

PART 2.2.

LANGUAGE CLASSIFICATION

143-144

2.2.0. PAPUAN LANGUAGE CLASSIFICATION PROBLEMS

S.A. Wurm and K. McElhanon

2.2.1. INTRODUCTORY REMARKS

The discussions given below in the various chapters of Part 2.3. point out the strong influences exercised by Papuan languages upon each other. It is clear that these influences bear heavily on the question of the classification of these languages and, in particular, call into question the value of the lexicostatistical approach which, in many instances, has been the basic tool for the classification of Papuan languages until fairly recently. While the importance of this approach for an initial crude attempt at classifying previously unclassified languages is obvious, its usefulness in the Papuan linguistic field must be regarded as very severely limited because of the almost ubiquitous presence of varied influences of languages upon each other on virtually all levels which can reach a magnitude unrealised in the framework of the lexicostatistical method and in the basic principles of genetic linguistics in general. Basic vocabulary items of one language which according to the postulates of lexicostatistics, are "unborrowable", are often encountered as quite obvious loan words in another unrelated language. Similarly, pronouns, singly and in sets, are taken over by languages from other languages, and the structures of many languages have apparently undergone quite drastic changes under the influence of other, sometimes unrelated, languages. Phonologies can change extensively under such an influence, and the only apparently relatively stable and persistent items and features in Papuan languages are constituted by verbs as lexical items, some structural features of verbs and their underlying principles, principles underlying pronominal systems - but to a much lesser extent the pronouns themselves as lexical

items - and semantic characteristics of the grouping of lexical items (e.g., *fire* and *tree* can, in one language, be different meanings of one lexical item, but in another, the meanings of two different lexical items: such principles are preserved in a particular language even if the lexical items themselves are borrowed into it from another language in which such items have semantic ranges which are completely different from those present in the borrowing language). In consequence, such items and features have considerable diagnostic importance in comparative and classificatory work. So for instance, there are several instances of languages, especially in the Sepik-Ramu Phylum, which show somewhat differing phonologies, relatively little cognation in nouns, and great differences in the form, but not the system, of their pronouns, but have many verbs in common, together with much of their verb morphology.

2.2.2. "MIXED" LANGUAGES

This is not the place to argue the case of "mixed languages" whose existence is strongly denied by many comparative linguists with the notable exception of A. Capell who has often been severely criticised for this (e.g. in the comments to Capell 1962), and of A. Pawley (1969), but in the light of Grace's (1965) statement that very little is known about language mixing, there seems to be considerable merit in Pawley's comments that the main reason for our continuing ignorance in this field is the fact that most linguists are reluctant to believe that extensively mixed languages might exist. This is of course largely attributable to the fact that the genetic model does not allow for mixed languages which therefore constitute an unacceptable concept as such. This is rather unfortunate because the New Guinea area may offer the probably best laboratory situations in the whole world for the study of language mixing: in several areas, two different, sometimes unrelated, languages are spoken in separate villages, whereas in villages with mixed populations hailing from the other villages, languages are spoken which appear to be "mixtures" of the two different languages. A thorough study of such languages by a comparative linguist and a sociolinguist would undoubtedly contribute materially to our knowledge of the nature of language mixing.

It must be pointed out in fairness, that most comparative and genetic linguists are prepared to admit that a certain amount of borrowing, or mixing, is present in all languages - apart from obvious loans on the lexical level - but they either appear to assume, at least by implication, that an inner core is present in every language

which is impervious to borrowing and which therefore constitutes the component which is relevant in a genetic classification of a language, (e.g. Hamp 1962), or they seem to feel that one component of a language is always dominant, even if borrowing or mixing on all levels of a language may be possible. This dominant element is in their view, the one which is drawn upon in a genetic classification of such a language, and it represents that particular language in such an exercise.

Pawley (1969) expresses the opinion that such a genetic classification which deliberately ignores elements in a language which lie outside the dominant component, so as to obtain an unambiguous, yes or no, classificatory result, provides no adequate basis for the reconstruction of linguistic events, nor does it give an adequate picture of linguistic relationships.

This is undoubtedly true in extreme cases. For instance, which is the dominant component in languages such as Magori (Dutton 1972, see also (II) 4.5.2.) in which the grammar is almost purely Austronesian, and the vocabulary mostly Papuan, or Maisin (Ray 1911, Strong 1911) in which the reverse appears to be true in some ways (see (II) 4.5.1.)? Or in the languages of the Reef Islands-Santa Cruz Family (East Papuan Phylum, 2.13.1.4.) in which about half the basic vocabulary is reconstructably Austronesian (Wurm 1970), and the grammar an intricate mixture of Papuan and Austronesian elements (Wurm 1969) (see (II) 4.5.3.).

