PACIFIC LINGUISTICS

Series B - No. 79

GLOTTOCHRONOLOGY WITHOUT COGNATE RECOGNITION

by

J.B.M. Guy

Department of Linguistics
Research School of Pacific Studies
THE AUSTRALIAN NATIONAL UNIVERSITY

Guy, J.B.M. *Glottochronology without cognate recognition.*
B-79, viii + 142 pages. Pacific Linguistics, The Australian National University, 1981. DOI:*10.15144/PL-B79.cover*

PACIFIC LINGUISTICS is issued through the *Linguistic Circle of Canberra* and consists of four series:

SERIES A - OCCASIONAL PAPERS

SERIES B - MONOGRAPHS

SERIES C - BOOKS

SERIES D - SPECIAL PUBLICATIONS

ALL CORRESPONDENCE concerning *PACIFIC LINGUISTICS*, including orders and subscriptions, should be addressed to:

The Secretary,
PACIFIC LINGUISTICS,
Department of Linguistics,
Research School of Pacific Studies,
The Australian National University,
Box 4, P.O.,
Canberra, A.C.T. 2600.
Australia.

First published 1980.

The editors are indebted to the Australian National University for help in the production of this series.

This publication was made possible by an initial grant from the Hunter Douglas Fund.

National Library of Australia Card Number and ISBN 0 85883 235 6

This monograph was produced on DEC-KL10 using VIDED, a display-oriented text editor written by Jacob Palme, of the Swedish National Defense Research Institute, and printed out on a Diablo-1640. Mathematical and phonetic symbols were typed in onto the Diablo printout. Unless otherwise acknowledged, all the software implementing and illustrating the methods described in this monograph was written by the author in Simula 67. Numbers in square brackets refer to tables or figures in appendix.

Editorial advice: Lois Carrington and Maureen P. Johnson.
Original style and layout: Maureen P. Johnson.

To G.Y.H.

TABLE OF CONTENTS

Guy, J.B.M. *Glottochronology without cognate recognition.*
B-79, viii + 142 pages. Pacific Linguistics, The Australian National University, 1981. DOI:10.15144/PL-B79.cover

CHAPTER 1: PRELIMINARY APPROACH

IDENTIFYING POSSIBLE SOUND CORRESPONDENCES

Consider the following wordlist from two little-known Austronesian languages:

English:	Titia:	Sese:
eye	mata	nas
sea	tasi	sah
father	tama	san
mother	mama	nan
tongue	mimi	nen
shellfish	sisi	heh
bad	sati	has
to stand	ti	se
to come	ma	na
with	mi	ne
not	sa	ha

The Titia and Sese words for "eye" ("mata" and "nas") might be related, and they might not. If they are, it would seem likely that Titia "a" would correspond to Sese "a", Titia "m" to Sese "n", and Titia "t" to Sese "s". But Titia "mata" and Sese "nas" could just as well be reflexes of, amongst many other possibilities, some protoform *amtas, e.g.:

	Titia:	Sese:
Stress placement	a'mtas	a'mtas
Loss of pretonic vowel	mtas	mtas
Cluster breakup	matas	
Loss of final consonants	mata	
Nasal assimilation		mnas
Cluster reduction		nas

If such a development is indeed true, Titia "a" would correspond to Sese "a", and Titia "t" to Sese "n", whereas no evidence of correspondences between Titia "m" and Sese, and Sese "s" and Titia, would have survived.

Given a pair of words suspected to be related such as Titia "mata" and Sese "nas", and in the absence of any other information, one can then only state that any one of the components of "mata" could possibly, even if very improbably, correspond to any one of the components of "nas".

Guy, J.B.M. *Glottochronology without cognate recognition.*
B-79, viii + 142 pages. Pacific Linguistics, The Australian National University, 1981. DOI:10.15144/PL-B79.1

The evidence for possible sound correspondences between Titia and Sese provided by the pair "mata", "nas", can be quantified in a 5x5 table (only five different phonemes happen to occur in each of the short lists given in this example):

Titia \ Sese	a	e	n	h	s
a	2	0	2	0	2
i	0	0	0	0	0
m	1	0	1	0	1
s	0	0	0	0	0
t	1	0	1	0	1

Each cell of the table shows the number of times a given possible sound correspondence was encountered when comparing the first word pair ("mata" and "nas"). Thus for instance, the number "2" at the intersection of row "a" and column "a" means that evidence of a possible correspondence between Titia "a" and Sese "a" was encountered twice (the first "a" of Titia "mata" against the "a" of Sese "nas" having provided the first case of such a possible correspondence and its second "a", again with the lone "a" of Sese "nas", the second case).

Likewise, the evidence for possible sound correspondences found when comparing the second word pair ("sea", Titia "tasi", Sese "sah"), is shown in the table:

Titia \ Sese	a	e	n	h	s
a	1	0	0	1	1
i	1	0	0	1	1
m	0	0	0	0	0
s	1	0	0	1	1
t	1	0	0	1	1

The total evidence provided by the first two items of the list is thus:

		Sese a	Sese e	Sese n	Sese h	Sese s
Titia	a	3	0	2	1	3
	i	1	0	0	1	1
	m	1	0	1	0	1
	s	1	0	0	1	1
	t	2	0	1	1	2

and that provided by all 11 items of the list:

		Sese a	Sese e	Sese n	Sese h	Sese s	Total:
Titia	a	10	0	9	3	6	28
	i	2	6	5	6	3	22
	m	5	3	12	0	2	22
	s	3	2	0	7	2	14
	t	4	1	2	2	5	14
Total:		24 +	12 +	28 +	18 +	18 =	100

FROM POSSIBLE TO PROBABLE SOUND CORRESPONDENCES

Consider the first row of this last table: Titia "a" is involved in 28% of all the possible sound correspondences with Sese. If Titia and Sese were unrelated, it should be reasonable to expect Titia "a" to show no special affinity with any particular Sese phoneme, so that the cell at the intersection of row "a" and column "a", which gives the number of cases of possible correspondences between Titia "a" and Sese "a", should account for just about 28% of column "a". In reality, it accounts for about 42% (10 out of 24) of all the observed cases of a possible correspondence between Sese "a" and Titia.

The same reasoning applied to the remaining 24 possible sound correspondences between Titia and Sese is summarized in the table below, where each cell contains two percentages: on the left side the proportion actually observed, and on the right side the proportion which would be expected if Titia and Sese were unrelated.

Titia \ Sese	a		e		n		h		s	
a	42%	28%	0%	28%	32%	28%	17%	28%	33%	28%
i	8%	22%	50%	22%	18%	22%	33%	22%	17%	22%
m	21%	22%	25%	22%	43%	22%	0%	22%	11%	22%
s	12%	14%	17%	14%	0%	14%	39%	14%	11%	14%
t	17%	14%	8%	14%	7%	14%	11%	14%	28%	14%

The cells where the discrepancies between observed and expected proportions are greatest are (not in order):

Titia:	Sese:	observed	expected
a	a	42%	28%
i	e	50%	22%
m	n	43%	22%
s	h	39%	14%
t	s	28%	14%

which happens to give precisely the substitution code which was used to generate the Sese words from the invented Titia glosses.

AUTOMATIC RECOGNITION OF SOUND CORRESPONDENCES: PRACTICE VS. THEORY

One could hardly expect such ideal situations ever to arise where, as in the fabricated Titia and Sese example, cognation is perfect and sound changes have somehow resulted in strict one-to-one correspondences. The method just outlined must therefore be tested on actual wordlists.

SELECTING REAL WORDLISTS

Two New Hebrides wordlists (Hiw and Toga), taken from Tryon 1976, were selected in a first test.

The manner in which these lists are given makes their inputting into a computer file an easy task, and I had long suffered from an extreme curiosity about what an accurate classification of the languages of the New Hebrides might reveal of their prehistory. But apart from this, there was no reason for this particular choice other than that Hiw and Toga happened to be the first two of the 178 lists available.

In fact, better lists could definitely have been selected. From fieldwork in 1974 I knew that Hiw had a richer than usual vowel system (probably nine vowels, possibly even more) which was not reflected these lists. Toga might have an equally complex phonology.

The lists are furthermore not phonemicized, and it might have been preferable to select lists elicited by myself (such as Vunapu, Tolomako, Roria, etc.) the originals of which were, unlike their published versions, phonemicized.

On the credit side of this choice, should the method prove itself on somewhat inaccurate wordlists, it could only perform better on fully reliable data; whereas a method requiring near perfect accuracy in the data to which it is to be applied would be of little practical value, as comparative linguists have all too often nothing but very deficient wordlists to work on.

ENCODING THE WORDLISTS

The Hiw and Toga lists were input into a computer file (see [1]) as they appear in Tryon 1976, with some spelling adaptations made necessary by the lack of phonetic symbols in the ASCII character set used by the DEC-KL10. Mostly, the same keys were used which on an I.B.M. Selectric typewriter equipped with a Camwil phonetic ball would have typed the required symbol, so that schwa for instance came out on the terminal screen as "7". No attempt was made at removing the prefixed articles or possessive suffixes present in the lists.

PROCESSING THE WORDLISTS

A computer program was written, which read the Hiw and Toga wordlists, produced a table of the possible sound correspondences encountered ([2]), calculated for each cell the probability that the discrepancy between observation and expectation was too large to be attributed to chance alone, and gave a list of the first 40 most likely sound correspondences in order of decreasing likelihood ([3]).

> The way in which the probabilities were calculated (using the normal distribution as an approximation to the binomial distribution) was chosen more for its simplicity and speed than for its accuracy, and would hardly meet with the approval of a statistician, as, for cells with low expectations or cells in near-empty rows or columns, the formula used grossly underestimates the probability of the difference between observations and expectations being due to chance alone. For

clear explanations and examples of the notions and mathematical formulae involved, see Spiegel 1972:122-124, 133-134.

RESULTS OBTAINED AND DISCUSSION

The first nineteen sound correspondences pinpointed by the method were right beyond any doubt. Especially encouraging was the fact that the correspondence between the Hiw velar fricative (denoted in the computer file by "G") and Toga "r" had been recognized, even though the Hiw velar fricative also corresponds to a velar fricative in Toga. The 18th and 19th correspondences in the list - e:3 and 7:) - are also right, the characters "3", "7", and ")" representing æ, ə, and œ respectively. Three other sound correspondences listed are worth commenting upon, although they appear at first as certainly wrong. They are:

	Hiw	Toga
(third column, second row)	n	æ
(third column, third row)	e	n
(fourth column, third last row)	n	i

One would object here that a vowel can hardly be expected to correspond to a consonant, and that these three correspondences are nonsense. But it must be remembered that the method followed compared, not symbols in corresponding positions, but each symbol of a word with all those of the other. Thus if in language X back vowels became fronted near dentals, but not in language Y, the method would be likely to find correspondences between the front vowels of X and the dentals of Y. Whether the three correspondences above do in fact reflect such a case in Hiw and Toga is besides the point here, made merely to suggest that the analysis of such tables of possible sound correspondences might provide clues which would normally be overlooked even by a trained comparative linguist.

PROSPECTS OF THE METHOD

Bearing in mind that these results were obtained on phonologically inaccurate wordlists, that the method used requires no phonetic information whatsoever (whether distinctive features, point or manner of articulation, etc.) about the symbols used in the transliteration of the wordlists, the automatic recognition of probable sound correspondences between related languages suddenly appeared a much less remote possibility than it could have been hoped.

FURTHER TESTS

The production of the table of possible sound correspondences and of the list of most likely sound correspondences between Hiw and Toga having taken only a few seconds of computer time, there was no valid reason not to submit it to further testing.

SECOND TEST: HIW VS. SAKAO

Hiw and Toga are fairly closely related (the computer-produced tables in Tryon 1976 give them as sharing 68.5% cognates), and the next logical step was therefore to select a more distantly related language. Sakao (NE Espiritu Santo), given as sharing 23.8% cognates with Hiw, and the diachronic phonology of which is known in detail (Guy 1978), seemed a likely choice.

The Sakao list was input in the same manner as Hiw and Toga, that is, without making any attempt at removing the affixes present in the lists.

The results obtained were incredibly bad ([4] and [5]). With hindsight, it seems that they could not have been otherwise. Whereas Hiw is not a particularly phonologically conservative language, Sakao is downright aberrant. Furthermore, Hiw and Toga are given as sharing 68.5% cognates out of 241 items compared, but Hiw and Sakao only 23.8% out of 239. The test to which the method had been put here was therefore no less than finding sound correspondences between two widely phonologically dissimilar languages, one of them inaccurately recorded, on the sole evidence of about 57 pairs of related words scattered amongst 182 unrelated pairs, without any clue as to which pairs were actually related, nor as to the system of transcription which had been used.

THIRD AND FOURTH TESTS: HIW, TOLOMAKO, AND SAKAO

Tolomako, a phonologically conservative language of NW Espiritu Santo, was selected for comparison with Hiw and Sakao in two final tests. The fear of a repetition of the failure of the Hiw-Sakao experiment made me input the Tolomako list phonemicized and stripped of affixes.

Tables [6] and [7], and [8] and [9] show the results obtained comparing Hiw with Tolomako, and Tolomako with Sakao. The sound correspondences between Tolomako and Sakao ([9]) are predicted with much better accuracy than between Hiw and Tolomako ([7]). This is not surprising: Tolomako is more closely related to Sakao, on the evidence not only of straight cognate counts (29.8% cognates with Hiw and 40.8% with Sakao), but also of more refined methods (Guy 1981).

FROM SOUND CORRESPONDENCE TO COGNATE RECOGNITION

METHOD

Taking for instance the Hiw and Sakao wordlists, one could, using some appropriate mathematical formula, calculate some probability of cognation for each word pair from their component phonemes and from the list of most likely sound correspondences already obtained. The items least likely to be cognate would then be temporarily removed from the lists, a new table of possible sound correspondences and a list of most likely sound correspondences would be computed from the shortened list; from these, probabilities of cognation would be calculated afresh for the original, complete wordlists; the least likely cognates pairs would again be temporarily discarded, and the process repeated until the very same items were found to be discarded on two successive iterations.

PRACTICABILITY: THE NEED FOR A DIFFERENT APPROACH

Even though such a method might eventually prove to identify cognates with tolerable accuracy, there are a number of reasons for believing that it would consume such large amounts of computer time as to be utterly unpracticable for classifying languages on the evidence of identified cognate pairs.

Traditional lexicostatistical methods suffer from an extreme sensitivity to unequal vocabulary retention rates, and even if vocabulary retention rates were constant, which they have long been shown not to be (Bergsland & Vogt 1962), the exceedingly small size of sample wordlists, usually representing at best 2% of the active vocabulary of the language sampled, would ensure that the phylogenetic classifications they yield would remain too unreliable to be of any practical use (Guy 1980).

Methods designed to cope with variable retention rates (Guy 1980:11ff, Guy 1981) are validly applied only to large numbers of languages and dialects.

The computation of a table of possible, and of a list of most likely, sound correspondences between two languages represented by a 250-item wordlist takes about two seconds of computer time on a DEC-KL10. This operation has to be carried out for each list pair, and would therefore require about 40 minutes of computer time for 50 lists, a sizable demand on computer resources, but by no means unthinkable. However, for each of the 1125 list pairs to be considered in attempting to classify just 50 languages, 250 word pairs would then have to be compared for possible cognation, with allowances made for sound loss, metathesis, the possible presence of irrelevant affixes, etc. This constitutes a much more complex task than the construction of a table of possible sound correspondences and the production of a list of most likely correspondences, and it is difficult to see it taking less than an hour of computer time (from an early attempt of mine at devising such an algorithm it seems that four hours would be a more accurate lower estimate).

Furthermore, there is no telling how many iterations would be needed for any two wordlists to see their putative inventory of shared cognates remain unchanged from the previous iteration. Even if a stable result were reached only after two iterations, a most unreasonable expectation, the processing of 50 lists of 250 items could by no stretch of the imagination take less than about three and a half hours of computer time, and would be more likely to take at the very least 10 hours. Two hundred lists would take 16 times as long, leading into quite unacceptable figures.

CHAPTER 2: LANGUAGE CLASSIFICATION WITHOUT COGNATE RECOGNITION

GENERAL PRINCIPLES

Consider the procedure followed in identifying most likely sound correspondences. Those cells of the table of possible sound correspondences which showed a suspiciously higher proportion of observations than should have been expected if the two languages under scrutiny had been unrelated, were often found to betray actual sound correspondences between these languages. Therefore, the more such cells, and the larger the discrepancies between the observed and expected values in each cell, the more closely related the languages should prove to be. Perhaps then, some measure of language similarity could be directly extracted from tables of possible sound correspondences and used, instead of proportions of shared cognates, to produce phylogenetic classifications.

SEARCHING FOR A DIRECT MEASURE OF LANGUAGE RELATEDNESS

THE χ^2 STATISTIC

An obvious answer is provided by statistical theory:

> A measure of the discrepancy existing between observed and expected frequencies is supplied by the statistic χ^2 (read chi-square) given by
>
> $$\chi^2 = \frac{(o_1-e_1)^2}{e_1} + \frac{(o_2-e_2)^2}{e_2} + \ldots + \frac{(o_k-e_k)^2}{e_k} = \sum_{j=1}^{k} \frac{(o_j-e_j)^2}{e_j} \qquad (1)$$
>
> (Spiegel 1972:201)

where o_1 stands for the number of cases observed and e_1 that expected in the first cell, o_2 and e_2 in the second cell, etc.

From the value of χ^2 thus obtained, one can then calculate the probability that the discrepancies found in the table are due to chance alone, i.e. that the two languages represented are not really related, but only appear to be so through sheer coincidence.

A comparative linguist, however, has no interest in learning the chances of two languages being related or not: he normally already suspects them to be related, and in comparing them for cognates is primarily concerned in measuring their degree of relatedness.

But, perhaps again, there exists some relationship between χ^2 computed from the sound-correspondence table of two languages and the proportion of cognates they share.

χ^2 AS FUNCTION OF THE PROPORTION OF SHARED COGNATES

THEORETICAL EVIDENCE

Observed Frequencies and Proportions of Cognates

Imagine that a 200-item wordlist from Titia and Sese happened to show exactly 50% cognates. We split this 200-item list into two 100-item lists such that the first list contains all those items which are cognate. A table of possible sound correspondences is then produced for each partial 100-item sublist and for the complete list. Take a cell of the whole-list, 50% cognate table, and call its observed frequency $o_{0.5}$ and its expected frequency $e_{0.5}$; call o_1 and e_1 the observed and expected frequencies of the corresponding cell of the table produced from the all-cognate sublist and o_0 and e_0 those of the corresponding cell in the table produced from the cognateless sublist.

