THE LANGUAGES OF SULAWESI

J. NOORDUYN

If one studies the language situation on the island of Sulawesi and the smaller islands surrounding it, one cannot but be struck by the extraordinary wealth of languages found there, especially if compared with its total population on the one hand, and with the situation in some of the other large islands of the Indonesian archipelago on the other.1

On the island of Java, with a population rapidly approaching one hundred million, only 4 languages are spoken: Javanese, Madurese, Sundanese and Jakarta Malay (and perhaps a few more if some considerably divergent dialects, such as Cirebon, are classified as separate languages). The much larger but far less populous island of Sumatra and adjacent islands is home to some 11 or, if the 6 Batak dialects are recognised as separate languages, 16 languages (cf. Voorhoeve 1955). In contrast, the number of languages spoken in Sulawesi and environs, with a population of no more than seven million people, is estimated to be around 80.

A few examples will suffice to illustrate this situation. In the province of North Sulawesi there is the narrow peninsula, called Minahasa, with a population of a few hundred thousand people, where 8 languages are spoken: 5 so-called Minahasan languages, 2 Sangiric languages and Manado Malay, which, though a Malay dialect if viewed from outside the region, functions as an independent language within the Minahasa context. In the tiny Sangir-Talaud archipelago 2 more Sangiric languages are found, and the Gorontalo-Mongondow group of languages to the west of Minahasa comprises 8 more languages, from Mongondow in the east to Buol in the west, just across the border with the province of Central Sulawesi, which makes the total number 18.

The situation in South Sulawesi is different in that it has a far denser population of over four million people. It contains not only the largest urban complex of Sulawesi (Ujung Pandang) and the surrounding Makasar region, but also the homeland of the Buginese sailors and the mountains of the picturesque Toraja people. Each of these has its own language, Makasarese, Buginese and Sa'dan Toraja respectively. But it may come as no surprise that a lexicostatistical survey of the entire South Sulawesi province, conducted by Charles and Barbara Grimes of the Summer Institute of Linguistics in 1983 (published 1987), reported 20 different languages of the South Sulawesi group plus 4 of the Muna-Buton group, making 24 altogether. These range from (at one extreme) the large group of some two and a half million people speaking a highly homogeneous Buginese with only a few dialect areas near the borders, to (at the other extreme): the small group of people speaking 2 non-contiguous dialects (Laiolo and Barang-Barang) of a Muna-Buton language on the southern tip of the oblong

---

1An earlier version of this paper was presented at the Sixth European Colloquium on Indonesian and Malay Studies, held in Passau, Germany, 22-27 June 1987. I am greatly indebted to Dr Jack Prentice for his useful comments on an earlier version of this paper and for his efforts to correct my English.


© J. Noorduyn
island of Salayar; some four and a half thousand Seko people living in the high mountains to the north of the Sa'dan Toraja (separated from the latter by 50km of uninhabited forest) who speak 3 Seko dialects (or languages) with 62-75 per cent mutual cognate relationship (Laskowske and Laskowske 1987:88,96); and the highly complex situation in the mountain region called Pitu Ulunna Salu to the north of the Mandar coast, where some forty thousand people, living in seventeen desas, speak some 14 dialects, which can be grouped into 3 languages, each consisting of a number of dialects connected by a chaining relationship. In this case the chaining means that “all relate to the neighbouring dialects at 89-90%, while the extreme points in the chain only relate at 84-85%” (Strømme 1985:11-12).

It should be remarked here that these numbers of South Sulawesi languages are partly dependent on the critical percentage used in determining what is identified as a discrete language and what as one of several dialects of the same language. If 70 per cent of cognacy is used as the ‘language limit’ (as is done by Dyen (1965:18) in his lexicostatistical classification of the Austronesian languages) instead of 80 per cent, then the 3 discrete languages of the Makasar subgroup become 1 Makasar language with 2 additional dialects (Konjo and Salayar); similarly the 5 languages of the Toraja-Sa’dan subgroup become 1 Sa’dan Toraja language with 4 additional dialects, and the total number of South Sulawesi languages is 14 rather than 20. I follow Grimes and Grimes (1987) in using the 80 per cent border line between language and dialect however, as this percentage has been part of lexicostatistics from the start, even though it is to a large extent an arbitrary decision. This does not mean to say that lexicostatistics should be regarded as the only method for determining this border line, nor that its results should be accepted without reserve. Some critical remarks will be made later on in this essay.

The situation in the province of South-East Sulawesi has its own peculiarities which fit in with the general pattern. On the mainland there are two languages spoken in well-defined areas: Tolaki with some five hundred thousand speakers, and Moronene, also spoken on the island of Kabaena, with some forty-five thousand speakers. On the other islands of the province one usually finds a single language per island or island group: Munanese is spoken by two hundred thousand people on the island of Muna, Wawonii by sixteen and a half thousand on the island of Wawonii, and a language with no generally accepted name by sixty thousand people on the four islands of the Tukang Besi group.

In contrast, the language situation on the island of Buton appears to be extremely complex, and as new data become known they only tend to add to the complexity of the picture. Here it is still possible to discover hitherto unknown languages.