However, for less extreme cases, the attitude could be taken, in contrast to Pawley's view, that a genetic classification based on the dominant element only can well be of value in the establishment of linguistic relationships, provided that the presence and nature of the non-dominant element is not ignored, but taken into account and its presence explained as far as possible in terms of sub-strata, borrowings, influences by other languages and the like. In extreme cases such as those mentioned above, Wurm suggests that the dominant element in the language structure, if one can be recognised, should be regarded as the dominant element in the language for the purpose of genetic classification, i.e. Magori should be regarded as Austronesian, and the Reef Islands-Santa Cruz Family languages as Papuan because the Papuan structural elements in them are more basic - e.g. concern the verb structure - than the Austronesian elements (Wurm 1969). Maisin is unclear even from this point of view, because in spite of its Papuan features, it may, when judging from a sociolinguistic point of view, probably be originally Austronesian rather than Papuan (Dutton, personal communication) (see (II) 4.5.2.).

The principles outlined above constitute, in essence, the basis for the classification of Papuan languages, but need to be elaborated on (see 2.2.4.).

2.2.3. LEXICOSTATISTICAL CLASSIFICATION AND STRUCTURAL EVIDENCE

The first linguist to discuss the shortcomings and difficulties of the application of the lexicostatistical method of Papuan languages was McElhanon (1970a, 1970b, 1971). He has found that the main problem of the use of the method as a survey tool is the fact that the presence of many loan words remains undetected and the results are often skewed in the direction of a chain relationship. He feels that the frequent presence of dialect and language chains in lexicostatistical classifications in the New Guinea area "provides considerable circumstantial evidence for concluding that the lexicostatistical method as usually applied is incapable of handling the phenomena caused by the wave principle and unable in many cases, if not in most cases, of providing an accurate sub-classification" (McElhanon 1970b:227).

The essence of this statement is that the lexicostatistical method can produce results which are only approximate.

One obvious way to increase the accuracy and reliability of classificatory results arrived at by the lexicostatistical method would be to identify loanwords in a lexical corpus which is to be utilized for lexicostatistical purposes, by the application of the comparative method, and to eliminate them from the corpus. However, such an approach is often not feasible in the Papuan linguistic field. There are two main reasons for this: a) detailed comparative work is only just beginning in some areas, with only preliminary results available to date, whereas in many other areas no comparative linguistic work has been undertaken; b) in many instances, loanwords may have entered a particular Papuan language thousands of years ago, and have undergone the same sound-changes as those words which are part of the original, directly inherited lexical stock of the language. It is usually not possible to detect such loanwords in the light of the low-level sophistication of present-day Papuan comparative linguistics. (However, it has been found that there has always been a very sharp increase in the number of obvious cognates shared by two or more languages, over those established by inspection only, as soon as some, however preliminary, reconstruction work had been carried out and possible proto-forms suggested. - see below 2.2.5.).

To overcome the difficulty presented by unrecognised loans, Thomas and Healey (1962) have suggested a special refined lexicostatistical

approach which has as its aim the detection of the presence of such unrecognised loans and their special treatment in computations for the purpose of circumventing the problem of chaining. McElhanon investigates the applicability of this approach to the lexicostatistical classification of the Huon Peninsula area languages, but questions the validity of its assumptions and expresses doubts about its usefulness (McElhanon 1970b).

Having challenged the value of the lexicostatistical method for the classification of the languages of the Huon Peninsula area, and, by implication, of Papuan languages in general, McElhanon advocates the study of typological and structural criteria for this purpose. He suggests that such criteria are particularly useful when a number of structural features form a set or block which is associated with a group of languages. On this basis, he finds it relatively easy to sub-divide the language of the Huon Peninsula Group - i.e. what has later been recognised as the Huon Peninsula Stock (McElhanon and Voorhoeve 1970) - into two families along lines which deviate strongly from the results of their lexicostatistical assessment and the principles of lexicostatistical classification in general (McElhanon 1970a). At the same time, a couple of obviously highly mixed languages are left whose exact classificatory position still remains indeterminable, and McElhanon points out that the final decisions in such cases are unavoidably arbitrary.

Wurm is generally in agreement with McElhanon's opinions on the subject, though he is inclined to take a somewhat more moderate view.

It may be mentioned that, with more results becoming available on the languages which have been regarded as constituting the Huon Peninsula and Finisterre Stocks (McElhanon and Voorhoeve 1970), the two stocks have been combined into a super-stock (see 2.5.3.3.2. in this volume). McElhanon himself prefers to regard them as constituting a single stock containing all the families and family-level isolates originally making up the two stocks (see 2.8.1.2. in this volume). The factors impinging on family-level classification as mentioned above remain however fully valid in spite of such a re-classification.

