

CONCLUSION

David C. Moody

1. INTRODUCTION

The preceding papers discuss the results of extensive data gathering and dialect intelligibility testing by the Summer Institute of Linguistics (SIL) to determine the present-day dialect and language relationships in Sabah. This summary article briefly reviews the SIL research and provides an overall perspective to these relationships.

Any scientific study rests upon assumptions held by the researchers whether or not they are explicitly stated. In this study it has been assumed that language is a code of meaning used for communication between at least two parties. One elemental basis of language as a code is that each unit of meaning has a particular phonetic shape or form, or set of forms. The similarity between two lists of these form-meaning composites reflects to some degree the relationship between the languages represented by the lists. It has been assumed that language use assumes intelligibility. The degree of intelligibility between two languages reflects to some degree the relationship between the two languages. A further assumption has been that for any two languages, the measure of their lexical similarity should roughly correlate with the degree of intelligibility between them. Lastly, it has been assumed that in the comparison of any two languages, there may be specific factors which skew the expected correlation between lexical similarity and degree of intelligibility (Pike and Pike 1977).

The starting point for these papers has been the data collected by SIL from 325 villages throughout Sabah and representing 344 speech samples. Smith (in this volume) based his computer-aided lexicostatistical study on this initial collection of data. Additional data has been collected and processed in the analysis of some groupings. As lexicostatistical studies go, Smith's has several significant advantages. The data was collected within a relatively short span of time. It was collected by a small group of technicians working in a common effort. The same basic wordlist was used for all collection points, and the wordlist was relatively long.

Using these 344 wordlists, Smith tentatively established a classification of Sabah's language in which he enumerated 51 languages and 83 dialects distributed into 16 language families and 36 subfamily groupings. Smith anticipated that with additional data, particularly data from other sources (e.g. intelligibility testing, comparative studies, etc.), his arrangement of the linguistic groupings would be improved by altering some of the affiliations he had proposed. Smith suggested that the designations of "language" and "dialect" would be most susceptible to change.

Julie K. King and John Wayne King, eds *Languages of Sabah: a survey report*, 325-337. *Pacific Linguistics*, C-78, 1984.

© David C. Moody

325

Smith assumed that there would be very little ability to communicate between speakers whose dialects represented separate subfamilies. Furthermore, he suggested that the definition of "language" implies the absence of mutual intelligibility with other languages. In the same way, the definition of "dialect" implies the presence of mutual intelligibility with other dialects representing the same language, though exception is made for chaining phenomena.

The relative merit of lexicostatistics as a tool for classifying languages has been much debated.¹ The approach taken by Smith and the authors of the preceding language-specific papers is that lexicostatistics can be used to indicate dialect and language relationships, but that any resultant classification must be considered tentative. The tentativeness of the lexicostatistical classification should not be a deterrent to presenting these findings. In his own defense, Dyen wrote, "The reason that the lexicostatistical classification is inconclusive would be the same that affects any classification; not all the knowable facts are known at the time of classification unless one chooses to wait hopelessly until all knowable facts are known. Based on fewer facts than that a classification remains open to correction as additional facts become available" (Dyen 1966:35). Although such a classification is tentative, it provides a data base and a set of hypotheses which may be used to further examine the relationships between dialect and language groups.

The language-specific papers supplement Smith's lexicostatistical analysis with findings from dialect intelligibility testing. In this testing Smith's analysis served as a starting point. His classification of particular dialects and languages served both as a set of hypotheses to be confirmed or refuted by the evidence provided from intelligibility testing, as well as the basis for determining the overall testing framework, that is, the test sets and test points.

The dialect intelligibility phase of the survey comprised a total of 965 tests conducted in 157 villages. A standard Sabah Malay reference tape was used in all but four villages; thus 812 non-Malay tests were conducted in the 157 villages where testing was done. Among these tests there are 62 reciprocal pairs, or 124 tests, where a reference tape from a test point A was tested in a test point B, and the reference tape from test point B was tested in the test point A.

Like lexicostatistics, dialect intelligibility testing has been variously evaluated as a means of determining the degree of linguistic similarity, or dissimilarity, between languages. It has been shown that intelligibility is a measure of social relations as well as linguistic similarity (Wolff 1964; Simons 1979). The component of linguistic similarity, furthermore, is complex, comprising lexical similarity as well as phonological, grammatical and semantic factors. Where social factors are not present, this component of linguistic similarity is directly related to intelligibility. Thus, in such cases, whereas lexical similarity is one component of linguistic similarity, intelligibility scores represent linguistic similarity directly, being composite measures of the several factors which determine linguistic similarity (Simons 1979:15, 57f., 67-87).

The present study uses intelligibility scores as a check on the language classification obtained by lexicostatistical analysis, the language classification itself being a set of hypotheses concerning the similarity, or dissimilarity, between dialects. The fact that the intelligibility scores do not incorporate any factor for social contact should be kept in mind. Nevertheless, in the majority of cases, the intelligibility scores have confirmed the lexicostatistical hypotheses. There is still a need to complete statistical analyses of the data to determine a more precise correlation between the lexicostatistical data and the intelligibility scores.