The very method of producing a table of possible sound correspondences ensures that:

$$o_{0.5} = o_1' + o_0'' \tag{2}$$

Imagine now that Titia and Sese had shared 75% cognates. It would have been possible to subdivide the 200-item list into four sublists one of which would have contained only non-cognates, the other three containing all the cognate pairs. And again, the observed frequency $o_{0.75}$ of any given cell of the full table would have been equal to the sum of the frequencies of the corresponding cells of the tables produced from the four sublists:

$$o_{0.75} = o_1' + o_1'' + o_1''' + o_0'''' \tag{3}$$

Even if they had shared 100% cognates, the full list could have been divided into four sublists so that this time:

$$o_1 = o_1' + o_1'' + o_1''' + o_1'''' \tag{4}$$

Provided that the freqencies observed are large, i.e. that the wordlist used is long enough, and that the items constituting each sublist have not been perversely selected, the observed frequencies of the cells of the all-cognate subtables can be expected to be in proportion with each other. Thus, in formulae (3) and (4) above, o', o'', o''', and o'''', having been computed from 50-item sublists, can be expected to be about equal. From which it follows that, approximately:

$$o_C = Co_1 + (1-C)\, o_0 \tag{5}$$

where o_C denotes the observed frequency of a cell of a table of possible sound correspondences between two lists which share a proportion of cognates C, o_1 the frequency which would have been observed in that cell if these lists had shared 100% cognates, and o_0 if they had shared none.

> Assigning items to a sublist on the basis of the length of the native glosses, or of the presence in them of unusually many (or few) sounds with a given point of articulation, would constitute instances of perverse selection of sublists. However, unless the phonotactics and semantics of the languages compared are so interwoven that such perverse distribution patterns cannot be avoided (and I doubt that such cases are possible in theory, let alone in practice), it is always possible to split the original list in a non-perverse manner, so that formula (5) holds true, the approximation becoming better as the size of the sample list increases.

Expected Frequencies and Proportions of Cognates

The expected frequency of a given cell is obtained by multiplying the sum of the observations in the row of the cell by the sum of the observations in the column of the cell and dividing this product by the grand sum of the observations in the table.

Assume that lists elicited from any two given languages X and Y contain no loanwords, so that a given item not cognate in X and Y is necessarily an innovation, perhaps of X, perhaps of Y, perhaps of both. Then, from what is known of human languages, the same phonotactics and the same relative phoneme frequencies can be expected to apply in any set of words from, say, X, whether these words have cognates in Y or not. Therefore, if the whole wordlist is divided into two sublists, one containing only cognate pairs, the other non-cognates pairs, word length, phoneme distribution, and relative phoneme frequencies obtained from each sublist can be expected to be nearly identical. Thus, in a table of possible sound correspondences computed from 200 unrelated word pairs selected from the vocabulary of two related languages, the sum of the rows, the sums of the columns, and the grand sum of the frequencies in the table can be expected to be approximately equal to those found in a table computed from 200 cognate pairs from the same languages. From which it follows that expected frequencies are independent of the amount of cognates.

Effects of Loanwords

If the lists contain loanwords, it is quite possible that these loanwords will somehow reflect the phonotactics and relative phoneme frequencies of the language of origin, and expected frequencies will not be strictly independent from cognation (or apparent cognation). However, if loanwords constitute a minority of non-cognate forms, then their numbers will distort the native phoneme frequencies only to a minor extent, and the formula for the computation of expected frequencies will still hold true, even if only approximately so.

Conclusion

The χ^2 statistic for a table of possible sound correspondences between two wordlists sharing a proportion of cognates C is therefore approximated by:

$$\sum \frac{(Co_1 + (1-C)\, o_0 - e)^2}{e} \tag{6}$$

Expression $(Co_1+(1-C)o_0-e)^2$, first expanded into a polynomial and then factorized, becomes:

$$C^2(o_1-o_0)^2 - 2C(o_0-e)(o_1-o_0) + (o_0-e)^2 \tag{7}$$

Expression (6) is therefore equivalent to:

$$\sum \frac{C^2(o_1-o_0)^2 - 2C(o_0-e)(o_1-o_0) + (o_0-e)^2}{e} \tag{8}$$

and hence to:

$$\sum \frac{C^2(o_1-o_0)^2}{e} - \sum \frac{2C(o_0-e)(o_1-o_0)}{e} + \sum \frac{(o_0-e)^2}{e} \tag{9}$$

Remember that o_0 denotes the observed frequency of a cell when the two lists compared share no cognates at all, in which case phoneme correspondences can be expected to be random, and therefore the numbers observed in the cells tend to be very close and even sometimes equal to those expected.

Consider then the second term of expression (9):

$$\sum \frac{2C(o_0-e)(o_1-o_0)}{e} \tag{10}$$

The difference between observed and expected frequencies, o_0-e, will be sometimes negative, sometimes positive, and sometimes zero. Expression (10) is then a sum of terms sometimes negative, sometimes positive, mostly quite small: its actual value can be negative or positive, but, phoneme correspondences being random, for each negative term we can expect to find a positive term of approximately equal absolute value, so that the expected value of (10) is zero.

Note that to say that the expected value of expression (10) is zero does not mean that it will be zero in most cases encountered in practice (on the contrary, it is extremely unlikely to be zero!). Imagine yourself betting a dollar on tails for 1000 tosses of a coin at even odds. This is a fair game, out of which you can expect to make no profit, and no loss: your expected pay-off is zero. And yet, it would be very extraordinary indeed if, on the thousandth toss, you were to break exactly even.

The third term of expression (9), however, being a sum of squares, is a sum of terms all positive, and its expected value is therefore also positive.

The expected value of the χ^2 statistic of a table of possible sound correspondences between any two wordlists with a proportion of shared cognates C is therefore (after cancelling the second term of expression (9), the expected value of which is zero):

$$\sum \frac{C^2(o_1-o_0)^2}{e} + \sum \frac{(o_0-e)^2}{e} \tag{11}$$

which is equivalent to:

$$C^2 \sum \frac{(o_1-o_0)^2}{e} + \sum \frac{(o_0-e)^2}{e} \tag{12}$$

The expected value of the χ^2 statistic of a table of possible sound correspondences is thus proportional to the square of the proportion of shared cognates, plus a quantity (the second term of the above expression) independent from the proportion of shared cognates. This quantity,

$$\sum \frac{(o_0-e)^2}{e} \tag{13}$$

is the value that the χ^2 statistic could have been expected to take if the discrepancies between observed and expected frequencies in the table of possible sound correspondences had been due to chance alone, i.e. if the two wordlists had shared no common vocabulary. When the number of degrees of freedom (here the product (m-1)(n-1) where m and n denote the number of rows and columns in the table) is large (greater than 30), this value is approximately equal to the number of degrees of freedom minus one half. In linguistic terms, the χ^2 statistic of a table of possible sound correspondences between two cognateless wordlists is thus, surprisingly enough, a function not of the size of the wordlist, but merely of the phonemic - or alphabetic - inventory of the languages represented.

The factor in the first term of expression (12),

$$\sum \frac{(o_1-o_0)^2}{e} \tag{14}$$

is the expected value of χ^2, had the two lists compared happened to contain only cognate pairs. I am unfortunately not mathematically competent enough to comment upon the expected value of this quantity beyond saying that it is directly proportional to the grand sum of the cell frequencies in the table, and that it appears to be some as yet unknown function of phonological decay.

PRACTICAL EVIDENCE (MONTE CARLO METHOD)

Coming from a fellow linguist rather than from a statistician, the lengthy demonstration above is not likely to sound much convincing to readers. And rightly so, as I was not myself satisfied until I obtained some more tangible evidence of the relationship between the χ^2 statistic and the proportion of shared cognates.

Method

To obtain such evidence, a computer program was used which, given a wordlist of some language X, would produce wordlists in a pseudo-language identical to X in its phonemic inventory, phoneme frequencies, and phonotactics, but which would share varying amounts of cognates with the original wordlist. Thus, original language and pseudo-language being phonologically identical, a gloss of the pseudo-wordlist would be considered cognate with the corresponding gloss of the original list if both were identical. Then, each pseudo-list was matched against the original list, a table of possible sound correspondences computed, and its χ^2 value stored into another computer file along with other relevant statistics (proportion of shared cognates, total observations in the table of possible sound correspondences, degrees of freedom).

Problems Encountered: Minimum Expected Frequencies

Some non-trivial problems arose when writing the program which was to calculate χ^2 from tables of possible sound correspondences:

> If the expected frequency of a cell does not exceed 5, this cell should be combined with one or more other cells until the above condition is satisfied.
>
> [...]
>
> From a theoretical point of view it is legitimate to combine cells in any desired manner, provided that one is not influenced by the observed frequencies. In many applications, however, there are practical reasons for combining neighbouring cells [...].
>
> (Hoel 1971:229-230)

When χ^2 is computed on a rectangular table, it is not the offending cells which have to be combined with others, but the whole rows or columns containing them. The last remark in the quotation from Hoel must therefore be reinterpreted: "In many applications, there are practical reasons for combining neighbouring rows (columns)."

The reason why neighbouring, rather than randomly chosen, rows or columns should be combined is that, in most applications, a table will have been drawn up in some orderly manner so that neighbouring rows and columns will tend to contain the frequencies of not too dissimilar phenomena, thus resulting in like being combined with like.

In the case of a table of possible sound correspondences, however, no such ordering seems possible, for it can be argued ab absurdo that if a satisfactory ordering could be found, since it would necessarily be linear, whereas speech sounds are ordered in at least two dimensions (point and manner of articulation), some like sounds would still inevitably be kept several rows or columns apart, and this ordering would thereby be unsatisfactory. If for instance successive rows or columns were assigned the symbols y, i, e, a, o, u, w, one could still argue that fronting of "u" to "i" being such a common phenomenon, they should be assigned to neighbouring rows and columns. And further, what symbols should be placed to the left of "y" and the right of "w"? Palatals and dentals seem to present more affinity to "y", labials to "w", thus perhaps:

n d t s z y i e a o u w v f p b m

A seemingly satisfactory arrangement, until one realizes that nasals are poles apart, or tries to fit in velars.

Reflecting on the reasons why, in practical applications, like observations should be combined with like observations made the random selection of rows or columns an unappealing alternative. And so eventually, rather than combining with others the rows and columns containing the offending cells, the program written simply deleted them.

Secondly, the reason why no cell should have an expected frequency of less than 5, is that the χ^2 distribution used for estimating whether observed discrepancies are due to chance alone is only an approximation of the exact distribution of the quantity

$$\sum \frac{(o - e)^2}{e} \qquad (15)$$

and that experience and theoretical investigations have shown the approximation to be usually satisfactory when the above condition (no expected frequencies of less than 5) was met. However, in this particular application, what is sought is not a probability of cognation of two languages, but a measure of their degree of relatedness, for which purpose the knowledge of the distribution of the χ^2 statistic might eventually prove unnecessary. Therefore the program was made to calculate χ^2 values before and after deleting the rows and columns containing cells with frequencies of less than 5.

Finally, the deletion of some rows and columns from a table of possible sound correspondences reduces the total number of observations in the table, so that cells with an earlier expected frequency of 5 or over before deletion, may show after deletion an expected frequency of less than 5, thus calling for the deletion of new columns and rows. This iterative process was judged too expensive in computer time to be used in the program and, instead, a shortcut was used which proved successful in eliminating all such near-empty cells (in calculations involving 820 tables, not a single cell with a frequency of less than 5 failed to be deleted).

Results of the Simulations

Two set of data ([10] and [11]) were produced using the method discussed above, each consisting of the results obtained matching 200 computer-generated wordlists to the original model. The Tolomako wordlist was used as model for the first set, the Sakao list for the second set.

From each of these two sets of data two graphs were plotted, one showing the relationship between proportions of shared cognates and χ^2 values calculated before near-empty rows and columns were removed from the tables of possible sound correspondences, the other one showing the same relationship after the removal of near-empty rows and columns ([12] to [15]).

The pattern exhibited by all graphs shows a strong correlation between values of χ^2 and proportions of shared cognates. The line along which the paired observations cluster could betray a logarithmic function ($y = a \operatorname{Log} x + b$), an exponential function ($y = a b^x$), or a quadratic function such as $y = ax^2+bx+c$ or $y = ax^2+b$. But most importantly, the curves all meet the axis of the Y-coordinates at points closely corresponding to the number of degrees of freedom of the tables of possible sound correspondences, i.e. when proportions of cognates are very low or nil, the χ^2 statistic becomes close to the number of degrees of freedom of the tables. This is precisely what the theory had predicted.

EIGHT MEASUREMENTS OF LANGUAGE RELATEDNESS

Having shown beyond reasonable doubt that the χ^2 statistic of a table of possible sound correspondences between two languages is an approximate function of aNC^2-b, where N is the grand sum of the frequencies in the table, C the proportion of common vocabulary, b the number of degrees of freedom minus one half, and a an unknown quantity probably depending on the conjugated phonotactics and degree of phonological decay of the languages compared, it seems that $(\chi^2-b)/N$, being an approximation of aC^2, might provide some useful measurement of language similarity.

The fact that b is usually quite small compared to aNC^2 suggests that the simpler expression χ^2/N might be so close to $(\chi^2-b)/N$ as to provide an equally useful measurement.

Furthermore, as pointed out above, the purpose to which the χ^2 statistic is being put to use here might not necessitate the weeding out of those rows and columns containing cells with expected frequencies of less than 5. Two sets of measurements should therefore be obtained in a preliminary implementation of the proposed method: one computed before "weeding", and one after.

Finally, since the lengthiest part of the necessary computations is the construction of the tables of possible sound correspondences, the "weeding" procedure and the calculation of χ^2 itself, from which these measurements are then derived in a negligible amount of computer time, as a control, two additional measurements were used, one being χ^2 unmodified, the other χ^2-b.

SEEKING METHODS COMPATIBLE WITH THE NATURE OF THE MEASUREMENTS

INADEQUACY OF CLASSIC LEXICOSTATISTICAL METHODS

The measurements defined above, the least heterogeneous of which is an approximation of the product of the square of the proportion of common vocabulary by some unknown function of phonological decay, very obviously cannot be submitted to classic lexicostatistical techniques to yield meaningful results, as these techniques are based on the calculation of estimated time depths, which one could hardly hope to obtain from any of the measurements contemplated here.

MORE RECENT METHODS

The measure of phylogenetic affinity between languages elaborated and tested in Guy 1980 and Guy 1981 (henceforth to be referred to as linear-correlation measure), is based on the observation that, when any two languages A and B share an IMMEDIATE common ancestor, the proportion of cognates shared by A with any other language X and that shared by B with that same language X tend to be in a constant ratio, whatever the choice of X. The reason for this phenomenon being that, if, since their split from their immediate ancestor, A has retained a proportion rA of its inherited vocabulary and B a proportion rB, then the expected proportion of common vocabulary between A and X is rA/rB times that between B and X, so that there exists between paired proportions of cognates between A and B and any other related language a linear relationship of the type

$$y = ax$$

where y denotes the proportion of cognates between one member of the pair and the outsider, x the proportion of cognates between the other member of the pair and that same outsider, and a the constant ratio, independent from the choice of the outsider, but dependent upon the past rates of lexical retention of the two members of the pair since their split.

APPLICABILITY TO χ^2-BASED MEASUREMENTS

ROLE OF LEXICAL DECAY IN THE MEASUREMENTS

If linear-correlation measures are computed, not from proportions of shared cognates, but from their squares, they lose none of their validity, since the square of the proportion of vocabulary shared between A and any outsider X and the square of that shared between B and X still tends to be in a constant ratio, $(rA/rB)^2$ this time, just as independent of the choice of X as rA/rB is.

ROLE OF PHONOLOGICAL DECAY

Implications of a Multiplicative Property of Phonological Decay

The χ^2-based measurements elaborated so far all involve the square of the proportion of shared cognates. If they involved no other factors, they could therefore be submitted to the same methods that were developed for use on percentages of shared cognates.

Unfortunately, although two extraneous factors (expected χ^2 value for unrelated lists and total cell frequencies) can easily be abstracted, there seems to be no way, at least at present, of dissociating cognate proportions from that unknown element which is most probably an effect of phonological decay.

If this unknown element should happen to be mainly a measure of phonological conservatism with a multiplicative effect, i.e. such that phonological conservatism between two languages A and B were equal to the product of the phonological conservatism between each and their CLOSEST common ancestor, in the very same manner in which the expected proportion of common vocabulary between two languages is equal to the product of their proportion of common vocabulary with their CLOSEST common ancestor, then, for any two languages A and B with an IMMEDIATE common ancestor, there would also exist a constant ratio, based on shared phonology, and parallelling that based on shared vocabulary. If so, the measurement $(\chi^2-b)/N$ proposed above being the product of the degree of phonological conservatism by the square of the proportion of common vocabulary, would enjoy the same multiplicative property as those based on proportions of common vocabulary, and could just as validly be translated into linear-correlation measures for which methods of phylogenetic classification have been discovered.

Implications of Non-Multiplicative Properties

Phonological decay might follow the rules of lexical decay; but it seems more likely that it does not, and that, if we denote by p(A,B) some measure of the amount phonological structure retained between two successive states A and B of a given language, then p(A,C) is some very complex function of p(A,B) and p(B,C) rather than a mere product - for which case no theoretically exact method if available at present.

Approximating Non-Multiplicative Properties

Polynomial theory brings a glimmer of hope: under certain conditions (the existence of derivatives), which would appear most likely to be satisfied under any reasonable theory of phonological decay, any function can be approximated by a polynomial (Taylor polynomial) consisting of the first terms of an infinite series (Taylor series of the function). The more terms used, the closer the approximation. One could then hope that, whatever the unknown rules governing phonological decay may be, the first term of the Taylor series of the corresponding, unknown function might provide a good enough approximation of the true function for practical purposes. If so, phonological decay would function roughly as if it had the same multiplicative property as lexical decay.

In fact, since the lexemes of a given language do not stand equal chances of replacement over a unit period of time (Dyen, James, & Cole 1967), even lexical decay does not strictly obey the product rule on which glottochronology and lexicostatistics are based (Guy 1981:7-11).