In the early 1930s it was discovered that what had been known as Butonese was, in fact, the language called Wolio, which was spoken only in the Sultan's residence and a few villages surrounding it, but was also used as the official language of communication with the other parts of the Buton sultanate (Cense, ed. 1954:165). Outside the tiny Wolio area quite different languages are in daily use. Most vernaculars spoken in central and southern Buton, apart from Wolio, seem to be dialects of two languages: Central Butonese or Pancana, and South Butonese or Cia-Cia. But beside these there are at least four other languages (sharing 40-50 per cent cognates with each other), two in the north-east and two in the south-west. One of the latter has only recently been discovered (together with another in northern Buton and a third (Kaimbulawa) on the island of Siompu) in a brief survey conducted in only a part of the region (van den Berg 1988:3,15-17). The complexity of the situation is demonstrated by the fact that occasionally four different languages are found to be spoken in the same village. It is clear that only a thorough and complete survey of the whole area can establish the real situation.
Altogether there are at least 15 different languages in the whole of the province of South-East Sulawesi.

The area with the largest number of languages is the province of Central Sulawesi. It is in this region that, in the first quarter of this century, linguistic work was carried out by Adriani, who in his time was one of the greatest linguists in the field of Indonesian linguistics although he is seldom mentioned nowadays. He not only made an extensive study of the language spoken in the area of Lake Poso (then known as “Bare’e”, but which its modern speakers prefer to call “Pamona”) and posthumously published a voluminous dictionary, grammar and texts (Adriani 1928,1931,1932-33), but also surveyed the whole area covering the present-day province of Central Sulawesi and published a meticulous report of the language situation with a detailed language map, which includes the greater part of the entire island of Sulawesi (Adriani 1914). If the situation as described by Adriani is compared with those presented by recent surveys, such as that by Barr, Barr and Salombe (1979), it can be seen that essentially there is no difference. For instance, Adriani mentions 9 Tomini languages (1914:348-350), and Barr, Barr and Salombe (1979:23) 8, but the latter add that they had not obtained data from Petapa, which was described by Adriani. They furthermore list Dondo, which is not mentioned by Adriani, as a separate language, though admitting that it has 80 per cent shared cognates with Toli-Toli and may thus also be considered a dialect of the latter (1979:28-29). They furthermore list Lauje and Tialo, which in Adriani's presentation (Adriani 1914:348-349) are 2 languages called Tinombo and Tomini, as 2 dialects of Tomini sharing 89 per cent cognates (Barr, Barr & Salombe 1979:30). The differences clearly lie more in interpretation of the data than in the actual language situation. It may be concluded that there are 9 or 10 Tomini languages.

Slightly different is the situation of the languages which were called West Toraja in the time of Adriani, Kruyt and Esser, but at present are preferably, though less informatively, called the Kaili group. All 12 languages listed by Adriani as West Toraja still seem to be spoken today. Even Tawaelia, which in Adriani's time had only three hundred speakers (Adriani 1914:108), apparently still exists, as it turns up as Sedaq with six hundred speakers in a recent survey (Barr, Barr & Salombe 1979:34,96-98). One of the results of Esser's research into the languages of this region in the 1930s was that many of the vernaculars which according to Adriani were separate languages proved to be dialects of one language, which was called Kaili. Esser's conclusions can be found in his language map (1938 sheet 9b), in which only 3 of Adriani's 12 West Toraja languages are mentioned (Kaili, Kulawi and Pipikoro), while more detailed information, on the basis of notes supplied by Esser, is contained in A.C. Kruyt's (1938) four-volume publication about the West Toraja. Surveys conducted in the 1970s have confirmed that most of Adriani's languages must be

---

1This is Volume III of the collection of volumes which Adriani wrote jointly with A.C. Kruyt on the Bare’e-speaking of Central Celebes (1911-1914), but I refer to it as ‘Adriani 1914’ because he, being the linguistic professional of the two, was the sole author of this third volume, whereas Kruyt, the ethnologist, wrote the other two volumes. Later on, Kruyt published a revised and much expanded version of his volumes. This second edition also consists of three volumes, but these were all written by Kruyt and did not include the language volume. Adriani 1914 was never revised or reprinted.

2The language named by Adriani ‘Petapa’, ‘Tadje’ or ‘Andje’ (the last two names being words for ‘no’) was spoken by a tiny community within the Kaili dialect area, Petapa being a village located due south-west of Pelawa near Parigi on the Tomini coast. Unlike the Kaili-Pamona languages it had consonants occurring in word-final position (ng, s, t, l, r, g) and thus was classified by Adriani as belonging to the Tomini subgroup (Adriani 1914:9,169,179-180,349). According to a recent study (Kaseng et al. 1979:66-99) it has in the meantime lost these final consonants (e.g. 1979 langit, ompo, uja ‘heaven’, ‘belly’, ‘rain’ respectively) and is considered to be possibly a Kaili dialect. This loss of final consonants is a clearly attested example of how a language may adapt to its linguistic environment by borrowing structural features, in other words, how an areal feature is spreading.