2. THE PROPOSED CLASSIFICATION OF THE LANGUAGES OF SABAH

The proposed classification of the languages of Sabah is presented in Figure 1. The revised classification (SIL 1982) is based on both the lexico-statistical evidence as well as the intelligibility data. For the sake of comparison this revised classification is presented alongside Smith's classification (in this volume).² The languages of this proposed classification are located on Figure 2.³

Among the dialects from which wordlists were collected, only those languages which were considered to be either indigenous to Sabah or whose tenure in Sabah has been long term were selected for inclusion in the dialect intelligibility testing phase of the project. On this basis, no testing was conducted in Chabacano, Butung, the five Javanese languages or the two Buginese languages. These languages would likewise have been omitted from testing on the basis of their lexical dissimilarity with the remaining speech samples in the whole study. Generally, testing was not conducted if the cognate percentages between two groups indicated they were in different subfamilies, that is, in the range below 75 percentage of shared cognates (PSC), nor if their cognate relations indicate that they were part of the same subdialect, that is 90 PSC or above. All of the dialect intelligibility testing was conducted in languages of Smith's North-western Austronesian superstock.

Testing within the Lundayeh language, including one sample from a village in Sarawak, confirmed a high degree of homogeneity. External testing of Lundayeh was conducted with a Murutic sample and confirmed the distant relationship which had been established by Smith and others.

Testing within the Illanun language indicated that high mutual intelligibility exists between the two dialects proposed by Smith. External testing with samples from the Philippines indicated a close relationship with some members of the Danao language family.

Testing within the Suluk (Tausug) language confirmed the homogeneity of the samples representing Sabah. External testing with two varieties of East Coast Bajau supported the distinction between these two Philippine-originated groups.

Testing within Smith's Ida'an subfamily did not confirm Smith's division of the samples into two separate languages. On the basis of intelligibility between the samples one Ida'an language is now posited. Dialect-level relationships remain unclear.

Testing within the Malayic family was conducted only in Brunei/Kedayan villages. Testing within the Brunei/Kedayan language indicated a division at the dialect level. External testing supported the division which placed the Iban language in a separate subfamily.

Testing within the Bajau family was conducted in both West Coast and East Coast villages. Testing within the West Coast Bajau language indicated at least two (possibly three) dialects should be distinguished. Testing among the East Coast Bajau villages supported the identification of the Sabah samples with three of the Sama languages distinguished in an earlier Philippine study. The three languages are Balangingi, Southern Sama and Kagayan. Most of the Sabah Bajau samples relate most closely to Southern Sama. Two dialects are proposed for Southern Sama.

The remaining testing was conducted among representatives of Smith's Bornean stock. Smith distinguished four language families within this stock-level grouping: Tidong (represented by a single language), Paitanic, Murutic and Dusunic.