CONCLUSION

Phonological decay can possibly be considered, for practical purposes only, to obey the same rule as ideal lexical decay, although it is almost absolutely certain that it does not. At best, the linear pattern evidencing closely related languages in graphs constructed according to the method explained in Guy 1980:1-2 might become less distinct, in which case the methods developed for the linear-correlation measures should still yield valid results; at worst, it will disappear almost entirely, in which case these same methods will prove worthless.

The answer to the question of whether valid classifications of related languages and dialects can be obtained directly from the χ^2-derived measurements of language similarity elaborated earlier in this chapter now rests on the examination of their performance on real data.

CHAPTER 3: SOME TEST APPLICATIONS

CHOICE OF THE TESTS

GENERAL CONSIDERATIONS

The nature of the measurements of genetic language similarity elaborated in the previous chapter necessitates the use of sizable wordlists; and the algorithmic methods appropriate for processing these measurements can be applied only to large numbers of wordlists.

It would have been desirable to test the methods proposed in this monograph on a family of languages, such as Indo-European, the history of which is known in its major part beyond reasonable doubt. However, wordlists satisfying the requirements of the method are not easily available. Alternatively, the method could be tested on data from a simulated language family, as in Guy 1980. The simulation of phonological decay is unfortunately a much more demanding problem than that of lexical decay.

For these reasons, a group of languages had to be used for which a sufficient number of long enough wordlists had been published in a form easily translatable into a computer file, and for which some evidence of past evolution, other than purely lexicostatistical, would provide some yardstick of the success of the methods, or of their failure.

This last condition can be considerably relaxed on the principle that, if the proportions of shared cognates are also available along with the wordlists themselves, once translated into linear-correlation measures, they can be used to yield a phylogenetic classification. Then, the validity of the classifications obtained from χ^2-derived measurements of language similarity can then be judged by comparing them to the cognate-based classification.

FIRST TEST: THE LANGUAGES AND DIALECTS OF ESPIRITU SANTO

REASONS FOR THE CHOICE

Although nothing is known of the prehistory of Espiritu Santo, a general classification of the communalects of the New Hebrides has proved to agree to a surprising extent with the geography of the area and with local toponymy and oral tradition (Guy 1981). According to this classification, the communalects of Espiritu Santo form a closeknit subgroup within the New Hebrides.

In the absence of previous experience with the method elaborated in the first two chapters, it was of course impossible to say whether 41 lists would be enough to yield a valid classification. But a classification of all 178 lists available would have taken around six hours of computer time, not to mention the time spent inputting the data into a file, whereas the processing of 41 lists could be expected to take only about half an hour.

THE WORDLISTS

As mentioned earlier, the wordlists published in Tryon 1976 are not phonemicized and some (Hiw) are somewhat phonetically inaccurate. The same observation is true of most of the Espiritu Santo wordlists. The difficulty is compounded by the fact that these wordlists were collected by different persons, some of them phonetically untrained, and reflect heterogeneous hearing idiosyncrasies. For instance, the wordlists for Amblong, Butmas, and Tur were collected by A.L. Jackson, an anthropologist, and seem to reflect English spelling habits. Thus we find in Amblong abatir (from *gapatia) for "moon", and in Tur sumsum for "food", where the two closely related languages of Lorediakarkar and Shark Bay have samsam. These apparently unexplainable sound correspondences might, on the other hand, simply be typing mistakes, as we also find in Sakao œaœð ɣœr instead of œnœð ɣœr for "dry coconut", pič∅ instead of pɨtj∅ for "to throw", in Tolomako βole instead of ɣole for "spear", in Nokuku ikʔa, a phonologically impossible form, for "fish", etc. However, judging from the procedure followed in constructing tables of possible sound correspondences, such mistakes should result only in minute errors in the values of χ^2 obtained.

More disturbing is the presence of irrelevant affixes in a number of glosses, presence which remains unmarked in any way. Thus for instance, in most lists, transitive verbs show an a or i final which is none other than the 3rd person singular object suffix. The fact that this ending does not belong to the gloss proper is sometimes, but not systematically, shown by the use of a hyphen. Thus we find Tolomako ɣati-a "to bite" and roɣo-i "to hear" but also ɣitsia "to squeeze" and tuɣia "to stab". In some cases subject markers have even been included, e.g. Amblong karonea (i.e. "we hear it") for "to hear".

Since it was impossible to restore these wordlists without spending a prohibitive amount of time on the task, it was decided to input them as they were, affixes included, for it was reasoned that, if the methods to be tested proved successful when applied to tainted data, they should perform all the better when applied to accurately recorded wordlists.

TRANSLITERATION SYSTEM

Since some of the wordlists to be processed were phonetically inaccurate, there seemed to be no compelling reason for insisting on a one-to-one transliteration such as the one used for the Hiw, Toga, Tolomako, and Sakao lists for testing the method of automatic recognition of sound correspondences and illustrating the relationship between χ^2 and proportions of shared cognates.

Digraphs and postposed diacritic symbols were therefore used liberally, which greatly increased the readability of the data file. Hyphens and spaces were deleted from the glosses and missing items were represented by an asterisk (*). A sample of the data thus encoded is given in the appendix ([16]).

CLASSIFICATIONS

Eight tables of χ^2-derived measurements were produced, one for each of the measurements explained in chapter 2. These were then transformed into linear-correlation measures, from which eight phylogenetic trees ([17] to [24]) were obtained using the n-way splitting algorithm described in Guy 1980:16-17.

As a control, the proportions of cognates shared between these 41 lists, based on Tryon's cognate recognition, were translated into linear-correlation measures to give, by the same method, a phylogenetic tree ([25]).

In comparing these trees, it must be remembered that linear-correlation measures only aim at estimating the order in which successive splits occurred in a language family, that the lengths of the branches of these graphs are therefore not proportional to the time elapsed between successive splits, and that the nature of the n-way splitting algorithm is such that the earliest splits of the family are reconstituted most accurately.

BEST EXPECTED RESULT

Before proceeding to assess the results obtained, it might be worth to try and predict which one of the eight measurements used is likely to yield the classification closest to the tree reconstructed from proportions of shared cognates ([25]).

Measurement $(\chi^2-b)/N$ seems the most likely candidate, since it is our closest estimate of the quantity aC^2, where C denotes the proportion of shared cognates and a that unknown quantity which is most probably mainly a measure of retained phonological structure.

Each of the four basic measurements, however, was taken before and after the removal of near-empty rows and columns from the tables of possible sound correspondences. Now the quantity b in expression $(\chi^2-b)/N$ denotes the expected value of χ^2 for lists which share no cognate forms. Statistical theory gives this expected value, in those cases with about 30 degrees of freedom or more, as approximately equal to the number of degrees of freedom minus one half. But, the condition regarding a minimum number of degrees of freedom satisfied, that regarding the minimum expected frequencies of invidivual cells still remains, since it governs how closely the χ^2 distribution approximates the true distribution of the quantity:

$$\sum \frac{(o - e)^2}{e}$$

Thus, on theoretical grounds alone, it would seem that it is the $(\chi^2-b)/N$ measurement, taken after the deletion of near-empty rows and columns from the tables of possible sound correspondences, which should give the most reliable results.

BEST OBSERVED RESULT

The classification based on the $(\chi^2-b)/N$ measurement is the only one which agrees with the cognate-based classification in reconstructing an earliest split of the Santo languages into two great subgroups, the larger one of which later split into three further subgroups ([17] and [25]). On this count alone, the $(\chi^2-b)/N$ measurement proves the most reliable, as predicted by the theory. Therefore, only the classification based on this measurement is considered in the detailed analysis of the results which follows.

ANALYSIS

NE Santo Subgroup

The seven NE Santo lists (Tur, Butmas, Polonambauk, Shark Bay 1, Shark Bay 2, Lorediakarkar, and Sakao) form a distinct subgroup as against the rest of Santo in both reconstituted genealogical trees ([17] and [25]).

Main Santo Subgroup

In both reconstitutions again, the larger subgroup, call it Santo Main, subdivides into three. Here, however, reconstructions differ in the lists assigned to each of these three subgroups. The cognate-based classification assigns Tolomako to the first subgroup (NW Santo), and Roria to the second (West Santo), whereas the classification based on $(\chi^2-b)/N$ assigns Tolomako to the second subgroup and Roria to the third (South Santo and offshore islands).

Position of Tolomako

Tolomako has dentals as reflexes of *p and *m preceding front vowels (Guy 1978). This particular sound change is found only in two South Santo languages (Roria and Tambotalo), in all the NE Santo languages, and in an intermediate stage of development (labials going to apico-labials before front vowels) in most offshore island languages (Mafea, Aore, Tangoa, and, sporadically, Araki). Tolomako has furthermore a fricative, ɣ, as a reflex of *ŋ in all environments, which particularity is shared only by Sakao, although only in well-defined post-tonic environments, and probably by earlier stages of Shark Bay and Lorediakarkar, where *ŋ is lost in certain post-tonic environments (Guy 1978). On the evidence of diachronic phonology, it seems that Tolomako should be classified with the NE Santo languages, or with Roria, Tambotalo, and the offshore islands. On purely impressionistic grounds, based on a working knowledge of Tolomako and Sakao, and on having elicited over half the Santo wordlists published in Tryon 1976, I would give the classification of Tolomako with the NW Santo languages as almost certainly wrong. But beyond this, there is no telling, in the present state of knowledge, where it belongs within Santo.

Position of Roria

Roria, as already mentioned above, has dentals as reflexes of *p, *m, and *v preceding front vowels (e.g. nata "eye" from *mata, ɣirið "rat" from *garivi, ⁿdeⁿd "butterfly" from *pepe), and it seems therefore more reasonable to group it, as does the $(\chi^2\text{-b})/N$ measurement, with the subgroup containing Tambotalo and the offshore islands rather than with the West Santo subgroup, as does the cognate-based classification.

Position of Tambotalo

Tambotalo is classified with Tangoa on the basis of proportions of shared cognates, whereas, on the basis of $(\chi^2\text{-b})/N$ measurements, it is given as having split early from an offshore island subgroup Mafea-Tutuba-Aore-Malo. This latter classification seems all the more likely that it also groups Tangoa with a nearby island, Araki, rather than with Tambotalo.

Subdivision of the Shark Bay Area

The classification of Lorediakarkar and the two Shark Bay communalects obtained using the $(\chi^2\text{-b})/N$ measurements is almost certainly wrong: the two Shark Bay communalects are very probably more closely related to each other than to Lorediakarkar.

Wusi and Kerepua

The two lists from Wusi, probably reflecting two very close dialects, are certainly more closely related to each other than to Kerepua. Here, it is the cognate-based classification which proves wrong.

CONCLUSION

The discrepancies between the genealogical reconstructions obtained on the evidence of identified cognates and that obtained by by-passing the process of cognate identification are quite slight, and, where discrepancies arise, the latter method seems to provide the more reliable classification.

SECOND TEST: NORTHEAST NEW HEBRIDES

REASONS FOR A SECOND TEST

However remote it may be, there is always a possibility that positive results such as those just obtained for Espiritu Santo could be a mere coincidence (such a coincidence was encountered once: see Guy 1980:28).

CHOICE OF THE DATA

The methods elaborated in this monograph having been successful in the case of Espiritu Santo, the next logical step was to apply them to a slightly more difficult case. The fact that the phonologies of a sizable proportion of the 34 communalects of the Northeast New Hebrides used in this second test are not properly reflected in the wordlists published in Tryon 1976 (Lehali, Lehalurup, and Hiw certainly, Toga, Apma, Seke, and Sa probably), and that affixes are very often incorporated in the glosses, sometimes with and sometimes without hyphenation (particularly the first half of the wordlists, being nouns about one third of which are obligatorily possessed, and items 154 through to 186, which carry mostly adjectival meanings), should make the task appreciably more difficult than in the case of Espiritu Santo.

The lists were input under the same principles as the Santo lists, that is, affixes included (see sample [26] in appendix).

CLASSIFICATIONS AND ANALYSIS

Nine classifications were obtained using the same methods as for the Santo lists ([27] to [35]). The discrepancies between the results obtained from proportions of shared cognates ([35]) and even the theoretically best χ^2-derived measurement ([27]) are much greater than those observed in the internal classification of the languages and dialects of Espiritu Santo.

Pentecostan Subgroup

The well-defined Pentecostan subgroup of languages (Apma, Seke, Sa, and Sowa; Raga, although on Pentecost, is, according to the cognate-based classification, more closely related to Aoban) is identified in all classifications.

Ureparapara

It would seem that the two lists from Ureparapara (Lehali and Lehalurup) should be likely to reflect communalects more closely related to each other than to any other communalect outside

Ureparapara. However, Lehali and Lehalurup are grouped together only in the classification based on proportions of shared cognates, whereas the χ^2-based classifications mostly give Lehalurup as more closely related to Motlav than to Lehali.

Bek

Bek is an extinct language of Vanua-Lava and should presumably be grouped with the other languages of Vanua-Lava (Vatrata, Mosina, etc.). It is grouped so only in the classification based on proportions of shared cognates, but most χ^2-based classifications give it as more closely related to Motlav and Ureparapara than to Vanua-Lava.

The Dichotomy Banks-Torres vs. Aoba-Maewo

All nine classifications agree in showing the same dichotomy between the languages of the Banks and Torres on one side and Aoba and Maewo on the other. The subsequent subdivisions of these two groups, however, show considerable disagreement, especially in the case of the Banks and Torres.

Position of Mota

The cognate-based classification groups Mota with the neighbouring island Motlav, within a subgroup encompassing Vanua-Lava and Ureparapara ([35]), whereas the classification based on the $(\chi^2\text{-b})/N$ measurement has it grouped with the languages of more distant Gaua ([27]).

CONCLUSION

The performance of the χ^2-based measurements is here all the more puzzling in that the results which they had given in the case of Santo appeared more likely to be true than those obtained from proportions of shared cognates, whenever discrepancies arose.

Whether the Northeast New Hebrides test should be interpreted as a failure of the χ^2-derived measurements is unclear. If these methods have indeed failed here, it could be argued that this failure should be blamed on the much greater number of irrelevant affixes in these wordlists than were present in the Santo lists (for the full lists see Tryon 1976:175-537), in which case a new classification should be obtained from the same wordlists, this time stripped of as many affixes as can be identified. It could also be blamed on the suspected presence of a large number of communalects the complex phonologies of which are inadequately reflected in the lists, to which it should be retorted that all the NE Santo languages except Sakao were represented by phonologically inadequate lists (some, such as Tur and Butmas, having been taken by phonetically untrained people, some, such as

Shark Bay and Lorediakarkar, having such extraordinary phonologies (Guy 1978) that subtle and yet phonemic distinctions would normally be missed in the too often hurried process of eliciting a sample wordlist); still, these languages appeared to have been correctly classified.

THIRD TEST: CORRECTED LISTS FROM THE NORTHEAST NEW HEBRIDES

Affixes were edited out of the computer file containing the 34 lists of the Northeast New Hebrides in order to test the hypothesis that the discrepancies observed between cognate-based and χ^2-based classifications were due to the presence of irrelevant affixes in these lists. The modifications affected mostly the first part of the lists, composed mainly of nouns from which prefixed articles and pronominal possessive suffixes were stripped, and list items number 154 through to 186, from which a number of different preposed or prefixed markers, apparently required by the adjectival meanings of these items, were removed. Undoubtedly, a few affixes are bound to have been overlooked in the process, and a few words abusively shorn of beginnings and endings which only appeared to be affixes.

CLASSIFICATIONS AND ANALYSIS

The classifications obtained from these corrected wordlists are given in figures [36] to [43]. There is surprisingly little difference between the classifications obtained before and after the removal of affixes from the wordlists. The $(\chi^2\text{-b})/N$ measurement taken after the removal of near-empty rows and columns from the tables of possible sound correspondences again shows the Pentecostan languages as a distinct subgroup, and classifies Raga with the languages of Maewo rather than with those of Aoba, a geographically more plausible grouping.

FOURTH TEST: PHONEMICIZED, CORRECTED LISTS

Many communalects of the Northeast New Hebrides, unlike those of Espiritu Santo, are characterized by the presence of a labio-velar series variously transcribed in the lists by k^w, kp^w, m^w, $ŋ^w$, etc. Furthermore, in most, if not all, of these communalects, neither voicing nor prenasalization are emic (ʈ and nd, phonemically distinct, differ in their point (dental vs. alveolar or cacuminal) and manner of articulation (stop vs. flap) and do not therefore constitute an exception to this observation). Single phonemes being thus very often represented in these lists by two or even three symbols, the transliteration system used in the computer file might have so obscured the pattern of sound correspondences that the χ^2 measurements computed from tables of possible sound correspondences could only have been grossly inaccurate.

A new file was therefore prepared by editing the affixless wordlists, and in which putative phonemes were represented by single symbols ([44]).

CLASSIFICATIONS AND ANALYSIS

Upon examining the classifications obtained from these near-phonemic corrected wordlists ([45] to [52]), one is again struck by how closely they agree with the previous classifications, so much so that little information has been gained. Raga is again in most reconstructions classified with Maewo, Bek with Motlav and Ureparapara (rather than with Vanua-Lava), and Mota sometimes with Gaua ([45] and [49]), sometimes in a Vanua-Lava-Ureparepara-Motlav subgroup ([46], [47], [48], [51], and [52]), and once in a Motlav-Ureparapara subgroup ([50]).

APPRAISAL OF THE χ^2-BASED METHODS

COGNATELESS VS. COGNATE-BASED CLASSIFICATIONS

It will be remembered that the purpose of the tests conducted was to evaluate the validity of methods of language classification not requiring the identification of cognate forms by comparing their results against those obtained from proportions of identified shared cognates.

The phylogenetic reconstructions obtained by those methods in the case of Espiritu Santo were extremely encouraging in that the best reconstruction was given by the measurement which the theory had predicted to be the most accurate, and that, where discrepancies arose between this classification and the one produced from proportions of shared cognates, the former appeared more likely to be true.

In the case of the Northeast New Hebrides, however, no such clear-cut assessment could be made.

Bek and Mota, for instance, seem best classified in the tree reconstructed from proportions of shared cognates ([35]), but Raga in that obtained from the $(\chi^2\text{-b})/N$ measurements ([36] and [45]).

The $(\chi^2\text{-b})/N$ measurements grouped Lehalurup not with the other Ureparapara communalect (Lehali), but with Motlav. Such a classification appears almost certainly wrong, until one notices that Lehalurup is on the side of Ureparapara closer to Motlav, and Lehali on the opposite, far side.

