes of Malayo-Polynesian D," *Language,* Vol. 23, 1947. pp. 227ff.

5. Eggan, Fred, *A Lowland Tinguian-Ilocano Grammar* based upon Waterman, *A Study of the Igorot Language as Spoken by the Bontoc Igorots,* 1935. MS.

6. Eggan, Fred, *A Sagada-English Vocabulary,* 1950. MS.

7. Elbert, Samuel H., "Internal Relationships of Polynesian Languages and Dialects," *Southwestern Journal of Anthropology,* Vol. 9, No. 2, 1953.

8. Fox, Robert, *A Brief English-Gaddang Vocabulary,* 1946, MS.

9. —————, *A Brief English-Isneg Vocabulary,* 1946, MS.

10. —————, *An English-Sambal Vocabulary,* based upon Conklin, *Lexical Check List for Philippine Languages,* 1948. MS.

11. —————, *An English-Ipági Egongot Vocabulary,* based upon Conklin, *Lexical Check List for Philippine Languages,* 1948. MS.

12. —————, *An English-Pampangan Vocabulary,* based upon Conklin, *Lexical Check List for Philippine Languages,* 1949. MS.

13. Institute of National Language, Commonwealth of the Philippines, *A National Language-English Vocabulary,* Manila: Bureau of Printing, 1945.

14. Lees, R. B., "A Method of Dating With Lexical Statistics," Michigan Linguistic Institute, Ann Arbor, Michigan. 1951, mimeo.

15. —————, "The Basis of Glottochronology," *Language,* Volume 29, No. 2, 1953.

16. Macaraeg, Austria Anastasio, *Vocabulario Castellano-Pangasinan,* Manila: Amigos del País, 1898.

17. Richardson, Arthur, *A Balbalusan-English Word List,* 1935, M.S.

18. Rutter, Owen, *The Pagans of North Borneo,* London; Hutchinson and Company, 1929.

19. Scheerer, Otto, "The Batán Dialect as a Member of the Philippine Group of Languages," Bureau of Science, Division of Ethnology Publications, Vol. V, Parts I and II, Manila: Bureau of Printing, 1908.

20. Swadesh, Morris, "Salish Internal Relationships," *International Journal of American Linguistics,* Vol. 16, No. 4, 1950.

21. —————, "Lexico-Statistic Dating of Prehistoric Ethnic Contacts, "*Proceedings of the American Philosophical Society,* Vol. 96, No. 4, 1952.

22. Vanoverbergh, Morice, *A Dictionary of Lepanto Igorot or Kankanay, As It Is Spoken At Bauco.* Verlag der Internationalen Zeitschrift "Anthropos" Modling bei Wien, St. Gabriel, Oesterreich, 1933.

23. —————, *Some Undescribed Languages of Luzon,* Nijmegen: Dekker and van de Vegt, N.V.:, 1937.

Table I

Tentative Separation Dates

for

Seventeen Languages in Central and Northern Luzon

	Tagalog	Pampangan	Sambal	Pangasinan	Ibaloy	Kankanay	Bontok	Sagada	Ifugao	Kalinga-1	Kalinga-2	Iloko	Tinguian	Isneg	Ibanag	Gaddang	Ilongot
Tagalog		755	450	205	282	264	445	840	252	290	122	226	132	109	111	59	267
Pampangan	755		356	99	362	441	613	691	172	172	240	152	402	109	59	115	324
Sambal	450	356		428	252	278	152	178	782	421	240	105	402	76	22	324	749
Pangasinan	205	99	428		579	219	125	47	74	170	180	511	229	450	111	59	82
Ibaloy	282	362	252	579		344	461	356	180	185	442	285	74	285	74	20	65
Kankanay	264	441	278	219	344		1406	1116	905	958	362	698	562	562	482	65	60
Bontok	445	613	152	125	461	1406		858	894	739	552	469	587	469	587	344	22
Sagada	840	691	178	47	356	1116	858		507	374	692	330	344	330	344	59	149
Ifugao	252	172	782	74	180	905	894	507		728	728	269	692	269	324	191	14
Kalinga-1	290	172	421	170	185	958	739	374	728		857	570	478	321	190	401	247
Kalinga-2	122	240	240	180	442	362	552	692	728	857		269	636	269	191	428	272
Iloko	226	152	105	511	285	698	469	330	269	570	269		787	478	269	425	579
Tinguian	132	402	402	229	74	562	587	344	692	478	636	787		908	324	579	421
Isneg	109	109	76	450	285	562	469	330	269	321	269	478	908		344	570	202
Ibanag	111	59	22	111	74	482	587	344	324	190	191	269	324	344		70	14
Gaddang	59	115	324	59	20	65	344	59	191	401	428	425	579	570	70		5
Ilongot	267	324	749	82	65	60	22	149	14	247	272	579	421	202	14	5	

* Underlined dates are B. C. All others are A. D.