2.2.4. CRITERIA FOR THE CLASSIFICATION OF PAPUAN LANGUAGES

In the light of what has been said so far in this chapter, two languages X and Y which share lexical cognates and/or show other similarities, will be regarded as genetically related if the following criteria apply:

2.2.4.1. LEXICAL CRITERIA

A) For language areas in which at least some comparative linguistic work, with reconstructions, has been carried out (e.g. the Trans-New Guinea Phylum area):

a) A number of the lexical items in Y can be shown to constitute reflexes of proto-forms reconstructed for the language group to which X belongs, and are, in the case of obviously only distant relationship, not just simple copies of the forms of these items in X when allowing for the phonological differences between the two languages on the synchronic level.

b) The lexical items of Y mentioned under a) contain at least one verb, most importantly *eat* or *say*, *speak*, and, rather less significantly, at least two of the range of pronouns present in Y, more importantly pronouns denoting the first or second persons singular or the first person plural, less so that of the third person singular.

c) The nouns amongst the lexical items of Y mentioned under a) contain at least some of the following: *arm* (or *hand*), *bone*, *breast* (*female*), *ear*, *eye*, *fire*, *louse*, and less importantly *mother*, *skin* and *water*, and do not lack a high proportion of these while a considerable number of other nouns in Y constitute reflexes of the X group proto-forms. The nouns listed have been found through empirical observation, to be more commonly cognate in related Papuan languages than others.

d) Semantic groupings of lexical items as observable in X are also present in Y even if the items concerned are themselves non-cognate.

B) For language areas for which no comparative linguistic work has been carried out, criterion a) reads as follows:

a) a number of the lexical items in Y appear to be cognate to their equivalents in X and, in the case of obviously only distant relationship, the sound correspondences observable in such apparently cognate items appear to be greater than may be expected in the light of the phonological differences between the two languages on the synchronic level.

Criteria b), c) and d) apply as under A).

2.2.4.2. TYPOLOGICAL AND STRUCTURAL CRITERIA

a) A number of typological and/or structural characteristics and features are shared by the languages Y, and X and other languages of the group to which X belongs. Of such shared characteristics, those connected with the structure and typology of verb forms are particularly

important and diagnostically more valuable than others except that the sharing of principles underlying the formation, distribution and function of verb forms has also high diagnostic value even if the verb forms concerned are formally different in the languages X and Y. At the same time, the sharing of principles underlying pronominal systems has also considerable diagnostic importance even if the pronouns themselves are formally different in the languages Y and X.

b) No typological and/or structural features are present in languages Y or X which are so contradictory as to be deemed incompatible and mutually exclusive, and which affect the basic verb structure and/or principles underlying it, and/or the principles underlying pronominal systems.

Note: this criterion b) is of low diagnostic importance because its applicability is very much dependent on personal judgment, and because of the fact that the influence of substrata in Papuan language may in cases manifest itself on this very level.

2.2.4.3. OTHER CRITERIA

In the consideration of the criteria listed in 2.2.4.1. and 2.2.4.2. in a particular instance involving two languages, the cumulative nature of evidence has to be taken into account and negative and positive evidence are mutually cancelling to some extent. In this, the criteria mentioned in 2.2.4.1. A) b), c), have much greater importance than that referred to in 2.2.4.1.A) a) if the latter appears by itself only without including b) and c). The criterion referred to under 2.2.4.1.A) d) is important, but only if other criteria pointing to the probable presence of a genetic relationship between the two languages are also observable.

Of the criteria discussed in 2.2.4.2., a) is much more important than b), and in general, the criteria listed in 2.2.4.2. carry greater weight than those given in 2.2.4.1., especially in doubtful cases, but their application is of somewhat doubtful value if none of those given in 2.2.4.1. provides positive evidence in a comparison of two given languages, unless 2.2.4.2. a) supplies formal structural evidence on the verb level, preferably accompanied by some on the pronominal level as well.

There may still be cases in which even the application of all these criteria may not yield reasonably clear results, be it on whether languages X and Y are genetically related, or on whether the genetic relationship between Y and X should be regarded as closer than that

prevailing between Y and Z, with consequent bearings upon the classification of Y. Such cases have to be decided on their merits by the classifier's judgment in the light of his experience with Papuan language classification.

To complete the picture, it may be mentioned that the nature and extent of the relationship of one of two languages studied for classificatory purposes, to a third language may sometimes decide a difficult issue: if it appears that the relationship of Y which shows relationship links with X, may be more plausible with Z in terms of what could be regarded as the "dominant component" of Y at least for this given situation, it seems more appropriate to classify Y as related to Z rather than to X.