3Esser's findings as contained in Kruyt's book and in his own quarterly reports are given in Noorduyn ed. (1963:338-352). They show that Esser, far from listing 12 different West Toraja languages (as alleged in Barr, Barr & Salombe 1979:30), was well aware of the linguistic situation. In his opinion, in the region north of Kulawi there were only dialects
considered dialects or subdialects. That is why Barr, Barr and Salombe (1979:23-24), too, list only 4 languages in the Kaili area, namely Kaili (7 dialects, among them Kulawi), Pipikoro (Uma), Lindu (Tado) and Sedoa (Tawaelia). Even this total is one too many since Lindu, which they included without having seen any data from it (Barr, Barr & Salombe 1979:33), had already been found to be a Kulawi dialect by Esser in 1935 (Noorduyyn ed. 1963:345). This was confirmed by Dyen, who gave 86.9 as their common cognacy percentage (Dyen 1965:28), and by Wumbu et al. (1983:46-55). In the Language atlas of the Pacific area, Lindu nonetheless figures as a separate language (Wurm & Hattori 1981-1983 no.43).

To these should be added 3 languages spoken in the mountains between the western and the eastern Toraja area (Napu, Bada'/Besoa, and Rampi' or Leboni), and Adriani's Bare'e, now called Pamona, in the Poso or eastern Toraja region.

In the eastern part of this Central Sulawesi province, there are 5 languages spoken on the eastern peninsula and surrounding islands. One of these, called Andi'o and spoken by some one and a half thousand people in the easternmost Balantak area, has only recently been discovered. There is an Andi'o list of 100 words in Barr, Barr and Salombe (1979:102-104) and another, differing from the latter by ± 50 per cent, in an unpublished survey report (Wumbu et al. 1983:135-138). It should be noted that this is not the same language as the one called Bobongko, despite Barr, Barr and Salombe's statement (1979:36) that they chose to use the name Andi'o in place of Bobongko because the latter has somewhat derogatory overtones. Bobongko is still spoken by a small group of people on one of the nearby Togian islands (Wumbu et al. 1983:16,29,84-87) as it was when Adriani first reported about it (1900:429-460), and clearly differs from the Andi'o spoken in Balantak, witness such vocabulary items as bagu and bu'o 'new' (Wumbu et al. 1983:56,137). In Wurm and Hattori (1981-1983 no.43) both are mistakenly given the name Andi'o as if they were one language. The 3 other languages spoken here are Saluan, Balantak and Banggai.

With 2 more languages spoken on the east coast of Central Sulawesi (Mori and Bungku), the sum total for this province amounts to 25.

Not surprisingly, the number of speakers of most of these languages is small to very small, not exceeding some thousands or tens of thousands. Only 2 languages have more than one hundred thousand speakers: Pamona and its dialects (c.110,000) and Kaili and its dialects (c.300,000), nearly half of the latter being speakers of the Palu dialect.

It we now complete the picture by totalling the numbers mentioned so far for the languages in the various islands and provinces of Indonesia, it has been posited that the number of languages spoken in Java is 5 and in Sumatra 16, which makes 21 in total. In contrast, the number of languages spoken in Sulawesi is more than three times as high: in the province of North Sulawesi 17, in Central Sulawesi 25, in South Sulawesi 23, and in South-East Sulawesi 15, which amounts to a total of 80. Of course, these figures are not to be considered definitely established and exactly comparable, because the principles and methods used to determine them, as well as the degree of sophistication of the linguistic research which produced them, may well have differed from area to area. They may nonetheless serve to illustrate the great contrast between the linguistic situation in Sulawesi and that in some of the other large islands.

Essentially the same difference also manifests itself in the matter of subgrouping. This is most strikingly demonstrated in the language map and survey of languages in Indonesia which were of 1 language, which he called Kaili, consisting of West Kaili, Central Kaili and East Kaili dialects, and he was inclined to consider Kulawi a South Kaili dialect.
published by Esser (1938) in the *Atlas van Tropisch Nederland*. Esser's survey, covering the whole of Indonesia, has 16 groups of Austronesian languages, among them 1 for Sumatra, 1 for Java and 1 for Borneo, but then 8 for Sulawesi, and another 4 for the rest of Indonesia, making 8 out of 15, or more than half, for Sulawesi. One should bear in mind, moreover, that the number of groups distinguished by Esser was already a reduction compared to that given by Adriani in 1914. Adriani had 11 groups for Sulawesi plus a few still unclassified languages. He had a West Toraja and an East Toraja group, which Esser combined into one Toraja group, and, secondly, Esser was the first to have one South Sulawesi group, which included 3 of Adriani's groups: the Makasarese-Buginese, the Sa'dan Toraja and the Mandarese.

Though a few changes have been made to Esser's list of language groups since he published it, and apart from Dyen's grand lexicostatistical classification (Dyen 1965), which has not had much influence, it is remarkable that after fifty years Esser's description of the language situation in Indonesia has remained essentially unchallenged until today, the more so because Esser himself was far from satisfied with the data he put into his map. This is quite clear from his comments to be found in one of his quarterly reports, from which some quotations were published only twenty-five years afterwards (Noorduyn, ed. 1963:335-336). In view of recent discussions on the subgrouping of western Indonesian languages (Blust 1981:461; Nothofer 1985:298), it may be interesting to cite some of Esser's casual remarks, translated from the Dutch.