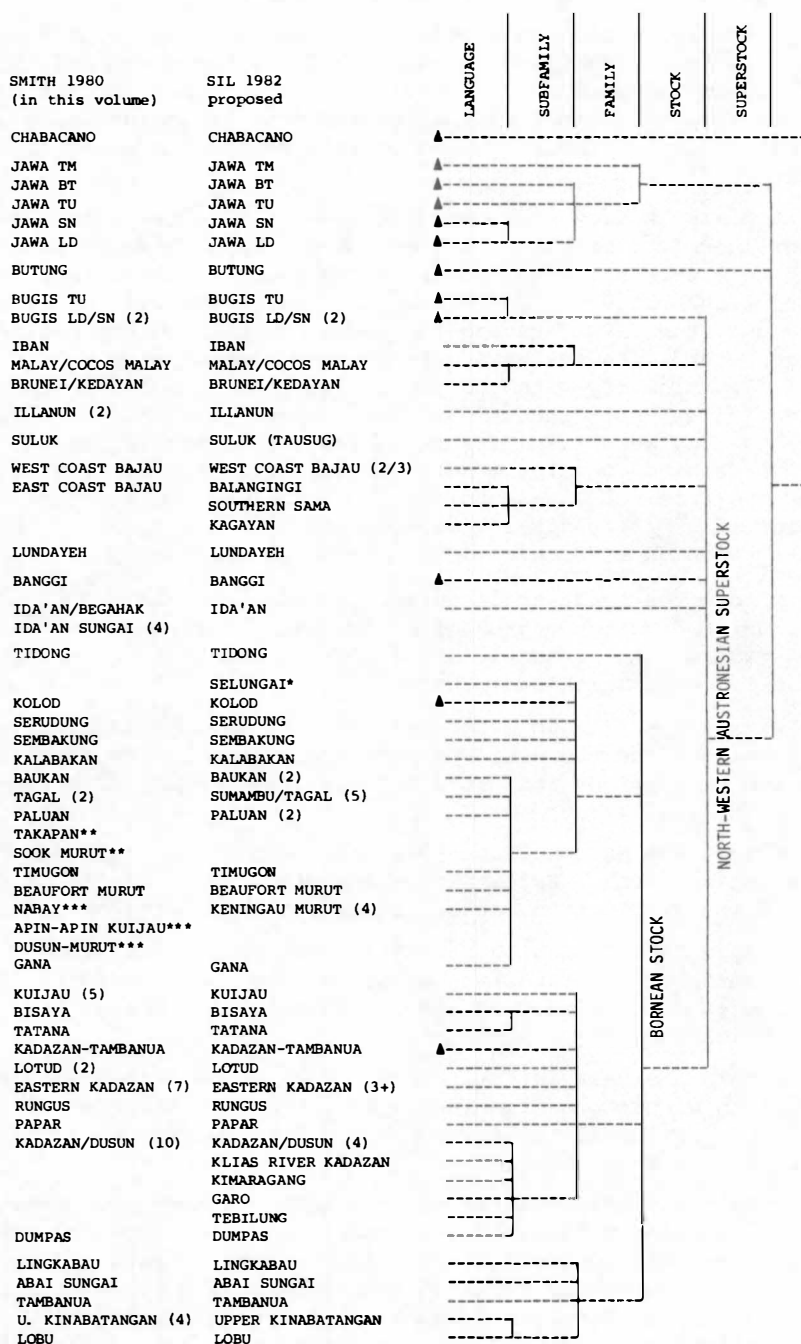


Figure 1: Classifications of the languages of Sabah: 1980 and 1982
 (* classified on the basis of data not available to Smith;
 ** Takapan and Sook Murut have been subsumed as dialects
 of Paluan in the 1982 classification; *** Nabay, Apin-Apin
 Kuijau and Dusun-Murut have been subsumed as dialects of
 Keningau Murut in the 1982 classification; (n) numbers in
 parentheses indicate the number of proposed dialects;
 ▲ indicates a language for which no dialect intelligibility
 testing was completed.)