2.2.5. DEGREES OF INTERRELATIONSHIP

To determine the degree of interrelationship of languages with a view to their inclusion into sub-families and families, stocks, and phyla, percentage figures denoting the extent of basic vocabulary sharing between them have been traditionally resorted to, though in the light of what has been said in 2.2.3., the value of such evidence by itself is often of a low order (see below). The following principles have been adopted:

Communalects have been regarded as dialects of the same language if their cognation percentages are mostly above 81%, and as separate languages if these percentages are mostly below 78% and rarely above 81%, with the decision regarding the presence or absence of cognation based on regular sound correspondences in the majority of the cases. They have been considered as members of the same sub-family if they show rarely below 45% and mostly above 55%, but below 70% basic vocabulary cognation, and as members of the same family if the percentages of basic vocabulary cognates shared by them are rarely below 20%, and mostly above 28%. They have been looked upon as members of the same stock if their cognation percentages are usually below 28% and more frequently below 20%, but only infrequently fall below 12%. They have been classed as members of the same phylum if the percentages are usually below 12%, but not lower than 5%. All these percentage figures are based on the use of a list of about two hundred basic vocabulary items which is a version of the Swadesh list (Swadesh 1955) as modified by Wurm.

In the light of what has been said in this chapter 2.2.0., especially in 2.2.3. and 2.2.4., it seems clear that the percentage figures mentioned above as denoting the extent of basic vocabulary sharing

between languages can no longer be regarded as determining the degree of the interrelationship between communalects by themselves, but that the evidence provided by such figures has to be critically weighted in the light of the general similarity between the communalects involved which manifests itself in similarities on the lexical, structural, semantic and typological levels, and especially in formal and system similarities and differences on the level of complex verb forms. Also, it has to be kept in mind that the application of comparative methods, however preliminary, and the work leading to the establishment of cognate chains within large groups of related languages has shown that, in general, lexical evidence for the interrelationship of languages is in many instances of a higher order than has been believed to be the case.

While the evidence for establishing interrelationship, and its degrees, between communalects has shifted considerably from the lexicostatistical sphere into wider and more varied orbits, the terms denoting the nature and extent of the relationship between languages and language groups, and within the latter, continue to keep their validity. Some remarks elaborating on them may be useful.

"X-level relationship" refers to the degree of relationship existing between members of the group represented by X and refers to their internal, intra-group relationship. For instance, a family-level relationship is that between various languages within a family, a stock-level relationship that between various families (or other family-level members such as family-level isolates) within a stock, etc.

An X-level group is a group which constitutes X with regard to its external, inter-group relationship. For instance a family-level group is a family whose members show family-level relationship internally, but whose external relationship, as a group, to other groups on the same hierarchical level is a stock-level relationship. Similarly, the external relationship of a stock-level family to other stocks is on the phylum level. In other words:

Group	Constituent members	Internal relationship of members within group	External relationship of the group to other groups
dialect	sub-dialects	dialect-level	language-level
language	dialects	language-level	family-level
family	languages	family-level	stock-level
stock	families	stock-level	phylum-level
phylum	stocks	phylum-level	unrelated

In the light of the degree and nature of the differences and similarities between languages belonging to the various stocks entering into the composition of various Papuan phyla, it has been found useful to introduce the concepts of sub-phyla and super-stocks to allow for greater taxonomic flexibility.

The following definitions may be given for 'sub-phyla' and 'super-stocks':

The term 'sub-phylum' denotes a section of a phylum whose stock-level members show the usual phylum-level relationship to each other, but which as a group displays marked differences from other stocks within the phylum, be it through greater distance in relationship between the members of the sub-phylum and the remaining members of the phylum, or through special characteristics of the members of the sub-phylum which sets them apart from the remainder of the phylum and increases the difference between them and the other members of the phylum. The use of the term 'sub-phylum' stresses the difference between the members of a sub-phylum and other members of the phylum of which the sub-phylum forms a part, and brings the external relationship of its stock-level members into focus. Each of these stock-level members of a sub-phylum is in a sub-phylum-level relationship to other members of the phylum which are outside the sub-phylum and can be looked upon individually as a sub-phylum in contrast to them.

The term 'super-stock' refers to a number of stocks within a phylum which can be grouped more closely together than other stocks within the phylum, be it through greater relational proximity between the members of a super-stock than is observable with regard to other stock-level members of the phylum to which they belong, or be it because of the presence of special characteristics shared by the members of a super-stock which makes them more similar to each other than is the case with other stocks of the phylum. Membership of a particular stock to a super-stock does not necessarily affect its relational distance to other stocks within the same phylum which are not members of that super-stock. The term 'super-stock' brings the internal relationship of its stock-level members into focus. Hierarchically, a super-stock is subordinate to a sub-phylum, and a sub-phylum can consist of a single super-stock, or even a single stock, family, or language.