"This language map", he says in 1931 when he was drafting the map, "gives an incorrect impression in that several things are included in it which are by no means incontestable, but that cannot be avoided in such a kind of work" (Noorduyn, ed. 1963:334). In 1938 he gives an example: "This subgrouping is still very preliminary, especially as regards the western part of the Archipelago. It will probably be more satisfactory to divide the Sumatra group into three, namely, 1. Acheh, 2. the languages to the west of Sumatra, and 3. the Malay group, while to the latter also Sundanese and some of the Dayak languages will have to be added. But as long as still so little is known about the borders in Borneo, it seems better provisionally to keep the old subgrouping (which I have adopted from older maps)" (p.336).

In these brief notes Esser naturally did not go into his reasons for preferring one classification over another nor into his subgrouping criteria. Another brief remark shows what he had in mind in this respect and how confident he was in the results obtained. After a brief visit to Bali in 1937 and a few days of discussion with Goris, who had studied the languages of Bali and Lombok, Esser concluded that "as to Balinese, it appeared that this should be detached from the group of Java languages and joined to that of Sasak and Sumbawanese, as it shares with the latter not only a great many words but also several grammatical particulars" (p.336). Not surprisingly this conclusion has not been confirmed by Dyen's lexicostatistical data; he considers Sasak and Balinese rather to be coordinate with Sundanese (Dyen 1965:49).

Since the classification of the Sulawesi languages which is currently in use chiefly stems from Adriani, it seems worthwhile to investigate his subgrouping criteria. This is not easy since Adriani is seldom explicit about his criteria. His usual method of comparing languages is to give a sketch of each of them by mentioning various features of their phonology, word structure, vocabulary and grammar, and to present these sketches as such with their differences and similarities as the basis for arriving at a certain subgrouping in a rather impressionistic way. Often his subgrouping decisions are given rather arbitrarily either after some weighing of the arguments, pro and contra, or with no arguments at all. His classifications often tend to emphasise that the borders between adjacent groups are vague and less rigid than previously thought, and the differences are gradual rather than clear-cut.
An illuminating example is Adriani's opinion on the presence or absence of tense forms as a subgrouping argument. There are West Toraja languages which have verbal prefixes with initial \( n \)-denoting past and present tense and with initial \( m \)-denoting future and optative tense, whereas East Toraja languages such as Bare'e or Pamona do not have such forms. Adriani (1914:102-108) reduced the importance of this difference, which previously had been considered a strong and significant dividing line between two large subgroups of languages, by using two arguments. First, he demonstrated that these two groups of Toraja languages are so closely related in other respects, lexically, phonologically and morphologically, that having tense forms or not cannot be such a decisive criterion. His second argument was a historical one. The past forms with initial \( n \)-such as \( na \)-had already been explained as abbreviations of compound prefixes such as \( mina \)-, consisting of a prefix \( ma \)- and an infix \(-in\)-, which are known in Philippine languages as past forms. The formal tense difference was preserved in these Toraja languages after this abbreviation took place because the \( na:-ma \)-contrast remained. But in languages in which the infix \(-in\)-had earlier changed into a prefix \( ni \)-, whereby the compound prefix became \( nima \)-, the tense difference disappeared when a similar process of abbreviation reduced this prefix to \( ma \)-. Examples of closely related languages, one of which has prefix \( mina \)-for past marker, the other having \( nima \)-, are known to occur in Minahasa (cf. Sneddon 1978:89-90). This ingenious diachronic argument serves to explain not only the difference between languages with and without such tense forms but also why this difference must not be considered an overriding subgrouping criterion.

This conclusion of Adriani's has been completely confirmed by recent lexicostatistical research. Calculations by Barr, Barr and Salombe (1979:26-27) on the basis of 100-item wordlists have established that there is one Kaili-Pamona subgroup, comprising all of Adriani's West and East Toraja languages, with a mutual relationship of upwards of 61 per cent shared basic vocabulary (with the 7 Kaili dialects having a relationship of 82 to 91 per cent), which coordinates with a similarly composed Tomini subgroup.

Not surprisingly in view of this multilingual area, lexicostatistical methods (establishing the percentages of shared cognates in basic wordlists) have been used fairly extensively for a first exploration of the complicated situation. In both South and Central Sulawesi and in the Minahasa area the relationships have been determined by lexicostatistics. Both the strengths and the weaknesses of the technique have come to light in the results.

On the one hand, the application of statistical methods has produced much more rigidity in the results than before. Among other things, different levels of relationships have been recognised. Sneddon has, for instance, established that the 5 Minahasan languages consist, first, of two groups: Tonsawang on the one hand, and the 4 others, called North Minahasan languages by him, on the other. The latter consist of Tontemboan on the one hand, and the 3 others, called North-East Minahasan by him, on the other (Sneddon 1978:20-69). The 5 languages of Sneddon's Sangiric subgroup are similarly arranged (Sneddon 1984), and the 6 Gorontalic languages are directly related on a higher level with the group of Mongondow and Ponasakan (Sneddon & Usup 1986:410). The next item on the agenda is clearly to establish the relationship between the Minahasan, Sangiric and Gorontalic/Mongondow subgroups. A study to this end has hardly begun. Sneddon (1984:12) collected a small amount of evidence, in the form of shared innovations, for the recognition of a Sangiric-Minahasan group by listing nineteen examples of Proto Sangiric and Proto Minahasan cognates – and one can find a few more – though without making reconstructions for Proto Sangiric-Minahasan. As yet, he states, comparative study has not been detailed enough to enable a systematic statement of shared phonological innovations in the two groups. It therefore seems premature when
in the *Language atlas of the Pacific area* (Wurm & Hattori 1981-1983 no.43) the Minahasan and Sangiric subgroups are combined into one supergroup. Still more difficult is the combination of these two with the Gorontalic/Mongondow subgroup, even though there are a few exclusively shared innovations, such as PMin, PSan, PGM *bungang* 'flower' and *uala* 'canine tooth', and more can be found, such as PGM *watok* 'step(-father etc.)', Sangir *uataq*, Tontemboan *tolo/atek* and Tombulu, Tonsea *lolo/atek*.