Equally perplexing is the classification by the $(\chi^2\text{-b})/N$ measurement of Mota with the languages of Gaua rather than with those of nearer Vanua-Lava. But here again, of the two offshore islands of Vanua-Lava (Motlav and Mota), Mota is the one closer to Gaua.

Finally, the wisdom of choosing the cognate-based classification as an absolute yardstick might be questioned, for such a classification is only as good as the accuracy with which cognates are identified.

Practically nothing was known of the diachronic phonologies of the languages of the New Hebrides when Tryon set out to identify cognate groups for his lexicostatistical classification, other than tentative sound correspondences set up on the evidence of Proto-Oceanic, rather than on the evidence of forms reconstructed within the New Hebrides.

It is difficult to imagine that, under such conditions, many cases of misidentification did not arise, fuelled by spurious resemblances such as encountered between, say, Latin dies and English day, or, on the contrary, by opaque phonological developments such as Modern Greek ματι "eye", from οφθαλμος via ομμα (plural: ομματα). In the absence of published data listing which words were considered reflexes of which putative protoforms, one cannot help wonder whether for instance Lehali nɛ "fish" was recognized as a possible regular reflex of *iga, Roria ðoðono "house" as a reflex of *vanua, Sakao ǀam "come" as a reflex of *lago-mai, but œ́pœs "dog" as a mere loanword.

Furthermore, whenever a word of a given list is wrongly identified with some putative protoform, a bias is introduced in the proportions of cognates between that list and all the rest. Past decisions as to which protoforms doubtful glosses were assigned to would seem more likely to influence future decisions in the same, rather than in the opposite, direction. Thus, instead of cancelling each other out, erroneous recognitions of cognates are likely to have a cumulative effect; whereas in the methods elaborated here measurements of relatedness are calculated afresh and independently for each pair of wordlists, so that a random error in the measurement obtained for one list pair does not affect the others.

χ^2-BASED AND COGNATE-BASED MEASUREMENTS COMPARED

The tables of proportions of cognates, $(\chi^2\text{-}b)/N$ measurements, and linear-correlation measures for the Santo and Northeast New Hebrides lists are given in appendix ([53] to [64]).

Linear-correlation measures computed fom $(\chi^2\text{-}b)/N$ are on the whole lower in absolute values than those computed from proportions of shared cognates (this confirms the suspicion mentioned in chapter 2 that the cumulative effects of phonological decay do not obey a multiplicative rule). They are also generally higher when computed on lists in which one phone or phoneme is represented by a single symbol ([64]) than when computed on lists in which digraphs and trigraphs are used ([62]).

Graphs [65] to [73] show the clear linear patterns formed by the communalects of Espiritu Santo when plotted according to their scores with Tutuba and Aore (which all classifications had as sharing an immediate common ancestor), whether those scores are proportions of shared cognates ([65]) or χ^2-based measurements ([66] to [73]).

A FINAL TEST: 26 LANGUAGES FROM ENGLISH TO JAPANESE

THE LISTS

The data for this test is taken from Bergman 1968, essentially a list of 1000 words in 26 languages set out in such an extremely convenient manner that the temptation to submit the method to yet another trial was irresistible.

Twenty-six languages ranging from English to Japanese via Swahili, Hebrew, Hungarian, etc., can under no circumstances constitute a statistically representative sample of the world languages, all the more so that the list includes Esperanto, an artificial hybrid, and thereby outside the scope of application of the theory of phylogenetic reconstruction developed here and in Guy 1980 and 1981. Furthermore, their vocabulary was sampled using the 196-item list in Dyen, James, & Cole 1967, a number of which items are not to be found in Bergman's dictionary, the 1000 words of which were selected for their everyday usefulness. Even after substitutes, such as "street" for "road", were used, the lists input were only 132 items long.

These lists were transliterated in the same undemanding manner as was used for the Santo wordlists (see sample [74]).

EXPECTED RESULTS

As I was inputting these lists, it became increasingly evident that most languages fell into one of three great groups (Romance, Germanic, and Slavonic), and that, inside Germanic, English showed so many cognates with Danish that it would very probably come out as a Scandinavian language, rather than being classified with German and Dutch. With the foreknowledge that Hungarian and Finnish belonged to the same language family, it was possible to recognize just a handful of cognates; so for Arabic and Hebrew, but it is doubtful whether these cognates could have been identified without such knowledge.

Greek appeared unrelated to anything else, except perhaps to Indonesian on the basis of two words apparently unique to both: "head" (Greek κεφαλι, Indonesian kepala), and "eye" (Greek ματι, Indonesian mata). Turkish yielded no convincing cognates with Hungarian or Finnish, only Arabic loanwords, and Swahili and Japanese appeared unrelated to anything else, once the Arabic loanwords in Swahili had been discounted.

If the method was to work at all on such short wordlists from such an unbalanced language sample, the best which could be hoped for was a three-way classification of the Romance, Germanic, and Slavonic languages, with the rest here and there in a haphazard manner.

ANALYSIS OF THE RESULTS OBTAINED

The classification obtained using the best theoretical measurement, $(\chi^2-b)/N$, is, in this light, surprisingly good ([75]). Rumanian is correctly grouped with French and Italian rather than with Portuguese and Spanish. Even more surprising, and pleasantly so, is the grouping of Danish with Swedish rather than with Norwegian, for, as I was inputting the lists into the computer file, and noticing the greater similarity of Danish with Norwegian than with Swedish, I could not help sadly reflecting on how the reconstruction was bound to classify it erroneously with Norwegian.

CONCLUSION

It is extremely doubtful whether a better lexicostatistical classification could have been obtained by the laborious process of identifying probable cognate forms on the sole evidence of these 132-item wordlists. As for the costs involved, the computations and the drawing of the phylogenetic tree took six minutes and ten seconds of computer time (CPU time), and, even if computer time is valued at $500 an hour, one may well wonder if the time spent in identifying cognates is worth the money saved and should not be better used in more creative endeavours.

APPENDIX:

TABLES and FIGURES

	Hiw	Toga		Hiw	Toga
1	kwutik	kwutuk	27	kiok	kilok
2	nuyuk	luk	28	moso	m7to
3	toi9ok	del9ok	29	Gayawa	G7law7
4	mitik	muduk	30	te	ta
5	Gamiek	garamek	31	teGtuw7	eht7
6	yiwok	luwok	32	taweyok	wokok
7	mutok	m7tok	33	kwuGo9ok	wuGuk
8	9gok	9uhuk	34	Go9ok	
9	nuypaysetayuw	ulpeleh	35	pioGo9ok	p7Grok
10	paitek	Gumik	36	woGo9ok	wul7Gok
11	pinik	pinik	37	taGa	d7rok
12	myok	limok	38	siGik	hurik
13	pusmyok	puhlaBa	39	Gitik	G7litpi7k
14	tinamok	G7rok	40	miniGa	m7nuG7
15	tit	hih	41	G7Gak7	n7Gerak7
16	Gak	G7Ga	42	yayot	lalih
17	kwok	tukwok	43	tuGmotok	hurm7tok
19	BoGok	nuwet	44	tomok	t7mok
20	wus7	Bar7	45	teta	r7me
21	wotikw7	tukok	46	moGoye	natuk
22	BaG7yaB7	mokwmekw	47	t7mwen	t79wen
23	nuy	Gurok	48	yekwen	l7kweBin7
24	Bat	hu9wok	49	t7mwen eki7	
25	yusuk	l7hok	50	yok	yek
26	popo	tihik			

[1]

First 50 items of the Hiw and Toga wordlists (Tryon 1976) respelt for computer processing

Hiw

Toga	)	7	9	B	G	U	a	e	g	i	k	m	n	o	p	s	t	u	w	y	
)	1	9	0	4	6	0	2	0	0	3	3	1	1	4	0	0	4	1	1	0	)
3	0	3	2	0	4	0	2	9	0	2	2	5	6	2	0	1	1	3	2	2	3
7	6	76	9	26	86	1	83	55	0	74	35	50	38	74	15	19	71	35	49	33	7
9	0	5	14	1	6	1	10	9	1	6	6	10	5	17	1	1	12	7	12	4	9
B	2	17	1	42	23	0	17	14	0	23	7	8	11	27	3	8	18	12	7	12	B
G	0	21	6	15	63	0	39	17	0	17	16	11	7	29	6	5	16	15	11	8	G
a	2	31	4	12	49	0	60	20	0	23	13	24	6	28	9	10	30	17	19	10	a
d	0	3	1	0	5	0	3	1	0	9	2	3	1	5	0	0	6	0	2	3	d
e	0	33	12	15	40	0	42	48	0	23	16	30	20	49	12	9	42	18	28	23	e
g	0	0	0	0	1	0	1	1	0	1	1	1	0	0	0	0	0	0	0	0	g
h	0	24	1	2	18	0	24	8	1	16	11	7	11	26	7	21	32	24	25	10	h
i	0	21	3	14	27	0	11	27	0	58	23	13	21	28	13	11	26	10	11	19	i
k	0	18	7	5	23	0	33	14	1	26	58	17	7	46	8	5	28	13	26	15	k
l	3	20	4	10	28	0	33	26	0	38	17	20	12	44	8	9	26	18	16	46	l
m	1	22	3	10	31	0	32	17	0	32	17	45	8	32	2	2	26	7	5	9	m
n	1	24	3	4	34	1	15	31	0	23	9	18	42	30	5	5	20	12	12	7	n
o	0	19	12	5	31	0	16	10	0	16	28	14	12	67	3	8	32	8	18	12	o
p	0	4	1	0	6	0	3	9	0	13	5	2	5	6	17	2	4	6	2	6	p
r	3	26	9	9	64	0	20	22	0	30	18	18	15	38	6	8	24	10	12	4	r
s	0	2	0	4	2	0	3	5	0	4	2	1	2	6	1	3	4	0	0	1	s
t	3	27	5	9	29	1	21	21	0	34	14	25	4	42	8	1	70	21	20	9	t
u	3	17	5	14	28	1	13	9	2	38	27	16	7	32	7	8	29	37	22	21	u
w	0	29	7	6	20	0	32	17	0	10	23	12	8	33	2	9	29	14	46	11	w
y	0	0	1	0	0	0	0	2	0	1	1	0	0	1	0	0	0	0	0	1	y
	)	7	9	B	G	U	a	e	g	i	k	m	n	o	p	s	t	u	w	y	

[2]

Table of possible sound correspondences between Hiw and Toga (Torres Is.) computed from the wordlists in Tryon 1976.

+ B:B	11.825	+ s:h	6.018	u:h	3.349	9:y	2.487
+ p:p	10.852	+ o:o	5.966	n:3	3.067	e:y	2.343
+ n:n	8.952	+ i:i	5.779	e:n	2.912	B:)	2.314
+ k:k	8.758	+ u:u	5.570	a:G	2.850	B:s	2.314
+ 9:9	7.867	+ a:a	5.514	9:o	2.833	p:i	2.220
+ y:l	7.634	+ G:r	5.340	U:9	2.821	+ s:s	2.142
+ w:w	7.131	+ e:e	3.648	g:9	2.821	+ 7:7	2.108
+ m:m	6.946	+ e:3	3.586	i:d	2.775	n:i	2.103
+ t:t	6.843	+ 7:)	3.583	w:h	2.657	+):)	2.096
+ G:G	6.066	g:u	3.416	k:o	2.496	G:a	2.049

[3]

Most likely sound correspondences between Hiw and Toga computed from table [2]

Sound pairs are listed in order of decreasing likelihood. For instance (first column, sixth line), Hiw y is given as corresponding to Toga l. The number following the pair y:l (7.634) is a (rough) measure of the probability of this correspondence being due to factors other than chance. A plus sign (+) precedes those sound correspondences which seem to apply beyond reasonable doubt.

Hiw

Sakao	)	7	9	B	G	U	a	e	g	i	k	m	n	o	p	s	t	u	w	y	
)	3	18	4	12	45	0	19	14	0	35	30	15	13	36	6	6	39	32	15	15	)
*	0	16	5	4	17	0	22	14	0	16	19	10	4	15	4	7	16	10	9	10	*
0	1	5	1	7	9	0	4	8	0	5	3	3	2	15	2	3	9	4	2	9	0
9	2	4	5	1	7	0	4	4	1	5	6	4	1	9	1	1	3	1	5	1	9
B	0	10	6	1	14	0	12	7	0	21	23	11	9	28	10	4	13	12	18	9	B
D	4	20	0	10	25	1	20	25	0	26	12	26	10	31	9	4	41	11	17	14	D
E	4	25	3	18	40	0	46	36	0	28	11	25	14	49	10	10	58	20	25	26	E
G	0	7	3	3	18	0	12	8	0	12	7	11	3	13	1	2	7	10	8	6	G
O	0	14	12	4	17	1	7	8	2	9	16	9	6	36	6	5	16	10	20	2	O
R	0	0	0	0	0	0	0	0	0	0	0	1	0	3	0	2	0	0	0	0	R
U	1	4	2	2	4	0	3	0	1	7	6	4	5	7	1	1	4	4	4	0	U
a	3	42	11	17	61	0	57	52	0	37	29	42	30	60	15	14	45	21	37	30	a
e	1	8	6	3	8	0	5	4	0	13	11	4	0	18	3	4	11	3	6	3	e
h	1	5	2	4	12	0	16	7	0	8	8	6	9	16	3	4	13	7	6	8	h
i	0	8	1	4	7	0	4	15	0	13	7	8	8	13	6	1	12	4	6	4	i
j	0	7	3	9	11	0	15	2	0	11	3	4	5	21	1	5	19	3	5	5	j
k	2	10	2	2	15	0	9	8	0	13	12	11	4	16	2	6	15	3	5	6	k
l	3	25	12	13	31	0	21	26	0	32	23	21	12	56	7	7	38	15	20	25	l
m	0	8	3	3	14	0	7	3	0	3	9	9	7	10	1	0	10	16	7	0	m
n	1	31	8	14	50	0	38	37	1	43	39	27	29	69	11	12	57	29	28	30	n
o	0	7	1	0	2	1	5	2	0	4	3	3	4	9	5	6	5	4	4	5	o
p	0	7	4	7	5	0	7	2	0	5	2	1	2	14	4	2	6	3	2	5	p
r	1	35	12	15	66	2	49	29	0	45	25	32	23	62	15	15	39	20	27	18	r
s	0	8	3	5	15	0	19	5	1	6	6	7	3	11	1	4	13	4	1	1	s
t	1	5	1	2	11	0	9	7	0	12	3	3	14	9	0	2	16	7	6	2	t
u	1	4	5	4	19	0	12	5	0	5	11	2	2	22	1	2	5	3	6	7	u
w	0	10	1	2	8	0	7	8	0	5	3	3	2	5	3	0	5	4	6	5	w
y	0	9	1	4	20	1	13	4	0	13	6	5	8	12	5	1	5	5	10	3	y
	)	7	9	B	G	U	a	e	g	i	k	m	n	o	p	s	t	u	w	y	

[4]

Table of sound possible correspondences between Hiw (Torres Islands) and Sakao (NE Espiritu Santo) computed from the wordlists in Tryon 1976.

	s:R	4.979	+	U:o	3.392		p:o	2.596	+	p:B	2.309
	u:m	4.763	+	9:9	3.134	+	B:j	2.588		m:D	2.298
	n:t	4.510	+	k:B	3.064		g:s	2.546		G:u	2.254
	g:O	3.920		B:p	3.021	+	o:O	2.539		o:u	2.223
	9:O	3.883		):9	2.877		B:O	2.528		G:y	2.205
+	u:)	3.778		e:i	2.841		y:O	2.436		k:*	2.129
	g:U	3.703		w:O	2.794		9:e	2.432		t:j	2.037
	g:9	3.539		o:R	2.704	+	t:D	2.406		w:B	2.005
	s:o	3.444		e:a	2.667	+	U:y	2.388		U:r	1.992
	a:s	3.394		t:E	2.624		7:w	2.350	+	G:r	1.987

[5]

Most likely sound correspondences between Hiw and Sakao computed from table [4]

A plus sign (+) precedes those sound correspondences which seem to apply beyond reasonable doubt.

Hiw

		)	7	9	B	G	a	e	g	i	k	m	n	o	p	s	t	u	w	y	
	a	5	84	18	33	95	86	73	0	76	55	67	32	125	22	24	92	22	50	49	a
	e	5	21	1	12	29	38	32	0	24	23	19	12	36	15	14	43	14	17	16	e
	g	0	23	10	3	35	21	25	1	31	16	19	10	34	7	8	31	10	17	15	g
T	i	2	49	11	33	74	56	42	1	55	23	37	22	78	14	15	56	21	33	29	i
o	k	2	7	3	2	16	10	4	0	13	10	8	6	13	3	0	7	1	5	3	k
l	l	5	19	5	14	18	24	14	0	17	5	12	10	32	2	2	32	7	9	22	l
o	m	0	9	4	1	12	12	5	0	2	6	6	1	9	2	6	7	6	4	3	m
m	n	2	23	2	9	35	29	22	0	26	18	29	19	48	4	5	34	9	14	18	n
a	o	3	19	11	5	37	25	18	1	29	22	17	14	50	8	8	30	16	15	15	o
k	p	0	9	6	2	10	7	8	0	9	15	6	6	19	1	4	10	7	16	8	p
o	r	1	18	5	9	38	26	15	0	23	13	16	7	42	8	10	22	13	11	12	r
	s	0	14	0	9	13	13	13	0	12	11	4	5	18	5	6	21	2	12	4	s
	t	2	25	1	13	28	18	34	0	25	17	24	10	39	15	7	38	14	19	15	t
	u	3	25	6	17	54	39	20	0	37	35	23	21	65	7	11	39	23	28	27	u
	v	3	18	2	23	26	28	10	0	22	13	10	16	41	11	7	19	14	9	16	v
	z	0	9	5	2	12	11	9	1	11	8	9	3	18	3	5	13	4	5	3	z
		)	7	9	B	G	a	e	g	i	k	m	n	o	p	s	t	u	w	y	

[6]

Table of possible sound correspondences between Hiw (Torres Islands) and Tolomako (NW Espiritu Santo), computed from the wordlists in Tryon 1976

+	B:v	4.170		B:i	2.268		B:l	1.823		):k	1.571
	w:p	3.396		9:o	2.237		k:u	1.821	+	k:k	1.570
+	y:l	2.992		p:e	2.147	+	7:a	1.805		t:s	1.568
	g:z	2.911	+	m:n	2.119	+	G:r	1.751		u:m	1.529
	):l	2.852	+	9:g	2.081		n:v	1.751		t:e	1.522
+	k:p	2.623	+	9:m	1.908	+	e:e	1.706		g:o	1.513
	e:t	2.588		t:l	1.901	+	u:u	1.683		a:m	1.473
+	p:t	2.456		9:z	1.898		s:e	1.645		i:k	1.458
	s:m	2.399	+	n:n	1.866		p:v	1.623		w:s	1.425
	9:p	2.338	+	):e	1.849		g:g	1.615		G:k	1.420

[7]

Most likely sound correspondences between Hiw and Tolomako computed from table [6]

A plus sign (+) precedes those sound correspondences which seem to apply beyond reasonable doubt.