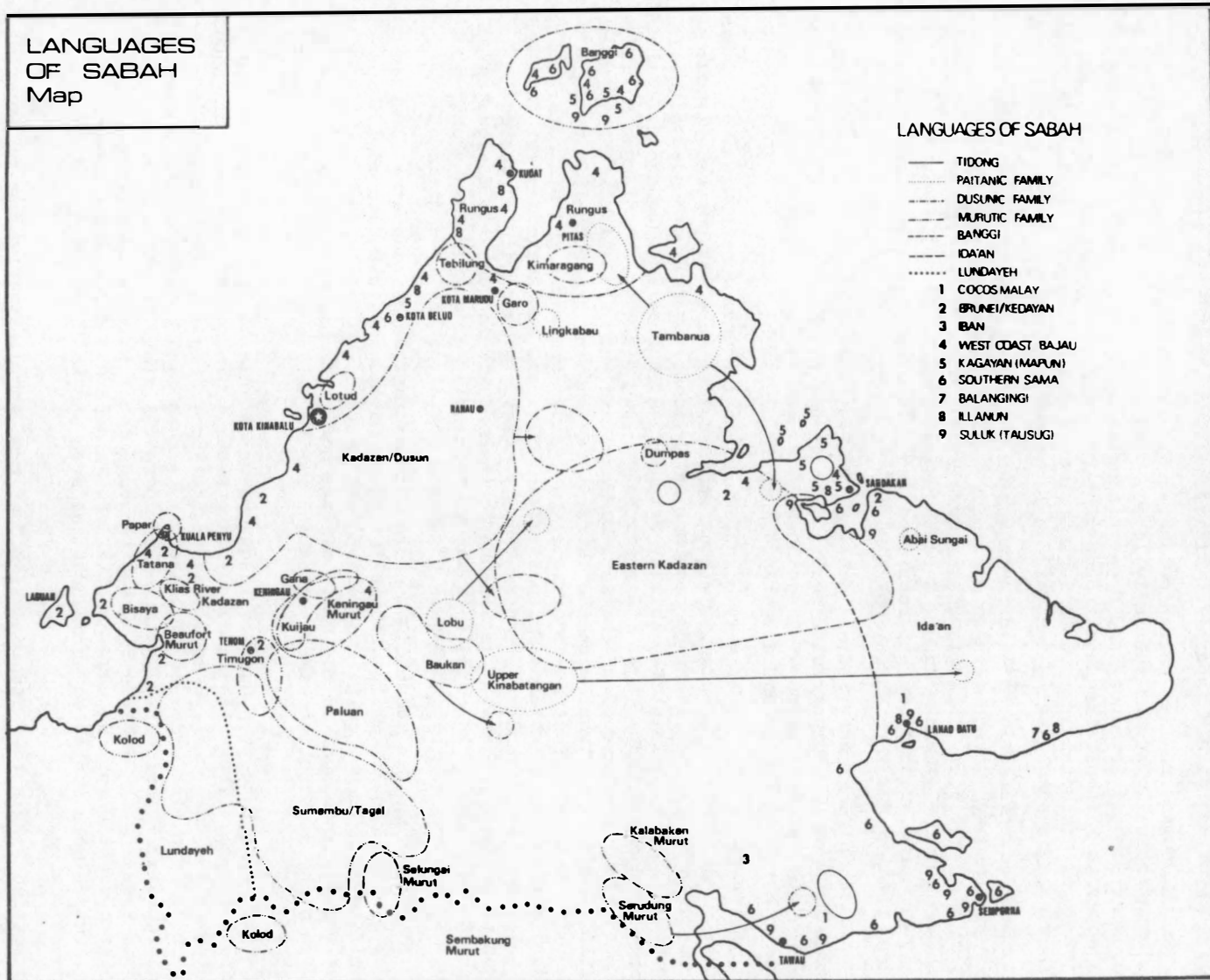


Figure 2: Languages of Sabah map

Testing within the Tidong language along with additional lexicostatistic data confirmed Smith's single-language classification. A possible dialect-level division was also proposed. External testing with representatives of the Murutic family supported Smith's separation of Tidong from the Murutic languages.

Smith enumerated five Paitanic languages, each relating to the others lexicostatistically at the subfamily level. The five languages were Lingkabau, Lobu, Abai Sungai, Tambanua and Upper Kinabatangan. Lobu remains classified as a separate Paitanic language though testing indicated high Lobu intelligibility of Upper Kinabatangan. Dusun Segama remains classified as a dialect of Upper Kinabatangan, though testing indicated a language distinction may be appropriate.

Smith enumerated 15 Murutic languages which he distributed among eight subfamily divisions. Seven of these subfamily divisions were represented by a single language each. These languages were Kolod, Gana, Apin-Apin Kuijau, Kalabakan Murut, Sembakung Murut, Serudung Murut and Tagal. The eighth subfamily grouping, Central Murut, comprised the following eight languages: Takapan, Paluan, Timugon, Beaufort Murut, Dusun-Murut, Sook Murut, Baukan and Nabay.

Additional lexicostatistical data and analysis and dialect intelligibility testing has produced significant realignment of the Murutic family classification. This revised classification distinguishes 12 languages, which group into the Northern Murut subfamily and an eastern Murutic division.

The Northern Murut subfamily comprises seven languages. The distinctiveness of Gana, Tagal, Paluan, Baukan, Timugon and Beaufort Murut as separate languages is confirmed. Gana has remained distinct on lexicostatistical grounds, although it demonstrated mutual intelligibility with Kuijau, a Dusunic language. Tagal has been renamed Sumambu/Tagal and encompasses five dialects: Pensiangan Murut, the most widely understood of the Sumambu/Tagal dialects, Salalir, Rundum, Tomani Sumambu and Maligan Tagal with the latter three showing greater affinity with each other than with Pensiangan or Salalir. Paluan comprises two dialects: (1) Paluan proper, which includes Smith's Paluan, Takapan and Sook and (2) Pandewan, for which Smith had no data. For Baukan, two dialects have been suggested: Baukan proper and Tenggara. The seventh Northern Murut language is designated Keningau Murut and includes Smith's Nabay, Apin-Apin Kuijau and Dusun-Murut groups as well as Ambual; Nabay is the most central of the four dialects.

The term "eastern Murutic" was coined to designate the remaining five languages of the Murutic family without intending to suggest that they should be identified as a formal unit. No testing was done in the Kolod language. Selungai Murut was distinguished on the basis of data not available to Smith. Testing confirmed the distinctiveness of Sembakung Murut, Kalabakan Murut and Serudung Murut.

Smith enumerated ten Dusunic languages which related to each other at the subfamily level. These ten languages were Papar, Dumpas, Kadazan-Tambanua, Lotud (including two dialects), Bisaya, Tatana, Kuijau (including five dialects), Eastern Kadazan (including seven dialects), Rungus and Kadazan/Dusun (including 13 dialects). Researchers attempting to do intelligibility testing in the only kampung speaking the Kadazan-Tambanua language found that subsequent to the earlier collection of language material the kampung had dispersed. Speakers of the language were scattered throughout the area and no longer functioning as a language community.

Testing within the Dusunic family confirmed the distinctiveness of Papar, Lotud, Bisaya, Tatana, Kuijau, Eastern Kadazan, Rungus and Dumpas. Within this testing several anomalies were noted. Papar speakers demonstrated high intelligibility of Tatana. Lotud seems to function on the local level as a dialect of Kadazan/Dusun as indicated by high intelligibility scores. Furthermore, the

mutual intelligibility between Bisaya and Tatana indicated that a closer relationship may be appropriate.