Obviously, lexicostatistics and comparing lists of 100 or 200 words as the only material are only the first steps to subgrouping. Subsequently complete lexicons and morphological elements must be compared and considered and, as Sneddon has shown, reconstructions of proto-words, such as Proto Minahasan or Proto Sangiric, must be carried out.

On the other hand, the greatest risks of arbitrary decisions and downright mistakes are, as is well known, especially contained in the basic activities of collecting the wordlists and establishing shared cognates. Besides these things, the greatest enemy is borrowing, which introduces the problem of how to eliminate inconspicuous loan words. An example of how confusing the results can be is Kulawi, one of the Kai dialects of the West Toraja area, which nonetheless shows 86 per cent shared basic cognates with the neighbouring language Pipikoro, presumably through borrowing (Barr, Barr & Salombe 1979:34).

Another example, showing not only how basic words (that is words belonging to the core-lexicon, normally regarded as less subject to external influence, as envisaged in the 100-word and 200-word Swadesh lists) can be eliminated by borrowed ones, but also how real relationships can be obscured in the process, is Makasarese. There the words *baji?’ good*, *tinggi’ high* and *tuju’ seven* are basic words which are clearly borrowings from Malay; but *baji?’ has obviously pushed aside (and out of the basic wordlist) the Makasarese *pia*, now meaning ‘cured, healed’, and in this way has lowered the statistical relationship with Mandarese, Sa’dan Toraja, Mongondow, Sangir etc., which all show *mapia’ good*.

A clear example of the risks involved can be seen by comparing the relationship matrix for the South Sulawesi group of languages presented by Mills (1975:492,498) with that presented by Grimes and Grimes (1987:19). Not only do the percentages in the two matrices differ to some extent (although not significantly) but, even though they agree in positing only one subgroup within the South Sulawesi group (called proto-Sa’dan by Mills and Northern South Sulawesi by Grimes and Grimes), they also partly disagree as to which languages are included in this subgroup (according to Mills it includes Pitu Ulunna Salu, Mamuju, Seko, Sa’dan and Massenrempulu, whereas Grimes and Grimes include Mandarese and exclude Seko). Presumably the presentation of Grimes and Grimes is an improvement on Mills because of the larger quantity of their material, though their own conclusions have partly been improved again by new material from more recent surveys. This does not mean that Mills’s book of 900 pages about the South Sulawesi languages is useless. In spite of many unfounded speculations, unfair judgements and egregious errors, it contains much valuable material and many intelligent remarks and useful results – his 300 pages of Proto South Sulawesi reconstructions is an especially valuable collection.

On the basis of these examples it may not be surprising that the situation within the large subgroup of South Sulawesi is still far from clear because the material available is still decidedly insufficient.

Despite what has been said above, it is possible to mention here an example of a clearly successful application of the lexicostatistical method as a simple way of determining the relationships between languages. This is the case of the Wotu language, which is spoken by a few thousand people in two
villages on the northern coast of Bone Bay in a region bordering on the Buginese area to the west, the Pamona (Poso) area to the north and the Mori-Bungku languages to the east. The position of Wotu has long been uncertain, in the first place because what is known about it has not exceeded a few wordlists and secondly because there could not be found an obvious relationship with any of its neighbouring languages. Adriani (1914:90-91) came to the conclusion that Wotu showed some similarity to both Buginese and East Toraja, but on the basis of its vocabulary could not be classified as one of the Toraja languages; he was of the opinion that it was a transitional language between Bare'e and Buginese but should be classified in one subgroup with the latter. Esser (1938), on the other hand, put Wotu in the Toraja group as a separate subgroup next to the West and the East Toraja languages. However, a few years later he changed his opinion, after he had made a close study of the language, and concluded that it belonged to the Buginese group (Noorduyn, ed. 1963:356). Mills (1975:134,604-612) however, writing about the South Sulawesi languages and studying the little Wotu material published by Adriani, came to the conclusion that it does not belong to this group but did not propose an alternative.

The best solution to the problem appears to be that resulting from the lexicostatistical survey by Grimes and Grimes (1987). They found that Wotu shared its highest percentage of cognates not with any of the neighbouring languages but with Laiolo/Barang-Barang in southern Salayar, at a great distance from the Wotu settlement itself.

It is indeed striking that all groups of languages recognised as subgroups in Sulawesi appear to be most closely related to adjacent subgroups, and within each subgroup often the languages situated near its border appear to be most similar to languages located nearby across this border. As a result the entire language situation to a large extent gives the impression of gradual transitions, as Adriani rightly stressed, which can only partially be explained from the influence of borrowing and areal features. This overall picture evidently has not stimulated researchers to look farther away for closely related languages in a problematic case such as Wotu.