Tolomako

Sakao

	a	e	g	i	k	l	m	n	o	p	r	s	t	u	v	z	
)	50	20	17	50	2	13	5	22	25	15	29	10	21	33	12	8	)
*	41	17	18	28	0	13	3	16	10	5	6	4	21	10	4	8	*
0	21	4	5	14	1	8	2	9	4	0	7	1	8	15	10	0	0
9	11	1	4	5	0	1	0	1	2	2	2	0	3	6	4	1	9
B	32	21	13	23	6	5	1	8	20	16	10	8	14	14	10	4	B
D	62	21	19	47	5	15	3	21	23	7	15	7	44	18	11	3	D
E	84	56	16	53	14	26	10	43	31	4	33	19	26	31	21	13	E
G	25	10	10	18	8	5	2	9	11	2	12	3	4	12	7	0	G
O	27	15	8	17	2	5	2	7	18	14	9	6	10	19	14	4	O
R	1	0	0	1	0	0	0	0	1	0	0	0	2	2	2	1	R
U	11	0	3	5	4	1	0	2	4	1	3	1	2	4	2	2	U
a	148	27	42	74	18	37	8	36	43	14	38	14	32	43	38	13	a
e	17	7	6	10	3	2	0	3	6	3	2	5	7	8	5	2	e
h	17	8	10	12	1	5	1	9	10	1	8	8	3	12	5	1	h
i	25	13	14	25	2	7	0	6	9	3	7	9	16	15	6	3	i
j	30	13	7	16	4	5	3	12	8	1	9	4	7	10	10	5	j
k	31	14	9	21	7	11	2	9	7	7	8	3	7	12	8	3	k
l	84	31	23	39	14	43	6	24	33	9	20	9	21	26	20	8	l
m	22	6	1	12	4	1	7	12	8	7	8	2	6	10	2	3	m
n	130	32	34	67	9	30	11	56	33	8	32	14	33	50	34	15	n
o	11	7	2	7	1	1	2	3	7	2	3	3	6	4	7	1	o
p	8	6	2	12	0	6	0	6	8	4	3	3	6	4	4	1	p
r	90	29	28	72	9	18	4	31	24	14	46	7	28	62	26	17	r
s	13	17	7	18	3	2	3	4	15	5	9	14	4	9	7	6	s
t	19	6	9	8	2	9	0	4	12	0	1	1	14	6	6	0	t
u	27	9	8	9	4	6	1	6	10	3	10	3	8	18	7	1	u
w	10	3	3	9	1	3	0	1	3	2	5	3	7	10	2	1	w
y	12	6	5	12	4	3	1	5	9	3	3	4	10	15	4	2	y
	a	e	g	i	k	l	m	n	o	p	r	s	t	u	v	z	

[8]

Table of possible sound correspondences between Tolomako (NW Espiritu Santo) and Sakao (NE Espiritu Santo) computed from the wordlists in Tryon 1976

+	l:l	5.170	+	r:r	2.981		t:*	2.110		l:t	1.905
+	t:D	5.131	+	a:a	2.971		r:)	2.091		v:0	1.890
+	s:s	5.075		k:U	2.954		v:R	2.082		g:*	1.872
+	p:B	4.529	+	k:G	2.778		p:)	2.076		n:m	1.807
+	m:m	4.492	+	s:h	2.657		u:w	2.021	+	k:k	1.805
	p:O	4.284		e:s	2.398		s:i	2.018		o:s	1.797
+	e:E	3.936		p:m	2.334		o:t	2.016	+	z:r	1.774
	u:r	3.057	+	u:y	2.334		a:n	1.943		e:B	1.727
+	n:n	3.056		n:E	2.223		v:o	1.940	+	z:s	1.714
+	t:t	3.044	+	u:u	2.118	+	u:O	1.905		t:i	1.714

[9]

Most likely sound correspondences between Tolomako and Sakao computed from table [8]

A plus sign (+) precedes those sound correspondences which seem to apply beyond reasonable doubt.

Proportion of cognates	Before removing near-empty rows and columns χ^2	N	d.f.	After removing near-empty rows and columns χ^2	N	d.f.
0.000	252.962	6120	225	155.486	5095	121
0.028	251.463	5780	225	116.694	4850	121
0.122	399.398	5989	225	221.050	5215	132
0.159	301.091	6458	225	189.851	5600	132
0.146	389.417	6219	225	240.016	5197	121
0.228	423.549	6345	225	242.799	5393	121
0.285	562.113	6072	225	317.023	5153	121
0.321	676.607	6684	225	389.959	5618	121
0.362	778.166	6744	225	409.566	5692	121
0.451	1150.258	6521	225	732.136	5621	121
0.488	1346.215	6351	225	804.221	5263	110
0.581	1911.311	6816	225	1227.185	5933	132
0.646	1958.424	6538	225	1119.829	5425	110
0.614	1967.853	6961	225	1276.273	6141	132
0.707	2188.257	6611	225	1508.003	5615	121
0.797	2687.211	6909	225	1769.228	5880	121
0.805	2763.666	6921	225	1837.656	5958	121
0.829	3217.856	6948	225	2023.757	5959	121
0.911	3465.074	7021	225	2215.509	6071	121
0.955	3956.957	7025	225	2513.826	6052	121
0.000	204.773	5909	225	119.841	5213	143
0.073	323.486	6159	225	216.380	5276	132
0.077	292.477	6406	225	160.695	5613	132
0.207	438.192	6578	225	284.139	5601	121
0.191	390.567	6399	225	248.357	5563	132
0.268	522.601	6511	225	323.545	5578	132
0.317	926.476	6467	225	526.903	5558	121
0.313	692.055	6534	225	482.324	5519	121
0.419	835.413	6431	225	581.510	5478	121
0.459	1278.740	6549	225	896.436	5811	144
0.459	895.775	6371	225	589.611	5520	132
0.500	1211.330	6447	225	758.264	5442	121
0.638	2030.211	6731	225	1136.235	5590	110
0.675	2128.619	6670	225	1341.494	5703	121
0.626	1789.963	6658	225	1210.929	5721	121
0.768	2603.491	7048	225	1730.331	5988	121
0.793	2698.204	6785	225	1702.726	5871	121
0.825	3050.081	6911	225	1970.602	5927	121
0.886	3481.685	6898	225	2290.464	5932	121
0.972	3975.816	6881	225	2621.797	5940	121

[10]

Relationship between proportions of shared cognates
and the χ^2 statistic of tables of possible sound correspondences
Model: Tolomako (NW Espiritu Santo)

Proportion of cognates	Before removing near-empty rows and columns χ^2	N	d.f.	After removing near-empty rows and columns χ^2	N	d.f.
0.000	913.398	5526	729	140.454	2457	81
0.054	770.278	5370	729	73.011	2453	90
0.115	1023.805	5504	729	166.345	2455	81
0.151	1081.821	5792	729	192.723	2803	90
0.237	1465.654	5529	729	325.791	2617	81
0.269	1493.268	5750	729	271.754	2696	81
0.290	1596.360	5725	729	356.761	2741	90
0.312	1713.425	5899	729	496.174	2967	99
0.430	2409.794	6039	729	505.345	2792	72
0.444	2737.085	6034	729	814.258	3084	90
0.466	2777.892	6012	729	967.427	3044	90
0.556	3707.425	6180	729	1161.613	3134	90
0.581	3807.229	5840	729	1022.817	2778	72
0.695	4978.934	6057	729	1670.721	3081	90
0.703	5392.577	6048	729	1543.041	3096	90
0.746	5638.416	6187	729	1754.829	3304	99
0.814	7244.837	6412	729	2476.364	3514	100
0.846	7263.802	6361	729	1815.115	3031	72
0.907	8196.467	6425	729	2673.743	3573	100
0.950	8916.705	6401	729	3063.965	3526	100
0.000	852.937	5470	729	118.800	2284	80
0.050	870.909	5746	702	96.847	2743	90
0.125	946.061	6002	729	152.177	2990	100
0.161	1016.934	5723	729	155.371	2669	90
0.197	1187.418	5742	729	250.060	2767	90
0.233	1380.362	5930	729	245.535	2671	72
0.276	1801.735	5919	729	310.967	2637	72
0.362	2264.963	6055	729	413.261	2524	63
0.373	1932.040	5736	729	534.576	2780	90
0.484	3000.728	5948	729	837.128	3082	99
0.502	3032.138	6168	729	875.776	3089	90
0.548	4200.408	6286	729	1030.306	3201	90
0.581	3865.960	6141	729	1237.017	2974	81
0.659	4597.014	6298	729	1453.850	3308	100
0.685	4679.915	6065	729	1407.309	2962	81
0.703	5750.868	6256	729	1729.003	3229	90
0.810	6836.306	6171	729	2089.593	3444	100
0.832	6906.766	6345	729	2139.266	3120	81
0.867	7685.709	6425	729	2461.687	3208	81
0.950	8802.963	6450	729	2868.016	3409	90

[11]

Relationship between proportions of shared cognates
and the χ^2 statistic of tables of possible sound correspondences
Model: Sakao (NE Espiritu Santo)

```
4200 *
4000 *                                              X6X
3800 *                                             22X
3600 *                                            XX
3400 *                                         X 3XX
3200 *                                        X  3X
3000 *                                       23 2
2800 *                                    X 23
2600 *                                    X3 3
2400 *                                 XXXX X
2200 *                                X443
2000 *                            XXX3  X
1800 *                           XX XX
1600 *                           242X
1400 *                       3 3X
1200 *                      662 2
1000 *               X XXX2X  X
 800 *              22X3232 X
 600 *           33323XXX
 400 *   2 X244524
 200 9X2573  X
     *.........*.........*.........*.........*.........*
     0%       20%       40%       60%       80%       100%
```

Proportion of shared cognates

[12]

Relationship between χ^2 values of tables of possible sound correspondences and proportions of shared cognates.
Model: Tolomako (NW Espiritu Santo). χ^2 values computed on tables of possible sound correspondences before removal of columns and rows containing cells with expected frequencies of less than 5

(Graph adapted from output from GLIM © Royal Statistical Society, London)

```
2720 *
2560 *                                              27X
2400 *                                             3X
2240 *                                           42X
2080 *                                        XX XX
1920 *                                       X3 2X
1760 *                                    X226
1600 *                                 X  2XXX
1440 *                                X343
1280 *                            XX22 X2X
1120 *                           X22X2
 960 *                      X  2 23
 800 *                      562X3
 640 *               X  XX4 X3X
 480 *            XXX42533 X
 320 *      X X325223X
 160 9X2593XX532 2
   0 *
     *.........*.........*.........*.........*.........*
     0%       20%       40%       60%       80%       100%
```

χ²

Proportion of shared cognates

[13]

Relationship between χ^2 values of tables of possible sound correspondences and proportions of shared cognates.
Model: Tolomako (NW Espiritu Santo). χ^2 values computed on tables of possible sound correspondences after removal of columns and rows containing cells with expected frequencies of less than 5

(Graph adapted from output from GLIM © Royal Statistical Society, London)

```
    9500 *
    9000 *                                               4
    8500 *                                             24X
    8000 *                                          X 22
    7500 *                                         3X22
    7000 *                                       X25X
    6500 *                                     X22 X
    6000 *                                  2 22X3
    5500 *                                  X XX
χ²  5000 *                                X2X2X
    4500 *                           X 22523
    4000 *                          X 532
    3500 *                       X X23X
    3000 *                     2343
    2500 *                 2X2332
    2000 *             X  2X23
    1500 *          X75254XX
    1000 9236X445624
     500 *
       0 *
         *.........*.........*.........*.........*.........*
         0%       20%       40%       60%       80%       100%
```

Proportion of shared cognates

[14]

Relationship between χ^2 values of tables of possible sound correspondences and proportions of shared cognates.
Model: Sakao (NE Espiritu Santo). χ^2 values computed on tables of possible sound correspondences before removal of columns and rows containing cells with expected frequencies of less than 5

(Graph adapted from output from GLIM © Royal Statistical Society, London)

```
    3200 *
    3040 *                                              22
    2880 *                                             X23
    2720 *                                           XXX
    2560 *                                          X 22
    2400 *                                        X6XXX
    2240 *                                       X XX
    2080 *                                     X32XX
    1920 *                                  X 22
χ^2 1760 *                                 2X XX   X
    1600 *                               2X 22X  3
    1440 *                             4 323
    1280 *                            2X3
    1120 *                      X2  X23 X
     960 *                      X2 X2XX
     800 *                   2X32X3  X
     640 *                  X2XXX
     480 *          XX  X2342XXX
     320 *      22 3 53342
     160 9235X44342X X2
       0 *  X
         *.........*.........*.........*.........*.........*
         0%       20%       40%       60%       80%       100%
```

Proportion of shared cognates

[15]

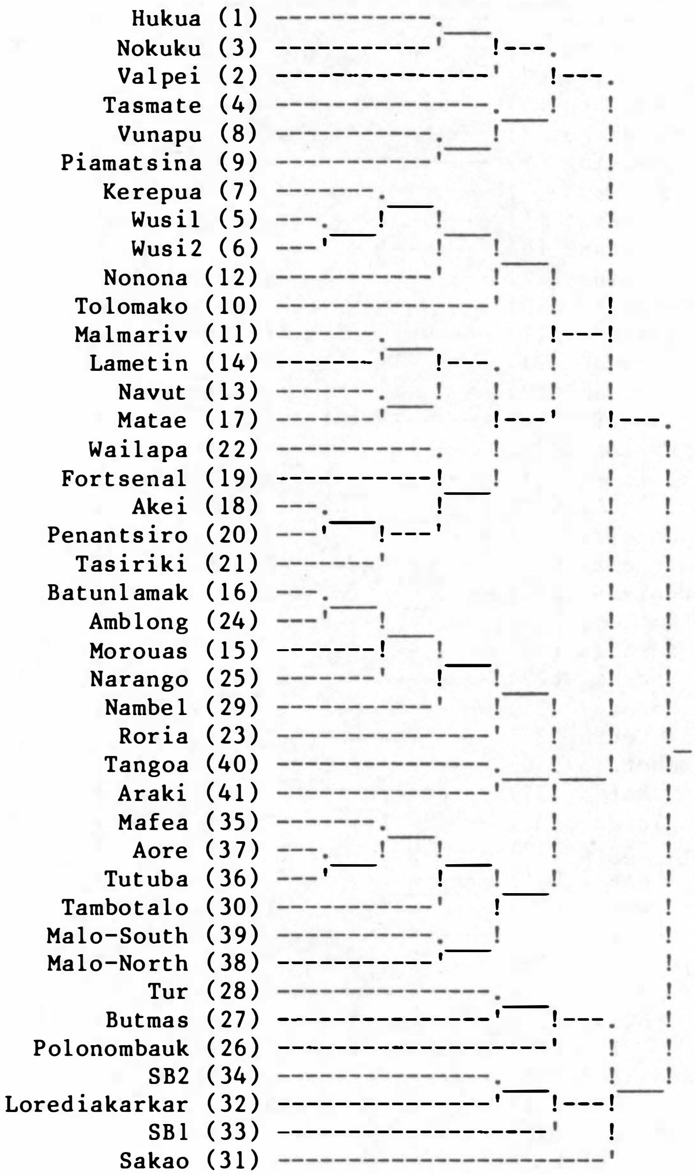
Relationship between χ^2 values of tables of possible sound correspondences and proportions of shared cognates.
Model: Sakao (NW Espiritu Santo). χ^2 values computed on tables of possible sound correspondences after removal of columns and rows containing cells with expected frequencies of less than 5

(Graph adapted from output from GLIM © Royal Statistical Society, London)

154:BLACK
le:loa loloa mae'to maeto lo
lo?o loo maeto taharea nanaeto
loo loo loo Mbo9Mbo9tuMb meto
mome'to loo viriviri viriviri viriviri
viriviri viri?a manaras daina mufir
mifir mofur mifil mofir leho
vu:r fud: fute movute fria
mamaeto virig:a vurixa vurixa=le'ho virixa
virixa
155:WHITE
wo'ka wuko wu?a wuko'o' hulu
hulu hulu vu?o voke voke
voNko voko vuso vso voku
mofoko vuso vuko voke alunu
lulu lulu movoNk fok mulul
mo'fo'o'k mofok mofok mofoNk vuso
wog: roro ror moror foko
vuso voke lulu lulu lulu
lulu
156:RED
hahara hahara kekara a?ara kara
kara kara Hara hahara g:ag:ara
Nkara kara Nkara Ngar NdRae
moNdRai kara ara ara ?a?ara
ha?ara ?a?ara me:Ngar moNdai muNkar^
makar makar mokar hahara hahara
kar kar kar makare: ara
memea g:ag:ara Ndaixa taiha xaxara
xaxara
157:YELLOW
9aora mWacina malke'tSa matSine mWatSina
matSina matSina matSina matSina matSina
marimaria matSina matSina mantSina matSina
mome'tSina matSina anoa matSina mo'mansna
ano?a tSori?a me:g:entS matS^ini mo'mansna
me'e'tS memetS e'ne mimisin momatSna matSine
ne'lya * mesine mesine a9o
ma9oa malasilasig:a c^orixa * tSorixa
tSorika