Following the table given above, the following can be added with regard to the super-stocks and sub-phyla:

Group	Constituent members	Internal relationship of members within group	External relationship of the group to other groups
superstock	stocks	super-stock-level	phylum-level
sub-phylum	stocks	sub-phylum-level	phylum-level

2.2.6. SOME ILLUSTRATIVE EXAMPLES OF THE APPLICATION TO DOUBTFUL CASES, OF THE CRITERIA DISCUSSED IN 2.2.4.

2.2.6.1. SKO STOCK LANGUAGES

Languages of the Sko Stock in the northern border area between Irian Jaya and the western Sepik District of Papua New Guinea, in particular Sko itself, contain a number of lexical items which are clearly proto-Trans-New Guinea Phylum reflexes, amongst them one verb (*sleep*). There are also three pronouns belonging to set I (see 2.3.3.2.) (the pronoun set which is by far the most prominent set in Trans-New-Guinea Phylum languages) including the very basic pronouns of the first persons singular and plural. At the same time, of the ten nouns mentioned in 2.2.4.1. A) c), *bone, eye, fire, louse* and *skin* are not proto-Trans-New Guinea Phylum reflexes, and *ear* is doubtful. On the basis of this lexical evidence, Sko could perhaps be looked upon as an aberrant member of the Trans-New Guinea Phylum. However, it has little in common with the languages of that phylum on the typological and structural levels, and shows typological and structural features in its basic verb structure which are entirely at variance with Trans-New Guinea Phylum characteristics. It has therefore been decided not to include the Sko stock in the Trans-New Guinea Phylum in spite of the quite sizable Trans-New Guinea Phylum lexical element in it which extends to verbs and pronouns, but which is being regarded as the result of the influence of Trans-New Guinea Phylum languages upon Sko Stock languages. At the same time, the links of the latter stock with members of other phyla are of a very much lower order than those which they show with the Trans-New Guinea Phylum, and the Sko Stock has therefore been classified as constituting an unrelated phylum by itself, the Sko Phylum (i.e. phylum-level Stock) (2.14.1.1.).

2.2.6.2. PAUWASI STOCK LANGUAGES

Members of the Pauwasi Stock in north-eastern Irian Jaya which Voorhoeve (1971) regarded as unrelated to the Trans-New Guinea Phylum,

have been found by Wurm to contain more lexical items which are proto-Trans-New Guinea Phylum reflexes than Sko. These items include one verb (*eat*), and two important set I pronouns (first and second person singular) are present. Of the ten nouns listed in 2.2.4.1 A) c), only *ear*, *fire*, *skin* and *water* are not proto-Trans-New Guinea Phylum reflexes. This is better lexical evidence than that presented by Sko which lacks four or five of the seven more important diagnostic nouns whereas the Pauwasi languages lack only two, and which contains the verb *sleep* as a Trans-New Guinea Phylum reflex whereas the Pauwasi languages have *eat* which is diagnostically more important (see 2.2.4.1.A) b)). Sko presents somewhat better evidence than the Pauwasi Stock languages on the pronominal level only. Unfortunately, the typological and structural characteristics of the Pauwasi Stock languages are not known, but on the basis of the good comparative lexical evidence, it has been tentatively decided to include the Pauwasi Stock into the Trans-New Guinea Phylum as a sub-phylic, i.e. aberrant, member.

2.2.6.3. KWOMTARI STOCK LANGUAGES

The languages of the Kwomtari Stock in the north-eastern part of the West Sepik District of Papua New Guinea contain about as many lexical items which are proto-Trans-New Guinea Phylum reflexes as is the case with Sko except that there are no verbs amongst them. There are also two set I pronouns (first and third person singular). Of the nouns listed in 2.2.4.1. A) c), *ear*, *louse*, and *skin* are not proto-Trans-New Guinea Phylum reflexes, and *eye* is doubtful. The structure of the languages is comparatively simple which in itself constitutes a deviation from the general Trans-New Guinea Phylum pattern, and neither striking agreements, nor disagreements, with Trans-New Guinea Phylum typological and structural characteristics and structural principles have been observed. Further studies may contribute towards the clarification of the problem of their classification, but on balance, it seems best to exclude the languages of the Kwomtari Stock from the Trans-New Guinea Phylum at present. The presence of Trans-New Guinea Phylum elements in them is regarded as attributable to Trans-New Guinea Phylum influence. Links between the Kwomtari Stock languages and members of other phyla are even less pronounced than those mentioned in 2.2.6.1. with regard to the Sko Stock, and the Kwomtari Stock has therefore been classified as constituting an unrelated phylum by itself, the Kwomtari Phylum (i.e. phylum-level Stock) (2.14.1.2.).