In this case, it is true that before the survey of Grimes and Grimes, Esser had already pointed to two other possibilities of close connection with Wotu: firstly, the Ledo or Palu language of western Central Sulawesi, which he had studied himself and which has a negative ledo strikingly similar to the Wotu negative laedo and, secondly, the Wolio language on the island of Buton, both at a considerable distance from Wotu. When Esser wrote this, his main arguments were local Wotu traditions, according to which Wotu people had formerly emigrated both to Palu and to Buton (Noorduyn, ed. 1963:356-359). Although he did not explicitly mention Laiolo on the island of Salayar, the latter was nonetheless implicitly also included in the Wolio connection.

It had been known for some time that Laiolo was quite different from the other languages of South Sulawesi. This had already been noted by Jonker, who did not know, however, to which other language(s) it might belong. A close relationship with Buton had already been mentioned by van der Stok more than a century ago (1865:423) and Esser (1938) classified it as one of the Muna-Buton languages on his map. But no material has been available until quite recently to substantiate these claims. With the help of wordlists recently published (Grimes & Grimes 1987, Wotu and Barang-Barang; Stokhof 1984, Layolo) and even a dictionary (Anceaux 1987, Wolio), the necessary comparisons can now be made. They presumably show that the basic vocabularies – with a strong emphasis on the word ‘basic’ – are closely related. This can be confirmed by a characteristic morphophonemic feature contained in my own Barang-Barang material which I collected a few decades ago in Makassar. The third person object suffix of the transitive verb is -a in both Barang-Barang and Wolio (Anceaux 1952:27); this suffix has in both languages an allomorph -ea which is
used with stems having -a as final vowel. This allomorph is also used when the definite suffix -mo or the irrealis suffix -po occurs at the same time, in which case the vowel o of these suffixes is replaced by -ea.

It is curious – though not surprising – to note that it was the same Holle wordlist, collected by Koopman in 1897 and published recently by Stokhof (1984:199-211), that was used in establishing the relationship with Wolio above and, as the only Laiolo material at his disposal, formed the basis for Adriani's declaration that he had no hesitation at all in classifying Laiolo material in the same group as Buginese. The reason for this rather incautious opinion of Adriani's is not hard to find. Both the wordlist in question and my own material not surprisingly show that the Laiolo/Barang-Barang vocabulary has undergone a massive influence from the neighbouring Makasarese dialect of the other inhabitants of Salayar and, via the vocabulary, the sound system, word structure and even some of the pronominal prefixes also show this Makasarese influence; for example, the pronominal prefix of the third person has the exceptional form of la- in both Laiolo/Barang-Barang and the Salayar方言 dialect of Makasarese. In spite of this large number of Makasarese loan words in its lexicon, there can be no doubt that Laiolo does not belong to the South Sulawesi group but to the same group as Wolio.

Laiolo and Wolio have also clear lexical correspondences with Wotu and with languages of western Central Sulawesi such as Palu and Pipikoro. Some of the most diagnostic examples are the following:

<table>
<thead>
<tr>
<th></th>
<th>Laiolo</th>
<th>Wolio</th>
<th>Wotu</th>
<th>Palu</th>
<th>Pipikoro</th>
</tr>
</thead>
<tbody>
<tr>
<td>leaf</td>
<td>tafa</td>
<td>tawa</td>
<td>(dau)</td>
<td>tawa</td>
<td>(rau)</td>
</tr>
<tr>
<td>person</td>
<td>ito</td>
<td>(mia)</td>
<td>ito</td>
<td>ito</td>
<td>(tauna)</td>
</tr>
<tr>
<td>wind</td>
<td>ngalu</td>
<td>ngalu</td>
<td>(angi)</td>
<td>(poi)</td>
<td>ngolu</td>
</tr>
<tr>
<td>armpit</td>
<td>keke</td>
<td>keke</td>
<td>?</td>
<td>keke</td>
<td>(kirikiq)</td>
</tr>
</tbody>
</table>

One thing must be evident from what has been said so far: an enormous amount of language material is to be found on the island of Sulawesi and much of it is still awaiting collection and investigation. There are questions of a more general purport which obviously cannot be answered until very much more material has become available for systematic study than is accessible at this moment. One of these questions is the most general one in the present context: is there a Sulawesi supergroup comprising all languages spoken on the mainland and surrounding islands, and all subgroups mentioned so far? Or, in other words, was there once a Proto Sulawesi language?

Although it is clearly impossible at present to give an answer to this question, or even to broach a real discussion of this problem, it will inevitably have to be posed and discussed at some time in the future. Nevertheless I wish to bring forward two questions which may be raised in this context as a preparation to a discussion of the larger problem.

The first question, which immediately arises, is to what extent the languages in northern Sulawesi may be counted as belonging in some way to the Philippine group or supergroup of languages.