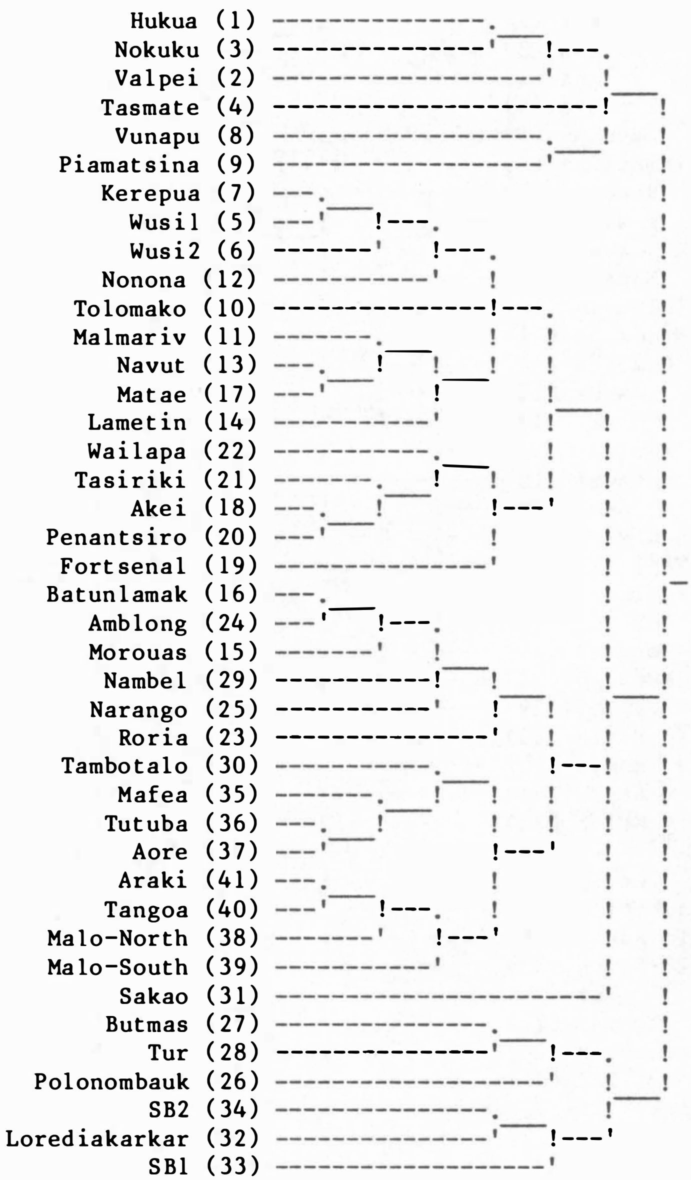
[16]

Four words in 41 languages and dialects of Espiritu Santo
transliterated for automatic processing
(an asterisk indicates a missing word)



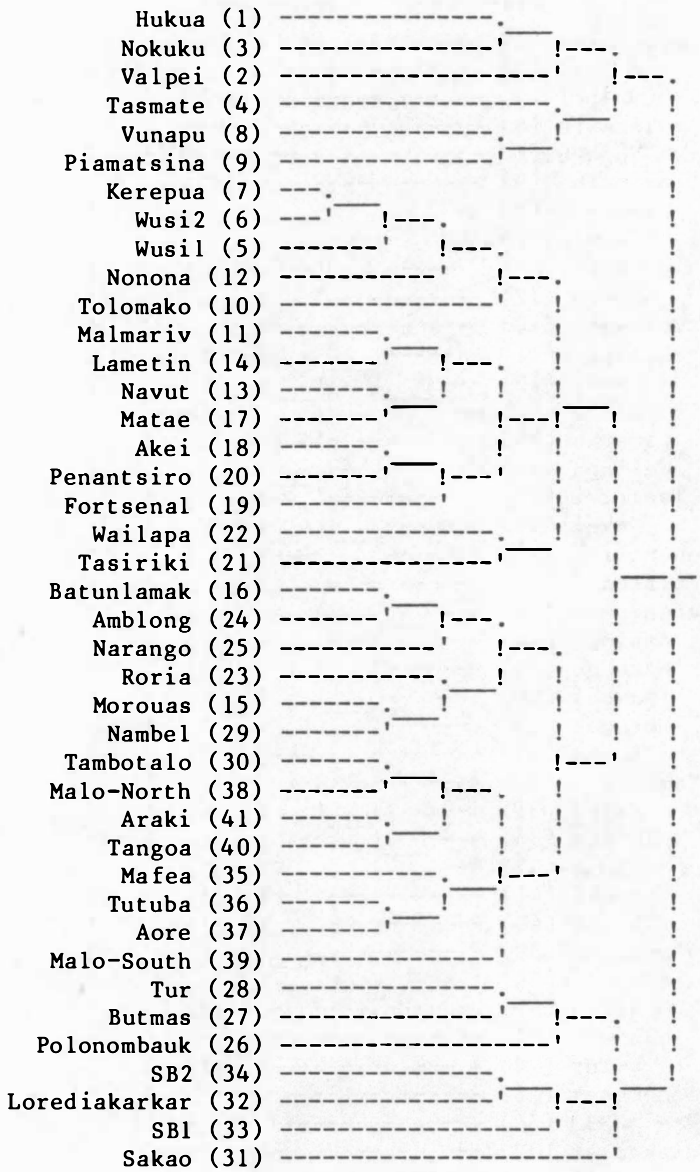
[17]

Phylogeny of 41 communalects of Espiritu Santo reconstituted from $(\chi^2-b)/N$ measurement, taken after removing near-empty rows and columns from table of possible sound correspondences



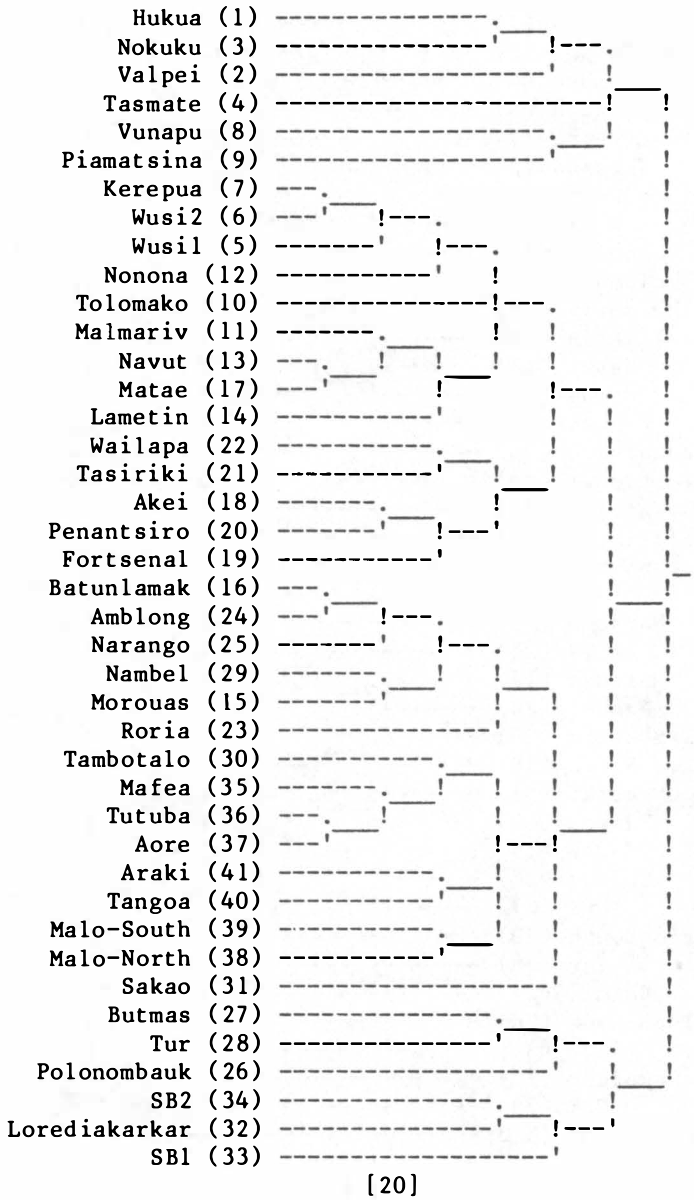
[18]

Phylogeny of 41 communalects of Espiritu Santo reconstituted from $(\chi^2\text{-b})/N$ measurements, taken before removing near-empty rows and columns from table of possible sound correspondences

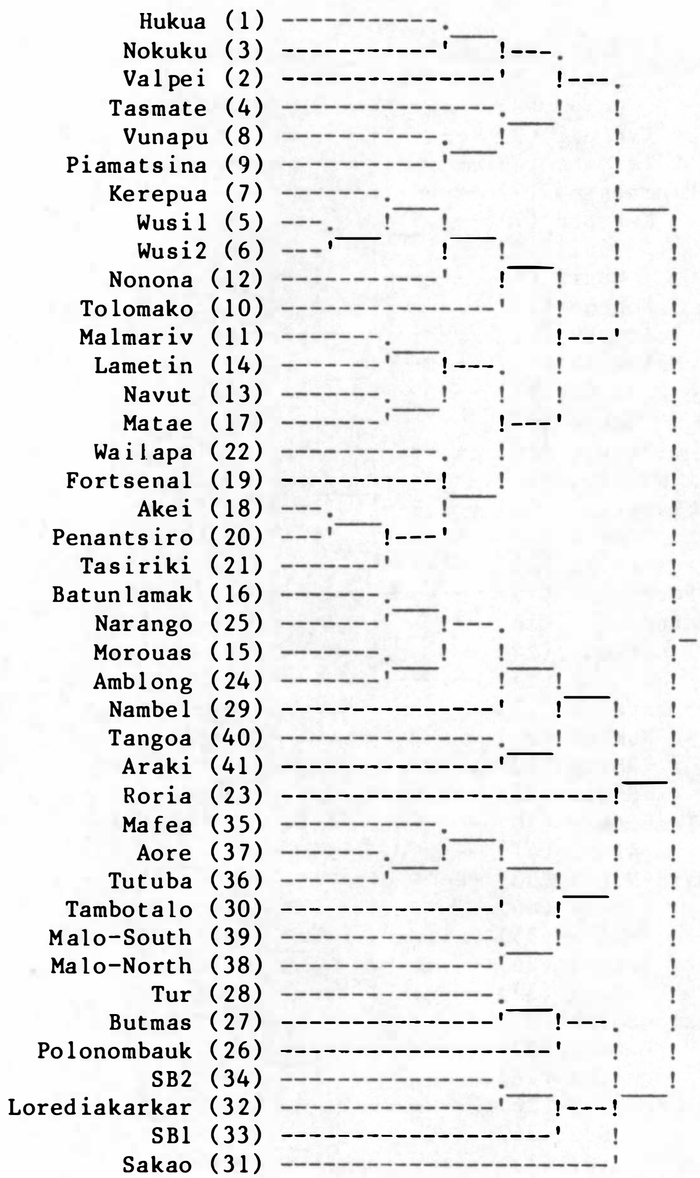


[19]

Phylogeny of 41 communalects of Espiritu Santo reconstituted from χ^2-b measurements, taken after removing near-empty rows and columns from tables of possible sound correspondences

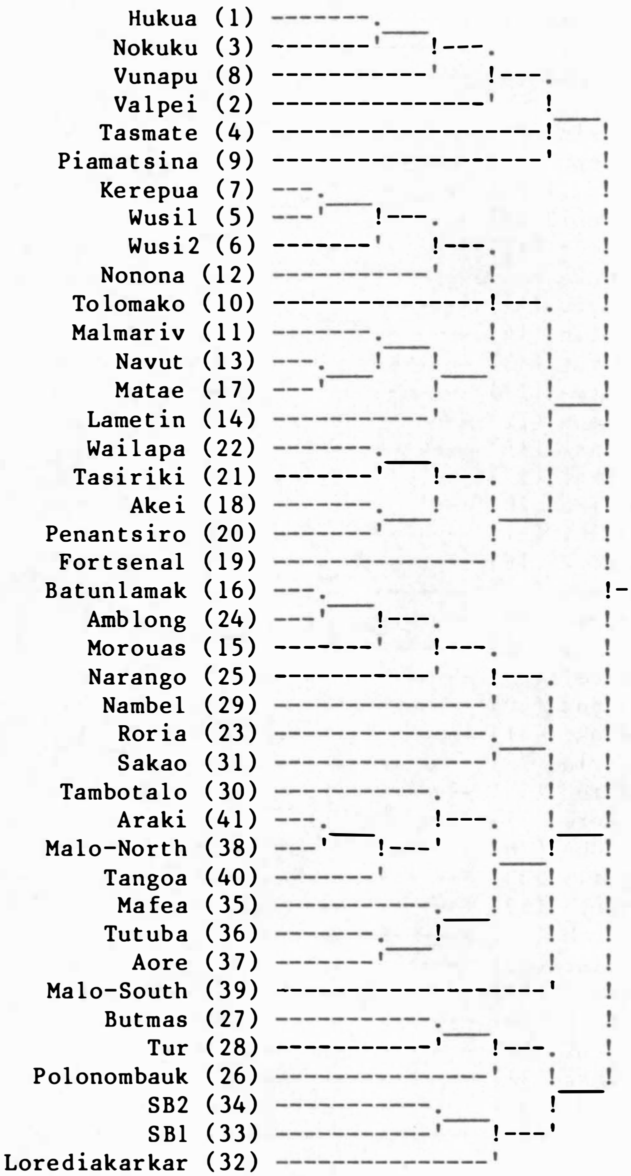


Phylogeny of 41 communalects of Espiritu Santo reconstituted from χ^2-b measurements, taken before removing near-empty rows and columns from tables of possible sound correspondences



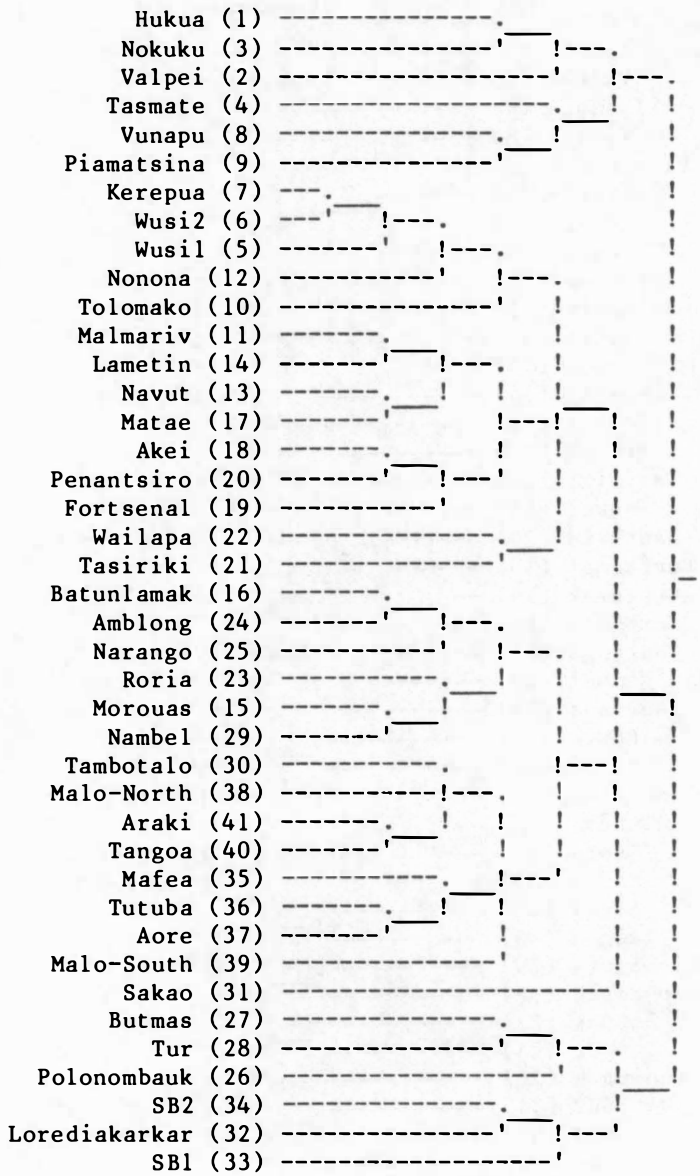
[21]

Phylogeny of 41 communalects of Espiritu Santo reconstituted from χ^2/N measurements, taken after removing near-empty rows and columns from tables of possible sound correspondences



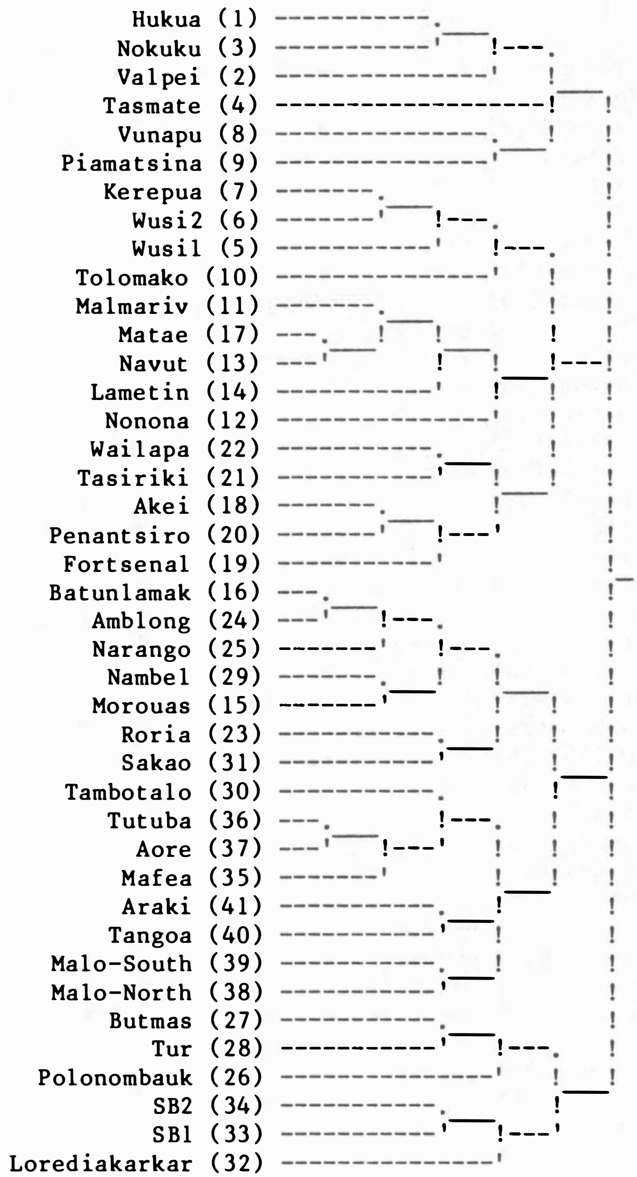
[22]

Phylogeny of 41 communalects of Espiritu Santo reconstituted from χ^2/N measurements, taken before removing near-empty rows and columns from tables of possible sound correspondences