2.2.6.4. MIDDLE SEPIK STOCK LANGUAGES

The members of the Middle Sepik Stock in the Eastern Sepik District contain a number of proto-Trans-New Guinea Phylum reflexes in their vocabulary, amongst them one verb (*sleep*). Four set I pronouns are present including the important pronouns of the first person singular and plural. Of the nouns listed in 2.2.4.1. A) c), *arm, bone, ear* and *eye* are not proto-Trans-New Guinea Phylum reflexes. This is fair lexical evidence, and in addition, some of the structural features of the Middle Sepik Stock languages are similar to Trans-New Guinea Phylum characteristics, such as the presence of sentence-medial verb forms which differ according to whether the subjects of the medial and final verbs are identical or different. There are however also typological features which contrast with Trans-New Guinea Phylum characteristics such as the absence of classificatory verbs, the universal presence of a two-gender distinction in the third, and usually also the second, person singular pronouns, and phonological features, e.g. the paucity of vowel phonemes.

In general, the lexical and structural-typological evidence would make it feasible to suggest the possibility of a distant relationship of the Middle Sepik Stock languages to the Trans-New Guinea Phylum, and this was in fact proposed in Wurm 1971. However, more recent studies (Laycock 1973) have clearly demonstrated that a very much closer relationship exists between the Middle Sepik Stock languages and other languages of the Sepik area which has made it possible to include the former into the Sepik-Ramu Phylum as a prominent member of the Sepik Sub-Phylum (see 2.11.).

2.2.6.5. LANGUAGES OF THE RAMU AREA STOCKS

Languages of a number of stocks in the Ramu River area in the western Madang District such as the Annaberg, Ruboni and Goam Stocks contain a good number of proto-Trans-New Guinea Phylum reflexes in their vocabularies, including verbs, as well as one or two pronouns of set I in individual languages. This lexical evidence is a little better than that presented by other stocks further west such as the Pihom, Josephstaal and Wanang Stocks which have been included in the Trans-New Guinea Phylum on the sub-phylum level. However, some structural and typological features of languages of the Ramu area stocks are quite markedly at variance with those of Trans-New Guinea Phylum languages, and at the same time, their general relationship appears to be considerably closer to languages of the Sepik area. They have therefore

been included into the Ramu Sub-Phylum of the Sepik-Ramu Phylum (see 2.11.).

2.2.6.6. SENAGI FAMILY LANGUAGES

Languages of the Senagi Family in the Irian Jaya-West Sepik District border area which was regarded by Voorhoeve (1971) as a phylum-level family isolate contains a fair number of proto-Trans-New Guinea Phylum reflexes including one verb (*sleep*). Three set I (and Ia - see 2.3.3.5.) pronouns are present, including the second and third person singular pronouns. Of the ten nouns listed in 2.2.4.1. A) c), only *breast* is not a proto-Trans-New Guinea phylum reflex, but an Austro-nesian loanword, and *ear* is doubtful. Lexical agreements between Senagi Family languages and member languages of neighbouring Trans-New Guinea Phylum stocks such as the Border and Pauwasi Stocks are of a rather low order. On the structural-typological level, the languages seem to follow the general Trans-New Guinea Phylum pattern without major deviations (though the agreements are not striking either) as far as can be judged from the very limited material available, but there is no information at hand on the existence or otherwise of sentence-medial verb forms and classificatory verbs. Some reference to gender in the third person singular appears to be present in verb forms in at least one of the languages of the family - this would reflect a substratum phenomenon observable in many of the Trans-New Guinea Phylum languages in the centre and central south of the mainland. On balance, it seems quite justifiable to include the languages of the Senagi Family into the Trans-New Guinea Phylum. However, it has been tentatively assigned sub-phylum status in view of the paucity of the information available on it and the possibility that major structural and typological deviations from the Trans-New Guinea Phylum may become evident upon further study, and also in the light of the tenuous nature of its lexical agreements with neighbouring Trans-New Guinea Phylum stocks.

2.2.6.7. KOLOPOM (OR FREDERIK HENDRIK ISLAND) FAMILY LANGUAGES

The languages of the Kolopom Family on Frederik Hendrik Island in south-eastern Irian Jaya show rather low lexical agreement with languages of Trans-New Guinea Phylum stocks of their general area except for their immediate neighbours, but the number of lexical items in them which are proto-Trans-New Guinea Phylum reflexes is not inconsiderable and includes verbs, amongst them *say* and *sleep*. Also, four to five set

I pronouns are present in them, including those of the first and second persons singular and the first person plural. Of the ten nouns listed in 2.2.4.1. A) c), only *bone*, *ear* and *water* are not proto-Trans-New Guinea Phylum reflexes. This lexical evidence would constitute a fair basis for the inclusion of the Kolopom Family languages into the Trans-New Guinea Phylum, but these languages differ rather markedly from the other languages of the Trans-New Guinea Phylum on the structural level in having particles instead of affixes as they are usually met with in the languages of that phylum. However, the principles underlying the functions of these particles compare quite well with those relating to affixes in other languages of the phylum. Only a few features such as the marking of the plural with nouns by a special particle are typologically in direct contrast with phenomena present in the bulk of the Trans-New Guinea Phylum languages, though some of the special characteristics of the languages of that phylum such as classificatory verbs and sentence-medial verb forms, seem to be absent.