The languages of Minahasa, Sangir and Talaud, and Bolaang Mongondow have been classified as "Philippine languages of northern Sulawesi" since the late 19th century, when Adriani studied first Sangirese and afterwards Tontemboan. Esser followed Adriani's example in his language map of 1938, and the term is still used even in Stokhof's 1983 edition of the Holle wordlists. Adriani and Esser do not include in this group the 4 languages to the west of Bolaang Mongondow which were known to them. The latter were considered by them to constitute a separate "Gorontalo" subgroup,
coordinate with and transitional to the so-called Philippine subgroup (Adriani 1914:184). Although Adriani (p.183) regarded Mongondow as the most southerly member of the last-named group, he also discerned Philippine traits in the still more southerly Tomini language group (pp.175-178).

Brandstetter did not yet know these "transitional" languages when he was asked by the anthropologist F. Sarasin for his opinion about the Minahasan languages. He wrote (1906) in reply a three-page argument that there existed a significant distinction between them and the other languages of Sulawesi and a similarly significant agreement with the Philippine languages. His obviously incomplete evidence was also one-sided in that he gave examples from Mongondow (which is not a Minahasan language!) as evidence for his statement that the Minahasan languages agreed with those of the Philippines in having g as reflex of Proto Austronesian *R, rather than r, as in Makasarese and Buginese, or zero, as in Toraja (Bare'e). He did not add that the normal Minahasan reflex, even in reconstructed Proto Minahasan, is h (Sneddon 1978:69). But other problems have arisen since Brandstetter. In the first place doubts have been raised about the existence of Proto Philippine as a "meso-language" (Reid 1982; Sneddon 1984:11). Sneddon (1978:11,1984:4) treats both the Minahasan and the Sangiric group as reflexes of Proto Philippine, although they are considered by some to lie outside the Philippine group (Charles 1974), while even the Sangiric languages are believed by other researchers (Walton 1979; Llamzon & Martin 1976) to be at best only distantly related to the Philippine languages.

On the other hand, Mongondow has recently been proved to be more closely related to the Gorontalic languages to the west than to the Minahasan languages to the east (Noorduyn 1982:258; Usup 1986:3). Further, new evidence has recently been adduced showing that all languages west of Mongondow up to Buol constitute one group of Gorontalic languages, which is most closely related with Mongondow (Usup 1986:216-220). This means that if Mongondow is a "Philippine" language the Gorontalic group must be also. In that case there must be a rather sharp border to the west of Buol, between Buol and Toli-Toli. If the available evidence is examined, there is at least a problem of interpretation here. Barr, Barr and Salombe (1979:26,30) give a 61 per cent cognate relationship between Buol and Toli-Toli and are anxious to explain this comparatively high percentage by influence from long historical contact. But Usup (1986:68) gives an estimate of 41 per cent for the same wordlists as those given in Barr, Barr and Salombe (1979:81-83) and counting the cognates in these lists myself, I find 43-45 per cent. More evidence and further research are obviously needed in this matter.

It must nonetheless be said that Brandstetter's argument is basically valid. It is true that the subgroups for which a proto-language has been reconstructed each have a different reflex of Proto Austronesian *R, namely Proto Sangiric R (Sneddon 1984:39-41), Proto Minahasan h (Sneddon 1978:69), Proto Gorontalic-Mongondow g (Usup 1986:280) and Proto South Sulawesi R (Mills 1975:357-366), while the majority of Central Sulawesi (Toraja) languages have zero (through y). This means that, when the reflexes of *R are used as a criterion, these five subgroups of Sulawesi languages cannot be combined into one exclusive supergroup. This conclusion agrees with that reached by Sirk (1981:33) who, when discussing the case of the connection between the South Sulawesi group and the Toraja languages, states that "the supergroup hypothesis finds little if any support in available data". Rather, items in a South Sulawesi language which show a zero reflex of *R, such as Bugis uae and Sa'dan uai versus Makasar Konjo éré 'water', and Sa'dan muané versus Bugis woroané and Makasar buraqné 'man', must be borrowings from a Toraja language or, rather, "substrata originating from aboriginal languages of their present-day area".
However, one wonders how ancient the splits between these five Sulawesi subgroups are, when one finds Philippine lexical items not only in the northern Sulawesi subgroups but also in South Sulawesi. Some examples are *pia* ‘good’ (South Sulawesi, Sangiric, Gorontalic-Mongondow), *butaq* ‘land’ (Makasar *butta*, Gorontalic-Mongondow *bataq*) and even exclusively shared innovations such as *lipu* ‘country’ (South Sulawesi, Gorontalic-Mongondow only).

An example of a different dimension is the beneficiary/causative suffix *-aken* (e.g. in Javanese), which occurs in Central Sulawesi as *-aka* (Uma, Pamona, but also Wolio, Laiolo) and in eastern Sulawesi as *-ako* (Mori, Tolaki). It is ancient because of Proto Oceanic *-aki(n)*, and especially so in Sulawesi if Sirk’s claim is correct that it originated in Sulawesi. Sirk (1978:265) argues that it has merged with the suffix *-ken* (cf. Malay *-kan*, Sundanese *-keun*) everywhere else but in the central Sulawesi languages, where the latter suffix occurs as *-ka* next to *-aka* and with different semantic functions. But it is completely absent from South Sulawesi and cannot be reconstructed for Proto South-Sulawesi. Here its semantic functions are carried out by *-an* (e.g. in Sa’dan), which is *-ang* in Makasarese and Buginese. This change from *-an* to *-ang* must have occurred through generalised velarisation of final *-n* (universal in Makasarese and Buginese) and (because of Sa’dan *-an*) not through the influence of *-k* on the final *-n* (as must have been the case with the Balinese suffix *-ang*, the only example of velarised *-n* in this language).