In addition, Smith's proposed dialect divisions for the Kuijau language were not confirmed by the dialect intelligibility testing. At least three dialects are confirmed for Eastern Kadazan, though others remain unclear. Internal and external testing of Dumpas indicated it may be more appropriately classified as a Paitanic rather than as a Dusunic language.

Smith's Kadazan/Dusun language designation has been significantly realigned. Testing within Kadazan/Dusun indicated four additional languages should be enumerated. These languages are Klias River Kadazan, Kimaragang, Garo and Tebilung. Four dialects are now posited for the Kadazan/Dusun languages; these include Central Dusun, Coastal Kadazan, Sugut Kadazan and Kuala Monsok (TA) Dusun.

Within the Bornean stock several languages seem to function transitionally between family groupings. Between Tidong and Murutic, the Kalabakan and Sembakung Murut languages share this position. Between the Murutic and Dusunic families, Murutic Gana and Dusunic Kuijau appear to be in the margin. Between Paitanic and Dusunic, Dumpas is marginal. Further study of these transitional areas between language families may provide valuable insights into the historical development of the Bornean languages.

In summary, the number of languages representing cultures of insular Southeast Asia has increased by two over Smith's calculation. The changes are identified as an increase from one to three East Coast Bajau languages, a decrease from two to one Ida'an language, a decrease from 15 to 12 Murutic languages and an increase from ten to 14 Dusunic languages.

3. NOTES ON BILINGUALISM

The survey data reflects numerous instances of bilingualism, where bilingualism is defined as the learning of a second language through social contact (a person may be "bilingual" in several languages). Bilingualism is reflected in the data by intelligibility scores which are significantly higher than corresponding PSC values. Sometimes bilingualism is unidirectional, or non-reciprocal, that is, speakers of language A learn language B, but speakers of language B do not learn language A. Bilingualism may also be bidirectional, or reciprocal, that is, speakers of language A learn language B and vice versa. This latter type may also be referred to as mutual intelligibility, though mutual intelligibility is not restricted to bilingualism; mutual intelligibility is expected to occur between very closely related speech samples. Thus bilingualism, whether reciprocal or non-reciprocal, does not contrast with intelligibility, but is rather a particular type of intelligibility (Simons 1979:3).

The survey data includes 62 reciprocal pairs of test and reference points. Among these pairs there are 20 cases where both A and B speakers registered intelligibility scores of 80% or higher on each other's language. These would be examples of mutual intelligibility. In only one case where the PSC relationship between sample A and sample B was below 75 PSC did speakers of both languages demonstrate mutual intelligibility above 80%. The Bisaya villagers of Padas Damit BT registered 85% intelligibility of the Tatana language of Bundu KP, and Bundu KP registered 88% intelligibility of Padas Damit BT; the cognate relation between the respective wordlists is 62 PSC. This appears a strong case for reciprocal bilingualism.

More frequently bilingualism is non-reciprocal. In the data on reciprocal pairs there are ten clear cases of non-reciprocal bilingualism where speakers of sample A registered an average intelligibility score of greater than 80% on sample B, and speakers of B registered intelligibility of less than 80% on A, and the spread between the two scores was greater than ten percentage points. (The range for the difference between the two scores of these reciprocal pairs is 14-35%. The average difference for the ten pairs is 27%.)

Bilingualism is largely a product of social contact between speakers of different languages, though the degree of linguistic similarity may facilitate the language learning process. In addition to cases of bilingualism which have already been cited, dialect intelligibility testing in the Keningau District demonstrated a high rate of bilingualism. Out of 70 comparable cases, there were 26 intelligibility tests where the difference between an intelligibility score above 80% and a cognate relation below 80 PSC was greater than ten percentage points. Within Keningau District, the town of Keningau serves as a market, educational and administrative centre for speakers of both Dusunic and Murutic languages, thus setting the stage for a high degree of social contact between language groups. The nature of the factors affecting bilingualism within Keningau District and elsewhere in Sabah is an area where further study is desired.

Inasmuch as the intelligibility data has not yet been analysed in a way which might account for the effect of social factors, the handling of the data has been subjective in some cases. The value of the intelligibility data, however, cannot be dismissed. In many cases Smith's conclusions were confirmed by the results of intelligibility testing where there is no reason to expect a skewing of the results due to sociolinguistic factors. On the other hand, in some cases the results of intelligibility testing indicate that a reclassification is appropriate. An example of the latter is the Dumpas situation. On lexicostatistical grounds, Smith had classified Dumpas as a Dusunic language; its highest percentage of shared cognates was with a dialect of Eastern Kadazan, a Dusunic language. Dumpas is now spoken in an area where Eastern Kadazan is the dominant language. Intelligibility testing, however, indicated that Dumpas is more closely related to the Paitanic family. This latter indication is likewise supported by sociolinguistic evidence, as well as the lexicostatistical evidence that Dumpas shows lower lexical relationships with Dusunic languages than it does with Paitanic languages.