It seems that a strong substratum is present in the Kolopom Family languages, and that their ancestral language which was probably unrelated to the Trans-New Guinea Phylum languages, was subjected to strong influence by the latter and took over from them the Trans-New Guinea Phylum set I pronoun forms and system as well as a considerable amount of basic vocabulary including verbs. They also seem to have adopted some of the basic Trans-New Guinea Phylum typological principles without getting strongly assimilated to the phylum languages on the structural level.

In the light of the criteria of classification discussed in 2.2.4., it appears that there are better reasons for including the Kolopom Family languages into the Trans-New Guinea Phylum on the sub-phylum level than to exclude them from it. The situation shows some parallelism to that of the Sko Stock (see 2.2.6.1.) which has been excluded from the Trans-New Guinea Phylum, but the lexical evidence which would support its inclusion is much weaker than that present in the case of the Kolopom Family languages, and its structural and typological features are to a great extent, completely in contrast with those present in the Trans-New Guinea Phylum languages which is not the case with those of the Kolopom languages to any comparable degree.

Recent work by Voorhoeve (Voorhoeve 1975) has demonstrated that the languages of the Kolopom Family show a fair amount of lexical links with languages of the southern Vogelkop area which have also been included into the Trans-New Guinea Phylum and constitute the South

Bird's Head Sub-Phylum in it (see 2.6.2.3.4. in this volume).

Some of the connections between the two language groups appear to be attributable to a non-Trans-New Guinea Phylum element which is present as a strong substratum in both. This has some bearing on linguistic prehistory (see 3.4.1.).

2.2.6.8. PAWAIAN FAMILY LANGUAGE(S)

The relationship of the language(s) (see 2.7.5.3.) of the Pawaian Family in the central north of the Gulf District to those of neighbouring groups used to be regarded as only very distant and doubtful (Wurm 1971) because of what was believed to be the low level of lexical relationship observable between them, though the relatively high degree of typological similarity between the Pawaian and these other languages had been recognised (Wurm 1964). Recent studies have shown that the Pawaian language(s) contain(s) a fair number of proto-Trans-New Guinea Phylum reflexes, amongst them several verbs including the important verb *say*. Two set I pronoun forms are found, i.e. of the first person singular and first person plural which are both very important. Of the ten nouns listed in 2.2.4.1. A) c), *arm*, *ear*, *bone* and *Louse* are not proto-Trans-New Guinea Phylum reflexes. On the structural and typological levels, a number of differences from the usual Trans-New Guinea Phylum pattern are in evidence such as for instance the very low development of sentence-medial verbal forms. The Pawaian language(s) also show(s) strong influence of a substratum which is widespread in many languages of the Southern Highlands, Western and Gulf Districts, as well as of the adjacent parts of the Western Highlands and Chimbu Districts, and which manifests itself in Pawaian in the presence of an abundance of nasal vowels, a limited occurrence of bound subject markers and of sentence-medial forms with verbs, as well as of a considerable number of aspectual distinctions within it (see 2.3.2.3.). However, in other structural and especially typological fields, the Pawaian language(s) show(s) a comparatively high level of agreements with the usual Trans-New Guinea Phylum pattern, and it seems quite justifiable to include the Pawaian Family into the Trans-New Guinea Phylum on the sub-phylum level. It may be pointed out that MacDonald (1973) suggests the existence of closer relational links between the Pawaian Family language(s) and those of the Teberan Family which is another sub-phyletic member of the Trans-New Guinea Phylum. In view of this, it has been decided to include the Pawaian Family with the Teberan Family into a sub-phylum-level superstock.