On a purely lexicostatistical basis, however, there may emerge a closer connection between the languages of South and Central Sulawesi, as is shown by Dyen’s lexicostatistical classification. Dyen (1965:27) finds a Celebes Hesion, consisting of Pamona (Barê’e) and Lindu on the one hand, and Makasarese and Buginese on the other, based on 26.5 per cent cognacy, though he also remains hesitant about this result (p.47). As to the other Sulawesi languages included in Dyen’s classification, he tentatively concludes that either Tontemboan (one of the Minahasan languages) and Gorontalic (Gorontalo and Suwawa) are each coordinate with the group of all the languages of Sumatra, Java and Bali, or Tontemboan is coordinate with this group including Gorontalic, while he believes that Sangir may be closer to Gorontalic than to Tontemboan.

These almost inconclusive results tend to show that the lexicostatistical method can hardly be expected to bring the problem of the relationships between the languages of Sulawesi closer to a solution. It is, moreover, obviously impossible to apply this method to the limited vocabularies of reconstructed proto-languages of subgroups, since the lexical material of such proto-languages is insufficient to provide the basic wordlists needed for the purpose.

The second question that may be asked is whether there are characteristic features common to all Sulawesi languages or to at least a large number of them extending beyond one subgroup.

In answering this question, I may mention one widespread phenomenon as a possible candidate for a common trait of such a larger group: the parallelisms in the inflectional forms of the verb. To show how widespread these are, and also how different in detail, I give the following examples from languages of different groups:

<table>
<thead>
<tr>
<th>Verbal inflection</th>
<th>Possessive suffixes</th>
</tr>
</thead>
<tbody>
<tr>
<td>he sees me</td>
<td>my house</td>
</tr>
<tr>
<td><strong>South Sulawesi group</strong></td>
<td></td>
</tr>
<tr>
<td>Makasar</td>
<td>na-cini-ka'</td>
</tr>
<tr>
<td>Bugis</td>
<td>na-ita-wa'</td>
</tr>
<tr>
<td>Sa’dan</td>
<td>na-kita-na'</td>
</tr>
<tr>
<td>Subgroup</td>
<td>Verb Form</td>
</tr>
<tr>
<td>------------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Muna-Buton group</td>
<td>la-longa-aku</td>
</tr>
<tr>
<td>Barang-Barang</td>
<td>a-kamata-aku</td>
</tr>
<tr>
<td>Wolio</td>
<td>no-wora-kanau</td>
</tr>
<tr>
<td>East Sulawesi group</td>
<td>no-toa-aku</td>
</tr>
<tr>
<td>Tolaki</td>
<td>i-kita-aku</td>
</tr>
<tr>
<td>Kaili-Pamona group</td>
<td>na-hilo-a</td>
</tr>
<tr>
<td>Uma</td>
<td>na-kitayaku</td>
</tr>
</tbody>
</table>

It should be remarked in the first place that one of the reasons why examples are here given from only ten languages is the lack of information concerning this subject matter in other languages of the same subgroups. Whether all languages of these subgroups possess similar verbal inflection is simply not known and an answer to this question, however probable it may seem, cannot be given until grammatical descriptions of these languages have been produced and published.

Secondly, however, there are other Sulawesi languages, found in the northern half of the island, which are known to have no comparable verbal inflection. In Gorontalo, for instance, ‘he sees me’ is expressed as waatia he bilohu-lio (I being seen by-him), in which the pronominal suffix of the third person -lio serves as agent marker. The fundamental characteristic of this inflection is that this verbal suffix is the same as the possessive suffix which is used with nouns. Whereas there is a formal difference between these verbal and nominal (possessive) suffixes in the languages from which examples are listed above, in Gorontalo they are the same. ‘His house’ is bele-lio in Gorontalo, with the same suffix -lio. For this reason this kind of inflection has often been termed the possessive inflection.

It may perhaps be concluded that the occurrence of a separate verbal inflection is widespread among Sulawesi languages, but only in the southern half of the island, that is, in the four subgroups mentioned above (and perhaps one or two more), whereas the possessive inflection is found in the northern half of the island.

Finally it may be remarked that this situation does not imply that the possessive suffixes are never used with verbs in languages which have special verbal pronominal suffixes. In fact when they are used they have special semantic or syntactic marking. An interesting example may be given from Makasarese. In that language ‘I see him’ is ku-cini'-ki and ‘he goes away’ is a'lampa-i, but these two sentences can be conjoined by replacing the pronominal suffixes of both verbs by the corresponding possessive suffixes, producing the sentence ku-cini'-na a'lampa-na, ‘when I saw him he went away’, in which the use of possessive suffixes expresses a special semantic-syntactic relationship between the two verbs. This construction is usually called the nominal construction, but it could equally well be termed a possessive inflection.

**BIBLIOGRAPHY**

1914, *See Adriani and Kruyl, 1914.*
1928, *Bare'e-Nederlands woordenboek...* Leiden: Brill.


BERG, R. van den, 1988, *The Muna dialects and the Munic languages: towards a reconstruction*. MS.