	Tagalog	Pampangan	Sambal	Pangasinan	Ibaloy	Kankanay	Bontok	Sagada	Ifugao	Kalinga-1	Kalinga-2	Iloko	Tinguian	Isneg	Ibanag	Gaddang	Ilongot
Tagalog		189 .589	191 .524	189 .471	194 .381	197 .384	197 .355	167 .299	194 .386	176 .380	141 .454	197 .475	168 .452	197 .452	194 .438	126 .484	184 .332
Pampangan	189 .598		185 .503	180 .450	185 .368	188 .356	188 .330	160 .319	182 .346	165 .400	134 .388	189 .407	163 .362	187 .417	183 .410	115 .496	177 .311
Sambal	191 .524	185 .503		183 .519	186 .366	191 .382	191 .403	162 .307	183 .399	167 .359	133 .451	194 .433	164 .476	188 .420	185 .427	115 .513	176 .295
Pangasinan	189 .471	180 .450	183 .519		186 .554	190 .474	189 .455	159 .440	182 .445	166 .464	143 .538	189 .466	168 .524	186 .419	183 .407	117 .487	173 .416
Ibaloy	194 .381	185 .368	186 .366	186 .554		194 .500	194 .526	167 .503	189 .466	169 .467	136 .522	195 .487	164 .445	192 .427	190 .500	115 .496	183 .393
Kankanay	197 .384	188 .356	191 .382	190 .474	194 .500		197 .793	166 .699	191 .638	176 .653	141 .504	197 .583	169 .550	196 .531	192 .505	116 .483	186 .355
Bontok	197 .355	188 .330	191 .403	189 .455	194 .526	197 .793		166 .620	192 .635	175 .594	140 .543	197 .528	171 .556	196 .500	194 .438	117 .496	184 .375
Sagada	167 .299	160 .319	162 .307	159 .440	167 .503	166 .699	166 .620		162 .537	146 .507	122 .582	167 .497	139 .496	169 .444	161 .428	95 .432	158 .304
Ifugao	194 .386	182 .346	183 .399	182 .445	189 .466	191 .638	192 .635	162 .537		171 .591	134 .552	192 .484	164 .530	192 .495	186 .468	113 .513	178 .337
Kalinga-1	176 .380	165 .400	167 .359	166 .464	169 .467	176 .653	175 .594	146 .507	171 .591		120 .625	176 .568	148 .568	176 .523	171 .468	104 .519	162 .383
Kalinga-2	141 .454	134 .388	133 .451	143 .538	136 .522	141 .504	140 .543	122 .582	134 .552	120 .625		142 .606	122 .639	140 .493	137 .518	92 .554	128 .359
Iloko	197 .475	189 .407	194 .433	189 .466	195 .487	197 .583	197 .528	167 .497	192 .484	176 .568	142 .606		170 .729	197 .645	194 .588	116 .586	186 .344
Tinguian	168 .452	163 .362	164 .476	168 .524	164 .445	169 .550	171 .556	139 .496	164 .530	148 .568	122 .639	170 .729		169 .556	166 .512	116 .552	159 .377
Isneg	197 .452	187 .417	188 .420	186 .419	192 .427	196 .531	196 .500	169 .444	192 .495	176 .523	140 .493	197 .645	169 .556		191 .644	117 .598	183 .322
Ibanag	194 .438	183 .410	185 .427	183 .407	190 .500	192 .505	194 .438	161 .428	186 .468	171 .468	137 .518	194 .588	166 .512	191 .644		116 .698	183 .366
Gaddang	126 .484	115 .496	115 .513	117 .487	115 .496	116 .483	117 .496	95 .432	113 .513	104 .519	92 .554	116 .586	116 .552	117 .598	116 .698		112 .438
Ilongot	184 .332	177 .311	176 .295	173 .416	183 .393	186 .355	184 .375	158 .304	178 .337	162 .383	128 .359	186 .344	159 .377	183 .322	183 .366	112 .438	

Table II

Table of Size of Samples (upper line) and Cognate
Percentages (lower line) for Seventeen Languages in Central and Northern Luzon

As sub-groups of the major languages are common, the specific areas from which the lexical lists used in this study were obtained are indicated by triangles and the names of the communities (e.g. Sagada).