2.2.6.9. THE OKSAPMIN ISOLATE

The Oksapmin language located in the central "hub" area of the mainland shows quite high cognation percentages with languages of the neighbouring Ok Family of the Central and South New Guinea Stock of the Trans-New Guinea Phylum, but these percentages decrease rapidly in direct proportion to geographical distance, and Healey (1964) looks upon them as the result of heavy borrowings in comparatively recent times and assumes that Oksapmin is genetically unrelated to the Ok Family. Almost all the proto-Trans-New Guinea Phylum reflexes in Oksapmin are obviously closely cognate with Ok Family forms. Structurally and typologically it differs to some extent from the Ok Family languages, but not strikingly, and shows some deviations from the usual Trans-New Guinea Phylum pattern, but not to a very great extent. Its position remains doubtful and it may well be classifiable as a phylum-level isolate which has been very strongly influenced by Trans-New Guinea Phylum languages, especially of the Ok Family, with this influence being so recent that it is still recognisable as a foreign element, at least on the lexical level. At the same time, it is not possible to say whether the relative similarity of Oksapmin to the usual Trans-New Guinea Phylum pattern on the structural and typological levels is attributable to such an influence, or is the result of a possible relationship of Oksapmin to the phylum, or is attributable to chance. Added weight to the classification of Oksapmin as a phylum-level isolate seems to be given by Laycock's (1973) suggestion that Oksapmin may be related to the Yuri Isolate situated further north-west in the Western Sepik District. Yuri is apparently unrelated to any other language and not a member of any of the established phyletic groups.

Oksapmin used to be classified as a doubtful member of the Trans-New Guinea Phylum (Wurm 1971), and until the situation has become clearer as a result of further study, this tentative classification will be adhered to, on the understanding that as an alternative, Oksapmin may constitute a small phyletic group with Yuri and be regarded as not related to the Trans-New Guinea Phylum.

BIBLIOGRAPHY

CAPELL, A.

- 1962 'Oceanic Linguistics Today', *CAnthr* 3/4.371-429.

DUTTON, T.

- 1971 *Magori - its Nature and Relationship to other languages in South-East Papua*. Paper presented at the Linguistic Society of Papua New Guinea in August 1971. Mimeographed.

FRANKLIN, K. ed.

- 1973 *The Linguistic Situation in the Gulf District and Adjacent Areas, Papua New Guinea*. PL, C.26.

GRACE, G.

- 1965 'On the Scientific Status of Genetic Classification in Linguistics'. *OL* 4.1-14.

HAMP, E.

- 1962 Comment on Capell 1962. *CAnthr*. 3/4.220-1.

HEALEY, A.

- 1964 *A Survey of the Ok Family of Languages*. Ph.D. Thesis Section, Australian National University. Canberra. Mimeographed.

LAYCOCK, D.C.

- 1973 *Sepik Languages - Checklist and Preliminary Classification*. PL, B.25.

MacDONALD, G.E.

- 1973 'The Teberan Language Family'. Franklin ed. 1973.111-48.

McELHANON, K.A.

- 1970a *The Selepet Language within the Finisterre-Huon Phylum (New Guinea)*. Ph.D. Thesis, Australian National University. Canberra.

- 1970b 'Lexicostatistics and the Classification of the Huon Peninsula Languages'. *Oceania* 40.214-31.

- 1971 'Classifying New Guinea Languages'. *Anthropos* 66.120-44.

McELHANON, K.A. and C.L. VOORHOEVE

- 1970 *The Trans-New Guinea Phylum: Explorations in Deep-level Genetic Relationships*. PL, B.16.

PAWLEY, A.

- 1969 *On the Internal Relationship of Eastern Oceanic Languages*. Paper presented at the Conference on Oceanic Prehistory in Fiji, in August 1969. Mimeographed.

RAY, S.H.

- 1911 'Comparative Notes on Maisin and other Languages of Eastern Papua'. *JRAI* 41.397-405.

SEBEEK, T.A. ed.

- 1971 *Current Trends in Linguistics, vol.8: Linguistics in Oceania*. The Hague. Mouton.

STRONG, W.M.

- 1911 'The Maisin Language.' *JRAI* 41.381-96.

SWADESH, M.

- 1955 *American non-cultural vocabularies*. Revised ed. Mimeographed.

THOMAS, D. and A. HEALEY

- 1962 'Some Philippine Language Subgroupings: a lexicostatistical study'. *AnL* 4/9.21-33.

VOORHOEVE, C.L.

1971 'Miscellaneous Notes on Languages in West Irian, New Guinea'.
PL, A.28.47-114.

1975 *Languages of Irian Jaya: Checklist*. *PL*, B.31.

WURM, S.A.

1964 'Australian New Guinea Highlands Languages and the
distribution of their typological features'. *AmA* 66/4,
part 2.79-97.

1969 'The Linguistic Situation in the Reef and Santa Cruz Islands'.
PL, A.21.47-105.

1970 'Austronesian and the Vocabulary of the Languages of the Reef
and Santa Cruz Islands - a preliminary approach'. Wurm and
Laycock, eds. 1970. 467-553.

1971 'The Papuan Linguistic Situation'. Sebeok, ed. 1971. 541-657.

WURM, S.A. and D.C. LAYCOCK, eds.

1970 *Pacific Linguistic Studies in Honour of Arthur Capell*.
PL, C.13.