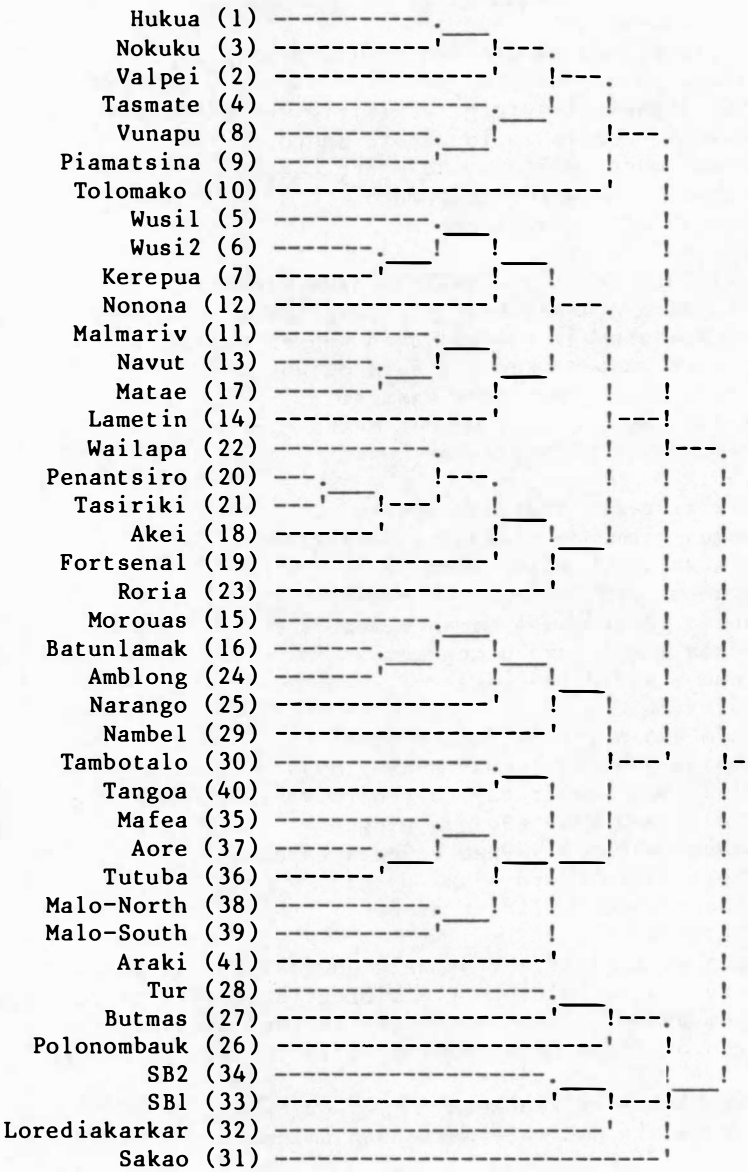
[23]

Phylogeny of 41 communalects of Espiritu Santo reconstituted from χ^2 measurements, taken after removing near-empty rows and columns from tables of possible sound correspondences



[24]

Phylogeny of 41 communalects of Espiritu Santo reconstituted from χ^2 measurements, taken before removing near-empty rows and columns from tables of possible sound correspondences



[25]

Phylogeny of 41 communalects of Espiritu Santo reconstituted from proportions of shared cognates

154:BLACK
miyig:e: mududut nemWyut milie nemleg:leg:
silsilig:a g:akorkor xekurkur korkor g:akurkur
KpWu9KpWu9 wiriwirig: vawiwirig: werweriu we:rwirig:
wewerig: Mbolo Mbolo osooso maeto
maeto maeto maito maeto maeto
maeto maeto meto maeat meto
teme aMbroh mamlek amee
155:WHITE
yui lulu nelul g:eKpWaKpWa nakWag:kWag:
aKpWag:a g:akWakWag:a g:eKpWag:KpWag: g:akWag: g:aKpWag:
g:eKpWauKpWag: wewend vawen wenwen vavWet
g:awet wenwen wenwen si9ara og:Wog:Wo
ovovo ovovo vWas voas wasig:i
mafute mevute lotu movuot maita
tedap asaf palpal adao
156:RED
mie mi nema: lawlaw nalawlaw
memea g:amemee g:elawlaw mame g:ameme
g:elaulau me vamme veme vamme
g:ameme meme meme memea memea
memea memea memea memea memeag:i
memea memea kara memea memea
tememe arfan 9i9i ame
157:YELLOW
9o9o e9o ne9oi 9oi9oi no9oi9oi
a9oa9o g:a9a9 g:ea9a9 a9a9 g:a9a9
g:e9or9or a9a9or va9 to:rto:ro va9a9or
g:aa9 a9a9 a9a9 a9oag:a a9og:a
ya9oxa ea9og:a ya9oxa ea9og:a ea9og:a
a9oka a9oka o9oro a9oak a9og:a
tesesede amsoh litlit aseset
158:GREEN
maig:ete: malag:ehe nemWalg:eh g:e9mal naKpWai
turuturug:a 'u'uruxo g:e'o'orog: ototorog: g:atortorog:
g:emWa9mWal veturturug: vato:te:rug: a9ra vlurlur
totorow NdoNdosma:t Ndosmat malag:esag:a malxesaxa
* * * * *
kesaka kesaka kesakesa * *
temamalkis amalmale maksansan amlamles

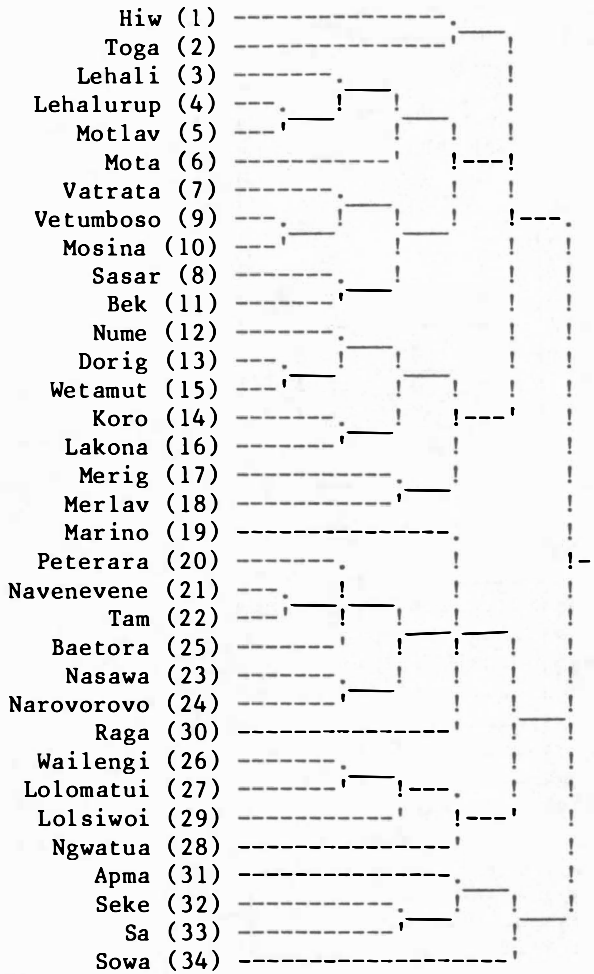
[26]

Five words in 34 languages and dialects of the Northeast New Hebrides, transliterated for automatic processing (an asterisk indicates a missing word)

```
       Hiw (1) -----------.___
      Toga (2) -----------'   !
    Lehali (3) -------.       !___
 Lehalurup (4) ---.___!___    !   !
    Motlav (5) ---'   !   !   !   !
      Bek (11) -------'   !---'   !
 Vetumboso (9) ---.___    !       !
   Mosina (10) ---'   !---'       !___
     Sasar (8) ---.___!           !   !
   Vatrata (7) ---'               !   !
      Mota (6) -----------.       !   !
     Nume (12) -------.   !---.   !   !
    Dorig (13) ---.___!___!   !   !   !
  Wetamut (15) ---'   !       !___!   !
     Koro (14) ---.___!       !       !
   Lakona (16) ---'           !       !
    Merig (17) -----------.___!       !
   Merlav (18) -----------'           !
   Marino (19) ---------------.       !
 Peterara (20) -----------.   !___    !-
Navenevene (21) -------.___!   !   !   !
      Tam (22) -------'   !---'   !   !
   Nasawa (23) ---.___    !       !   !
Narovorovo (24) ---'   !---'       !---!
  Baetora (25) -------'           !   !
 Wailengi (26) -------.___        !   !
Lolomatui (27) -------'   !---.   !   !
  Ngwatua (28) -------.___!   !___!   !
 Lolsiwoi (29) -------'       !       !
     Raga (30) ---------------'       !
     Apma (31) -------------------.   !
     Seke (32) ---------------.___!___!
       Sa (33) ---------------'   !
     Sowa (34) -------------------'
```

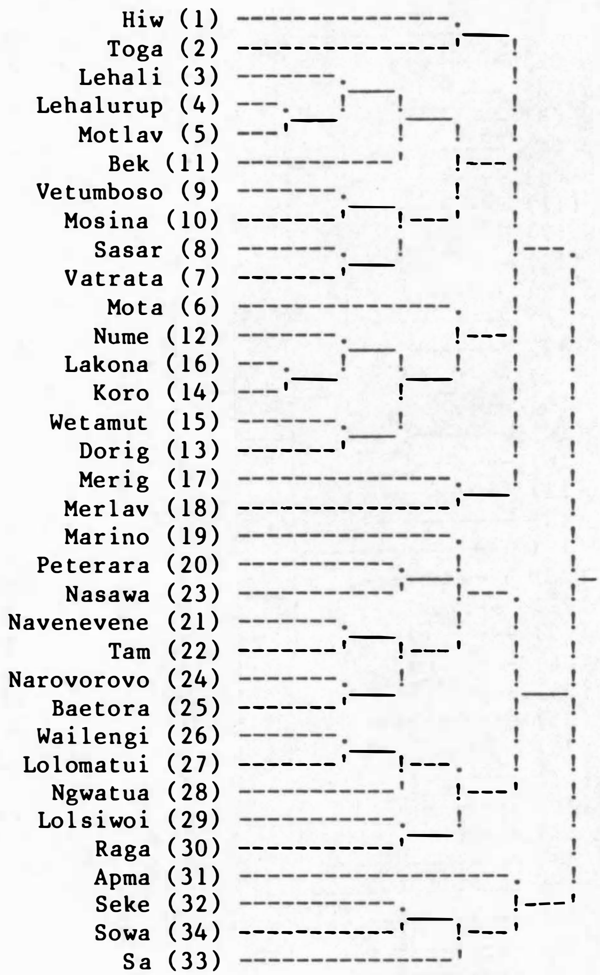
[27]

Phylogeny of 34 communalects of the Northeast New Hebrides reconstituted from $(\chi^2-b)/N$ measurements taken after removing near-empty rows and columns from tables of possible sound correspondences



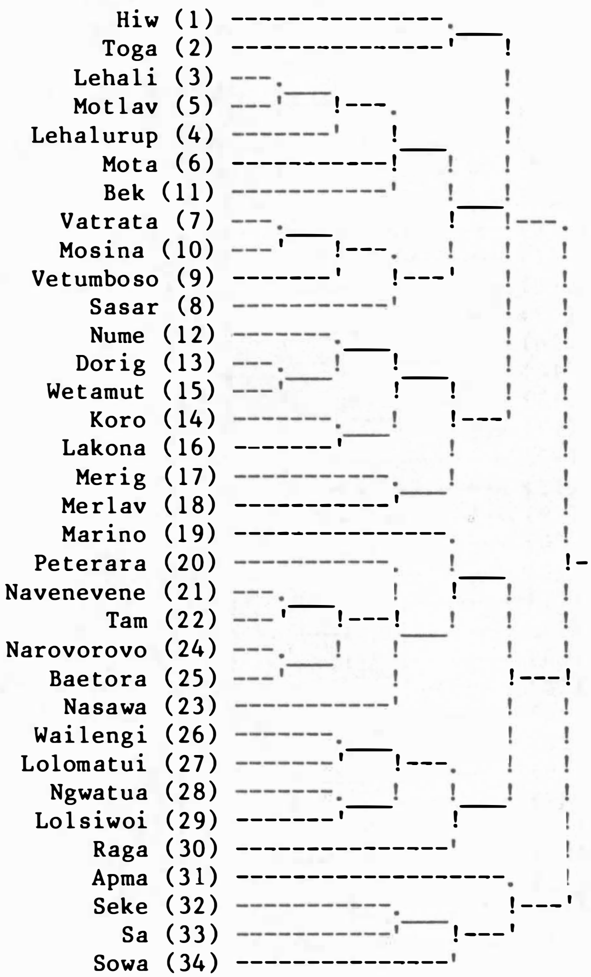
[28]

Phylogeny of 34 communalects of the Northeast New Hebrides reconstituted from $(\chi^2-b)/N$ measurements taken before removing near-empty rows and columns from tables of possible sound correspondences



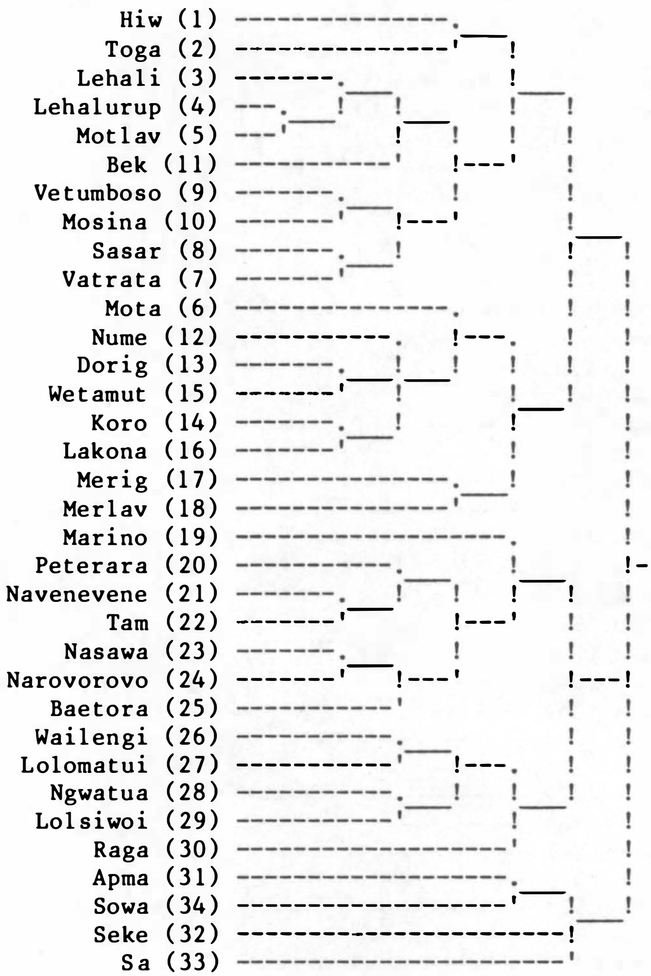
[29]

Phylogeny of 34 communalects of the Northeast New Hebrides reconstituted from χ^2-b measurements taken after removing near-empty rows and columns from tables of possible sound correspondences



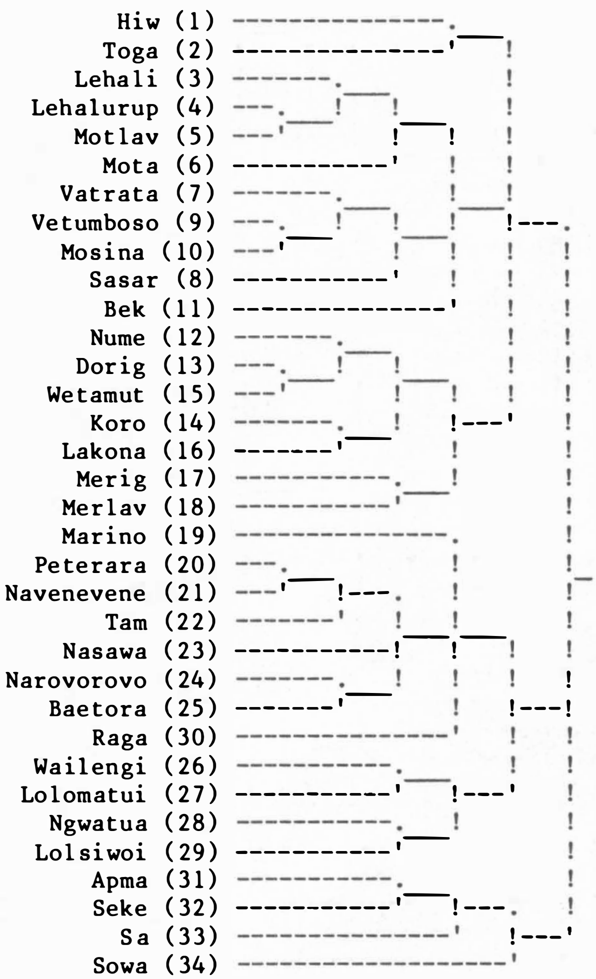
[30]

Phylogeny of 34 communalects of the Northeast New Hebrides reconstituted from χ^2-b measurements taken before removing near-empty rows and columns from tables of possible sound correspondences



[31]

Phylogeny of 34 communalects of the Northeast New Hebrides reconstituted from χ^2/N measurements taken after removing near-empty rows and columns from tables of possible sound correspondences



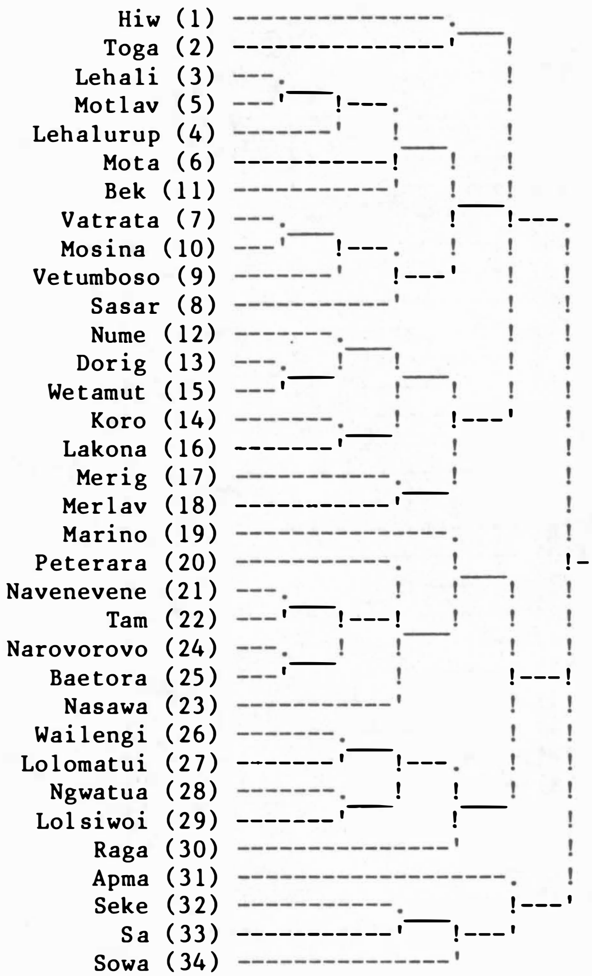
[32]