Testing within the Malayic family provided another example to demonstrate how intelligibility data can supply evidence for reclassification. Lexicostatistically, the Brunei/Kedayan samples tested appeared to be dialectally the same. Mutual intelligibility was expected. The testing results, however, show that the intelligibility is non-reciprocal. A dialect distinction, whether the basis is some yet undiscovered linguistic dissimilarity or sociolinguistic factors, should be recognised.

4. NOTES ON THE TESTING OF NATIONAL LANGUAGE COMPREHENSION

Testing for comprehension of the national language was conducted in all but four villages where intelligibility testing was conducted throughout the state. There were 146 such tests where there was also a cognate relation figured between the wordlist for the dialect being tested and the Sabah Malay wordlist. These scores divided almost evenly into two groups. Of the 146 correlations,

70 of the Malay test scores were between 75% and 100%, and 76 scores were between 12% and 70%. The average score for the higher percentage group was 86.5%, while the average score for the lower percentage group was 52%. The mean for all 146 scores was 68.5%.

Generally, throughout the testing several factors were noted as contributing to higher scores on the Malay reference tape. The most crucial factor seems to be education. In only a few locations did this correlation between education and higher Malay scores not hold. Age as a factor seems directly related to the education factor. The older test subjects generally (though with some notable exceptions) had not received much, if any, formal education. Education also seemed to level the difference in performance between males and females. Among the formally uneducated, males generally outscored females, which seemed to be a result of their greater mobility. Finally, the location of the test point with respect to population centres with their market, educational and administrative functions correlated well with the scores. It is evident that as the government of Sabah continues to develop its resources and its people through improvement of its educational institutions the percentage of people who competently and confidently can communicate in the national language will likewise continue to increase.

5. COMPARISON WITH DYEN

The proposed classification as it is presented here is similar to that produced by Dyen (1965), though there are some differences (Figure 3). All of the languages except Chabacano fit within Dyen's Hesperonesian Linkage. The Bugis and Malayic languages are most incongruous with Dyen's classification. He identifies Bugis (Buginese) with his Celebes Hesion. Butung, a language of southern Celebes which Dyen does not discuss, is placed within a Celebes grouping by Voegelin and Voegelin (1977), and presumably would also fit within Dyen's Celebes Hesion.

Dyen places Malayic coordinate with Javanese in his West Indonesian Cluster. Dyen groups his Celebes Hesion, West Indonesian Cluster and Northwest Hesion as coordinate members of Hesperonesian. In the present study the Bugis subfamily and the Malayic family are subordinate within Smith's North-western Austronesian superstock, which otherwise roughly parallels Dyen's Northwest Hesion. A possible implication of this is that either a great degree of lexical inflation is present in the data under consideration (that is, northern Borneo Bugis and Malay have become more like the remaining languages in the Northwest Hesion or vice versa), or there is actually a lesser degree of language variation for these languages than indicated by Dyen.

Most of Sabah's languages fit within Dyen's Northwest Hesion. Three languages of the present study, each representing a distinct linguistic stock (Smith's terminology), are not found in Dyen. Lundayeh, the most disparate language within Smith's North-western Austronesian superstock (which roughly parallels Dyen's Northwest Hesion) is, according to Clayre (1970), closely related to Kelabit which Dyen did not assign to one of his Hesperonesian divisions, though he notes that the closest relation between a member of his Kelabitic Cluster and a non-member is with Murut (30.2%, a figure which parallels Smith's figure for the relationship between Lundayeh and the Bornean stock languages). The Banggi language may represent Dyen's Palawan subdivision of his Sulic Hesion, as it was found to relate more closely with the Molbog language

spoken on islands in the southern Palawan area than with Sabah languages. This, however, would indicate a closer relationship with Suluk (Tausug) of Dyen's Bisayan Cluster than with Ida'an and the Bornean stock languages, contrary to Smith's findings. The third language representing a separate linguistic stock is the Ida'an language. Ida'an has no known counterpart in Dyen's work.

One other area of comparison between Dyen's study and the SIL study involves the languages Smith has grouped together within his Bornean stock. Dyen grouped Tidong with Murut, whereas the SIL study finds reason to keep them distinct. Although Dyen separated Murutic and Dusunic languages as coordinate members of his Northwest Hesion, they are grouped together within the Bornean stock in the present classification. The Paitanic languages in the present study are not represented in Dyen's work.