ISNEG

Kabugao ▲

ILOKO

IBANAG

TINGUIAN

Penarrubia ▲

KALINGA

▲ Balbalasan

▲ Lubuagan

BONTOK

▲ Bontok

▲ Sagada

GADDANG

▲ Bauko

Northern Kankanay

IFUGAO

KANKANAY

▲ Kiangan

IBALOY

ILONGOT

ILOKO

▲ Dibutunan

▲ Kadiyakan

PANGASINAN

SAMBAL

PAMPANGAN

TAGALOG

▲ Botolan
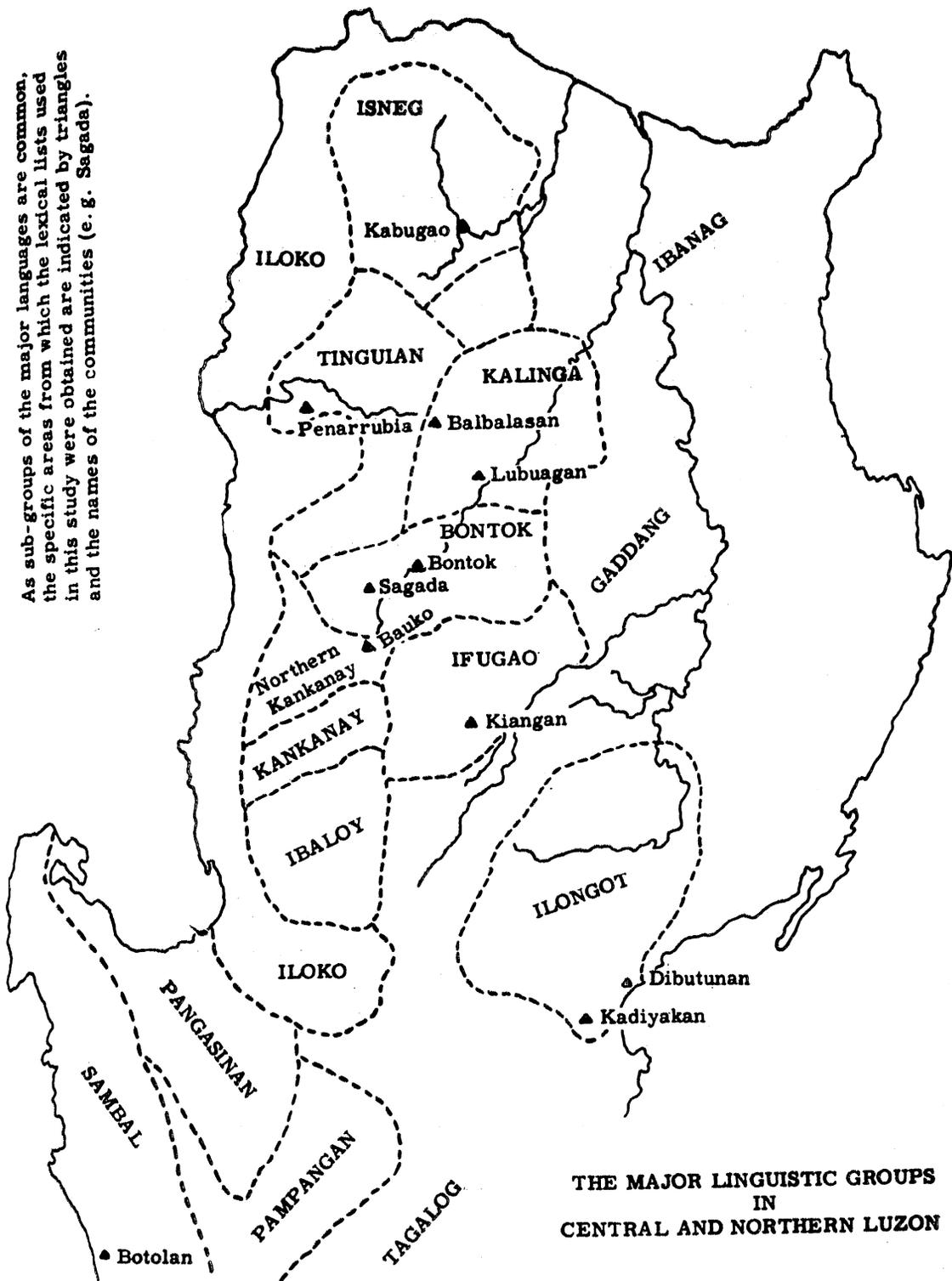
**THE MAJOR LINGUISTIC GROUPS
IN
CENTRAL AND NORTHERN LUZON**