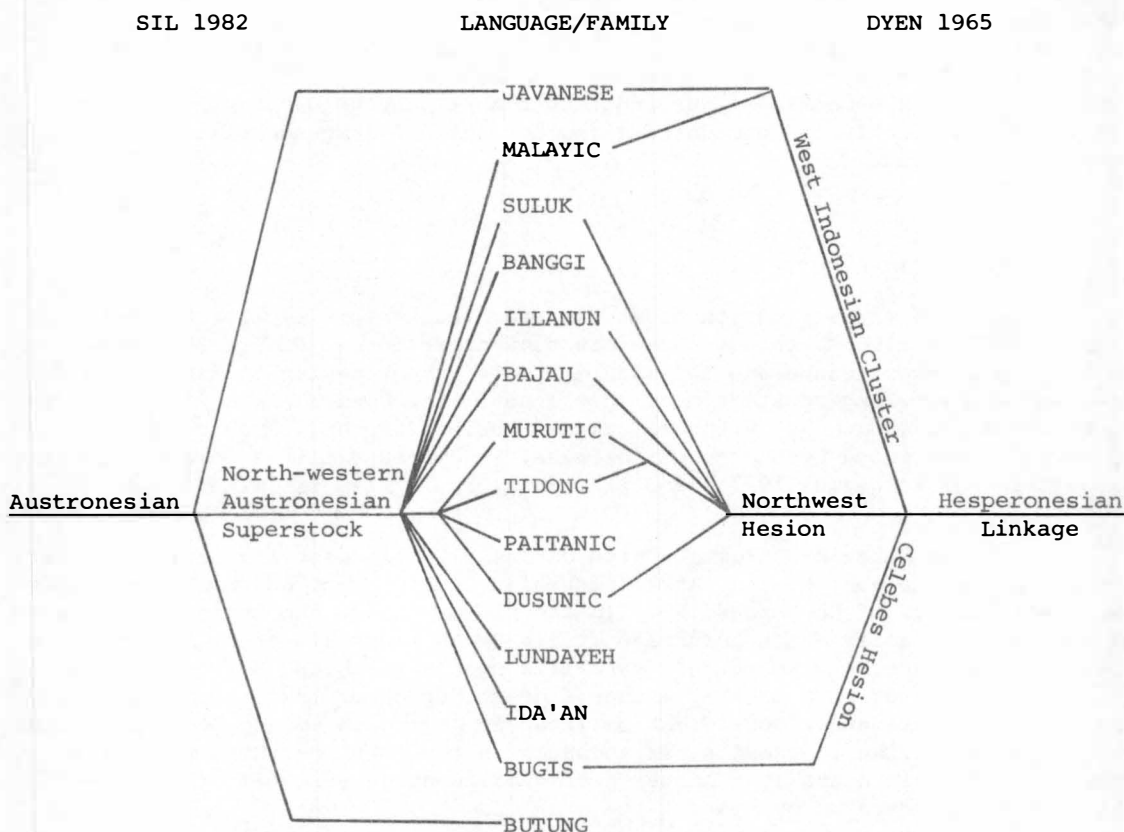


Figure 3: Comparison of SIL's classification of Sabah languages with Dyen's 1965 classification for the same language groups. (The language designations follow SIL's classification.)

6. COMPARISON WITH PRENTICE

The majority of the languages of this study are not known to be represented by communities located outside of northern Borneo. Thus it is most appropriate to compare these results with that of D.J. Prentice (1970) whose classification of Sabah languages has heretofore been the most thorough and detailed. Figure 4 presents this comparison in summary.

Prentice's classification comprises four taxonomic levels: subfamily, language-group, language and dialect. Prentice includes three higher order categories, the Murutic, Dusunic and Paitanic subfamilies.⁴ Using different thresholds (Smith, in this volume), the SIL classification includes stock and family groupings as well, but does not employ Prentice's "language-group" category, thus making "family" and "subfamily" in the current work roughly equivalent to Prentice's "subfamily" and "language-group", respectively.

Perhaps the most apparent of the differences between the two classifications is in the enumeration of languages. Prentice lists 19 languages whose cultural centres could be said to lie in Sabah. The corresponding SIL count is 34 languages. Most of the difference can be attributed to the more extensive collection of uniform-quality data in the SIL survey.⁵ Some differences also result from the application of different thresholds.

Within the two classification several differences can be noted. In his Murutic subfamily Prentice joins the Tidong and Murut "language-groups" (as did Dyen). In the present study, though Tidong's relation with Murut is regarded as closer than with Dusunic or Paitanic, the two are separated at the higher level. Two of Prentice's three Tidong constituents, however, are considered members of the Murutic family in this study. Prentice includes four Bisayan languages,⁶ including Tatana, while SIL distinguished only one Bisaya language which, as Prentice indicated, shows a closer relationship with Tatana than with other languages of the Dusunic family. Prentice includes the Banggi, Paitan and Buludupi languages within his Paitanic grouping. SIL has distinguished Banggi as a separate stock, which as noted above (as also by Prentice) may relate more closely to languages of the southern Philippines than to other Sabah languages. SIL's research has revealed five Paitanic languages rather than one. Buludupi, for which Prentice found nineteenth century wordlists and which is attested to in literature during the time of the Chartered Company rule, has not been positively identified by SIL investigations.⁷

Finally, this survey has attempted to define linguistic entities within the geographical confines of Sabah, not to link them in any formal way with languages found geographically outside of the state. The larger task of clarifying the higher order relationships which exist across Sabah's borders remains to be done. Even within the scope of the present work there are numerous relationships which remain tentative. These have been indicated by the various writers. It is hoped that the presentation of these papers will contribute significantly to the knowledge of Sabah's peoples and languages and will provide incentive to further linguistic research within insular South-east Asia.

PRENTICE 1970				SIL 1982 - proposed			
LANGUAGE -----Dialect	SUBFAMILY	FAMILY	STOCK	FAMILY	LANGUAGE -----Dialect		
TARAKAN TIDONG	MURUTIC	IDA'AN	BORNEAN	MURUTIC	TIDONG		
TINGGALAN TIDONG					SEMPAKUNG (E)		
TANGGARANG TIDONG					BAUKAN (N) (2)		
					-----Tenggara		
					-----Baukan proper		
					KALABAKAN (E)		
					SERUDUNG (E)		
					SELUNGA (E)		
					KOLOD (E)		
					SUMAMBU/TAGAL (N) (5)		
OKOLOD					-----Salalir (including Lumbis)		
SALALIR MURUT							
HIGHLAND MURUT					-----Tomani Sumambu		
-----Alumbis					-----Rundum Murut		
-----Sumambuq					-----Maligan Murut		
					-----Pensiangan Murut		
-----Paluan					PALUAN (N) (2)		
					-----Paluan proper		
LOWLAND MURUT					-----Pandewan		
-----Timugon							
					TIMUGON (N)		
-----Nabay					BEAUFORT MURUT (N)		
					KENINGAU MURUT (4)		
					-----Nabay		
					-----Apin-Apin Kuijau		
					-----Dusun-Murut		
-----Baukan*					-----Ambual		
MINANSUT DUSUN	DUSUNIC			DUSUNIC	GANA		
INLAND DUSUN					KUIJAU		
-----Menindaq					KADAZAN/DUSUN		
-----Tagaas					-----Kuala Monsok (TA) Dusun		
					-----Central Dusun		
-----Tambunan					-----Sugut Kadazan		
-----Ranau							
-----Kiau					-----Coastal Kadazan		
TAMPASUK DUSUN							
COASTAL DUSUN					RUNGUS		
-----Kadazan					LOTUD		
-----Marudu					TATANA		
LOTUD					BISAYA		
TATANAQ BISAYA					KADAZAN-TAMBANUA		
NORTHERN BISAYA					EASTERN KADAZAN (3+)		
SOUTHERN BISAYA					-----Lamag Sungai		
TUTONG BISAYA					-----Labuk/Mangkaak/Sukang		
					-----Malapi Kadazan		
					PAPAR		
					KLIAS RIVER KADAZAN		
					KIMARAGANG		
					GARO		
					TEBILUNG		
					DUMPAS		
PAITAN	PAITANIC			PAITANIC	LINGKABAU		
					ABAI SUNGAI		
					TAMBANUA		
					UPPER KINABATANGAN (2+)		
BANGGI					LOBU		
BULUDUPI							
					BANGGI		
					IDA'AN		

Figure 4: Comparison of SIL and Prentice classifications of Sabah languages. (The lines of each column of classification are not necessarily equivalent. The arrangement is intended primarily to show the similarities and differences within the higher taxonomic levels.) (E) indicates a member of the eastern Murutic group of languages; (N) indicates a member of the Northern Murut subfamily; (n) numbers in parentheses indicate number of established dialects; * Prentice's Baukan dialect is parallel to SIL's Baukan proper dialect. Language names are in upper-case letters; dialect names are preceded by five hyphens.)

NOTES

1. For an example of such debating as it also applies to a particular lexico-statistical classification, see the review of Dyen (1965) by Grace (1966) and the comments by Dyen (1966) and Hymes (1966) in the same volume.
2. Since Smith's lexicostatistical analysis (in this volume) was completed in 1980, the classification arrived at from that analysis is referred to as Smith 1980. Similarly, since the revised classification arising out of the intelligibility testing was completed in 1982, it is here referred to as SIL 1982.
3. The map serves as a general guide to the locations of the various language groups. Only those languages which were considered either indigenous to Sabah or of long-term residence in Sabah, and which have been discussed in this volume, have been included in the map.
4. Appell (1968) listed Bisaya as a separate subfamily along with Dusun and Murut in his more tentative earlier classification.
5. Prentice's study was based on the comparison of 73 wordlists, 64 of which he considered as "Ida'an" (=Bornean stock) languages; 54 of the lists were gathered from published sources dating as far back as 1872.
6. Prentice cites locations in Sarawak for two of his Bisaya groups, Southern Bisaya and Tutong Bisaya.
7. Prentice cites two wordlists, both of inadequate length, which exhibit 50 PSC between them. One of the lists demonstrates high lexical relations (83-87 PSC) with several of SIL's Ida'an wordlists. It is, of course, unlikely that the second list would likewise show such high relations with the same Ida'an lists. It is more likely that "Buludupi" is an early exonym applied to more than a single linguistically distinct speech group (John A. Spitzack, personal communication).