Phylogeny of 34 communalects of the Northeast New Hebrides reconstituted from χ^2/N measurements taken before removing near-empty rows and columns from tables of possible sound correspondences

```
        Hiw (1) ---------------.___
       Toga (2) ---------------'   !
     Lehali (3) -------.___        !___
  Lehalurup (4) ---.___!   !___    !   !
     Motlav (5) ---'       !   !   !   !
       Bek (11) -----------'   !---'   !
  Vetumboso (9) -------.___    !       !
    Mosina (10) -------'   !---'       !
      Sasar (8) -------.___!           !
    Vatrata (7) -------'               !---.
       Mota (6) -------------------.   !   !
      Nume (12) ---------------.   !---!   !
     Dorig (13) -----------.___!___!   !   !
   Wetamut (15) -----------'   !       !   !
      Koro (14) -----------.___!       !   !
    Lakona (16) -----------'           !   !
     Merig (17) -------------------.___!   !
    Merlav (18) -------------------'       !
    Marino (19) -------------------.       !
  Peterara (20) ---------------.   !       !
Navenevene (21) -------.___    !---!       !-
       Tam (22) -------'   !---'   !---.   !
Narovorovo (24) -------.___!       !   !   !
   Baetora (25) -------'           !   !___!
    Nasawa (23) -------------------'   !   !
  Wailengi (26) ---------------.___    !   !
 Lolomatui (27) ---------------'   !---'   !
   Ngwatua (28) ---------------.___!       !
  Lolsiwoi (29) -----------.___!           !
      Raga (30) -----------'               !
      Apma (31) -------------------.___    !
        Sa (33) -------------------'   !___!
      Seke (32) -----------------------!
      Sowa (34) -----------------------'
```

[33]

Phylogeny of 34 communalects of the Northeast New Hebrides reconstituted from χ^2 measurements taken after removing near-empty rows and columns from tables of possible sound correspondences



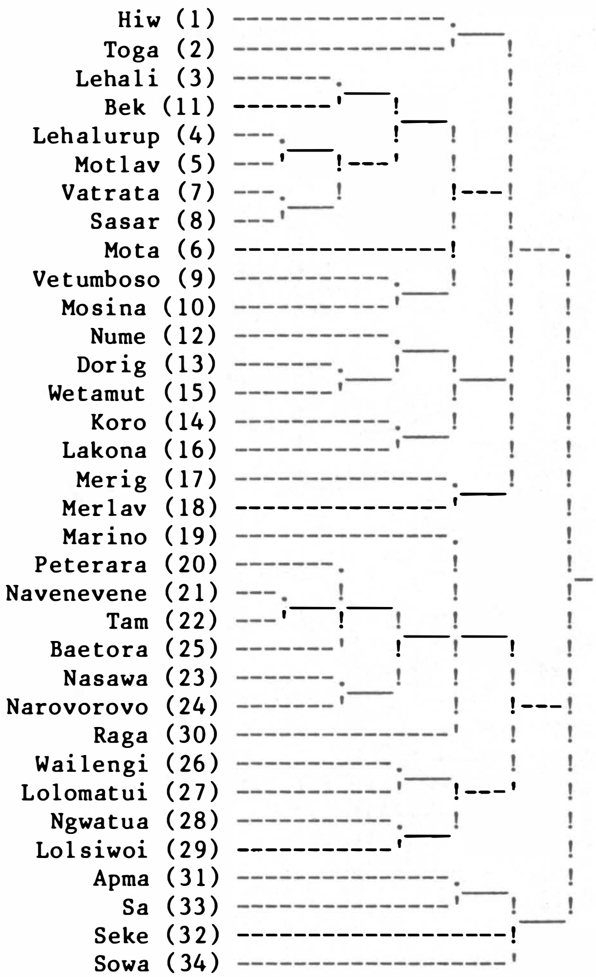
[34]

Phylogeny of 34 communalects of the Northeast New Hebrides reconstituted from χ^2 measurements taken before removing near-empty rows and columns from tables of possible sound correspondences

```
       Hiw (1)  ---------------.___
      Toga (2)  ---------------'   !
    Lehali (3)  -------.___        !
 Lehalurup (4)  -------'   !---.   !
    Motlav (5)  -------.___!   !   !
      Mota (6)  -------'       !   !
   Vatrata (7)  -------.___    !---!
     Sasar (8)  -------'   !---!   !---.
     Bek (34)   -----------'   !   !   !
 Vetumboso (9)  -----------.___!   !   !
   Mosina (10)  -----------'       !   !
    Dorig (12)  ---.___            !   !
  Wetamut (14)  ---'   !---.       !   !
     Nume (11)  -------'   !---.   !   !
   Lakona (15)  -------.___!   !___!   !
     Koro (13)  -------'       !       !
    Merig (16)  -----------.___!       !
   Merlav (17)  -----------'           !
   Marino (18)  ---------------.       !
Navenevene (20) ---.___        !___    !-
      Tam (21)  ---'   !---.   !   !   !
 Peterara (19)  -------'   !___!   !   !
Narovorovo (23) ---.___    !       !   !
  Baetora (24)  ---'   !---'       !---!
   Nasawa (22)  -------'           !   !
 Wailengi (25)  -------.___        !   !
Lolomatui (26)  -------'   !___    !   !
 Lolsiwoi (28)  -----------!   !___!   !
  Ngwatua (27)  -----------'   !       !
     Raga (29)  ---------------'       !
     Apma (30)  ---------------.___    !
       Sa (32)  ---------------'   !___!
     Sowa (33)  -------------------!
     Seke (31)  -------------------'
```

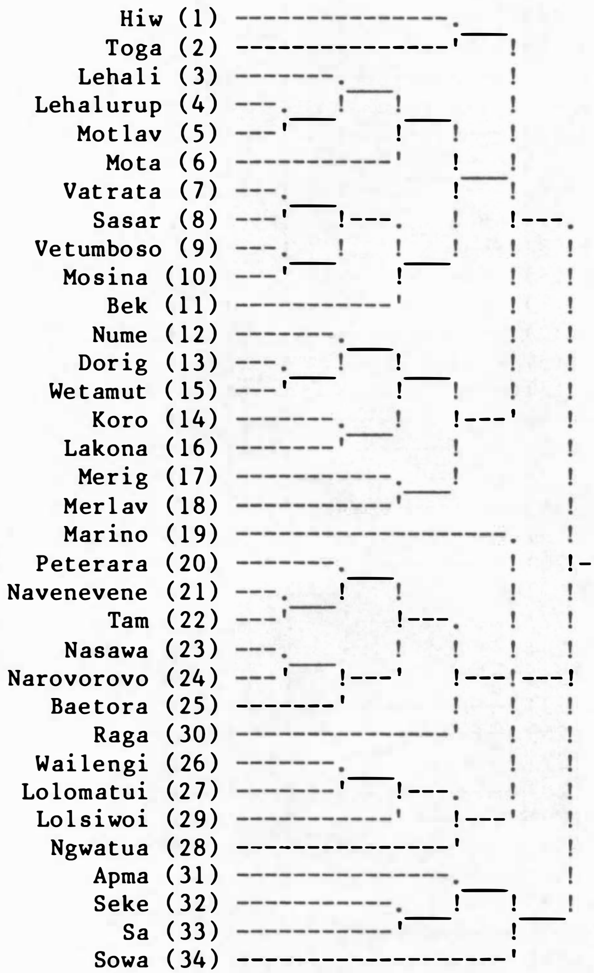
[35]

Phylogeny of 34 communalects of the Northeast New Hebrides
reconstructed from proportions of shared cognates



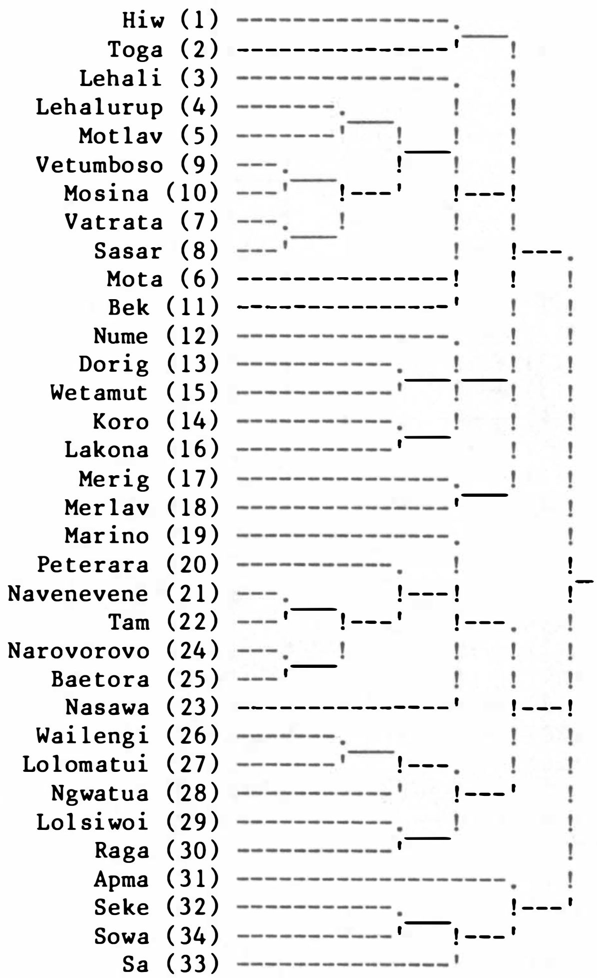
[36]

Phylogeny of 34 communalects of the Northeast New Hebrides reconstituted from $(\chi^2-b)/N$ measurements taken after removing near-empty rows and columns from tables of possible sound correspondences (affixes stripped)



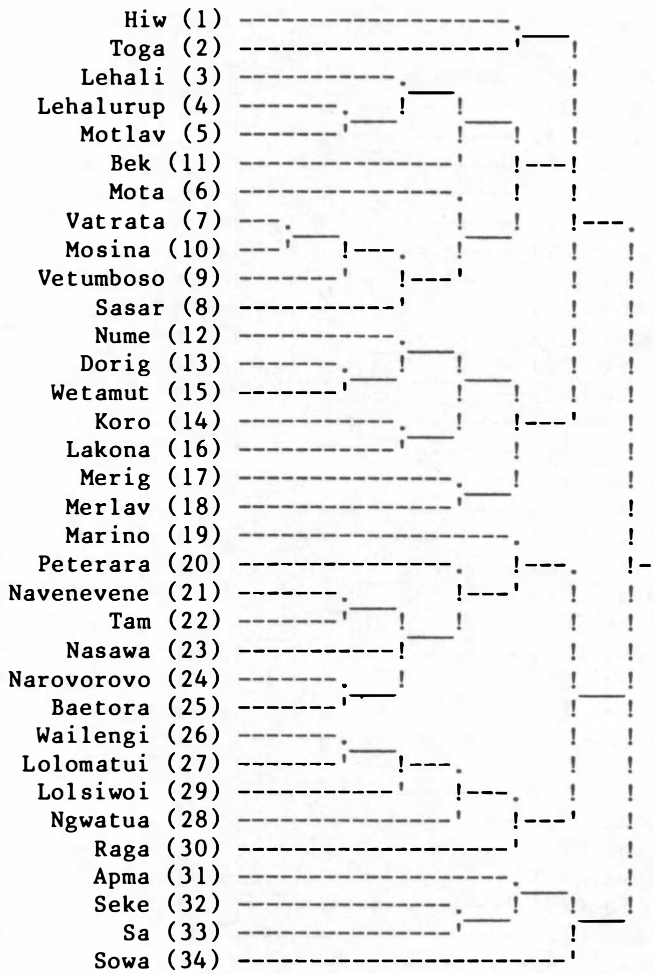
[37]

Phylogeny of 34 communalects of the Northeast New Hebrides reconstituted from $(\chi^2-b)/N$ measurements taken before removing near-empty rows and columns from tables of possible sound correspondences (affixes stripped)



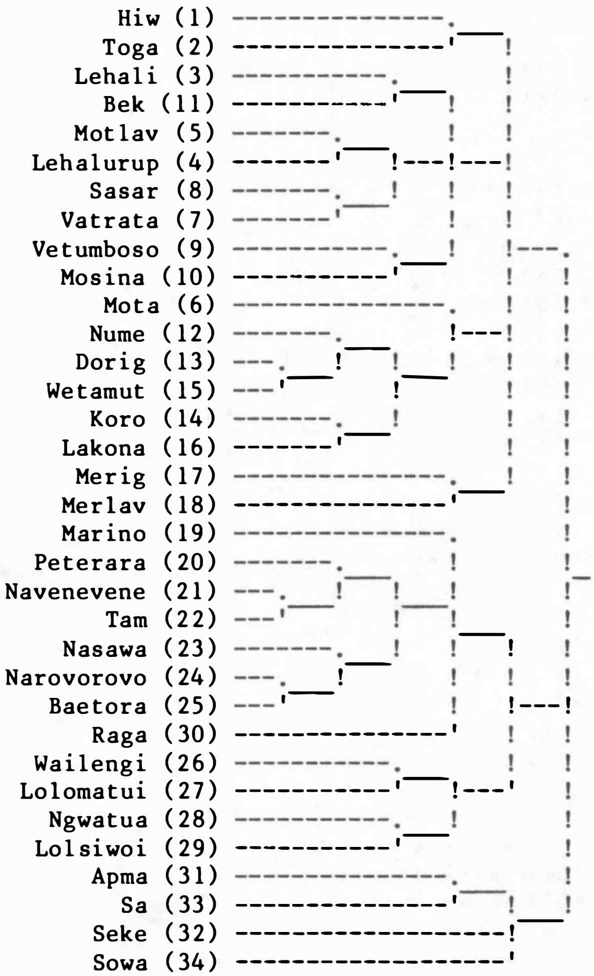
[38]

Phylogeny of 34 communalects of the Northeast New Hebrides reconstituted from χ^2-b measurements taken after removing near-empty rows and columns from tables of possible sound correspondences (affixes stripped)



[39]

Phylogeny of 34 communalects of the Northeast New Hebrides reconstituted from χ^2-b measurements taken before removing near-empty rows and columns from tables of possible sound correspondences (affixes stripped)



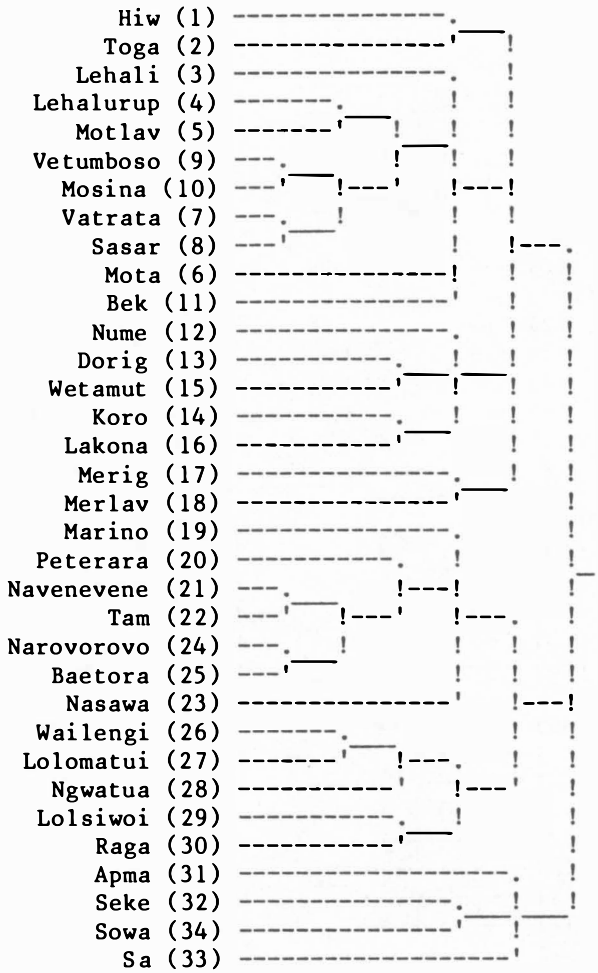
[40]

Phylogeny of 34 communalects of the Northeast New Hebrides reconstituted from χ^2/N measurements taken after removing near-empty rows and columns from tables of possible sound correspondences (affixes stripped)

```
       Hiw (1) ---------------.___
      Toga (2) ---------------'   !
    Lehali (3) -------.___        !
 Lehalurup (4) ---.___!   !___    !
    Motlav (5) ---'       !   !   !
      Mota (6) -----------'   !   !
   Vatrata (7) -------.___    !___!
     Sasar (8) -------'   !---!   !---.
 Vetumboso (9) -------.___!   !   !   !
   Mosina (10) -------'       !   !   !
      Bek (11) ---------------'   !   !
     Nume (12) -------.___        !   !
    Dorig (13) ---.___!   !___    !   !
  Wetamut (15) ---'       !   !   !   !
     Koro (14) -------.___!   !---'   !
   Lakona (16) -------'       !       !
    Merig (17) -----------.___!       !
   Merlav (18) -----------'           !
   Marino (19) -------------------.   !
 Peterara (20) ---.___            !   !_
Navenevene (21) ---'   !---.       !   !
      Tam (22) -------'   !___    !   !
   Nasawa (23) -----------!   !   !   !
Narovorovo (24) -------.___!   !---!---!
  Baetora (25) -------'       !   !   !
     Raga (30) ---------------'   !   !
 Wailengi (26) -------.___        !   !
Lolomatui (27) -------'   !---.   !   !
 Lolsiwoi (29) -----------'   !---'   !
  Ngwatua (28) ---------------'       !
     Apma (31) -----------.___        !
       Sa (33) -----------'   !---.   !
     Seke (32) ---------------'   !---'
     Sowa (34) -------------------'
```

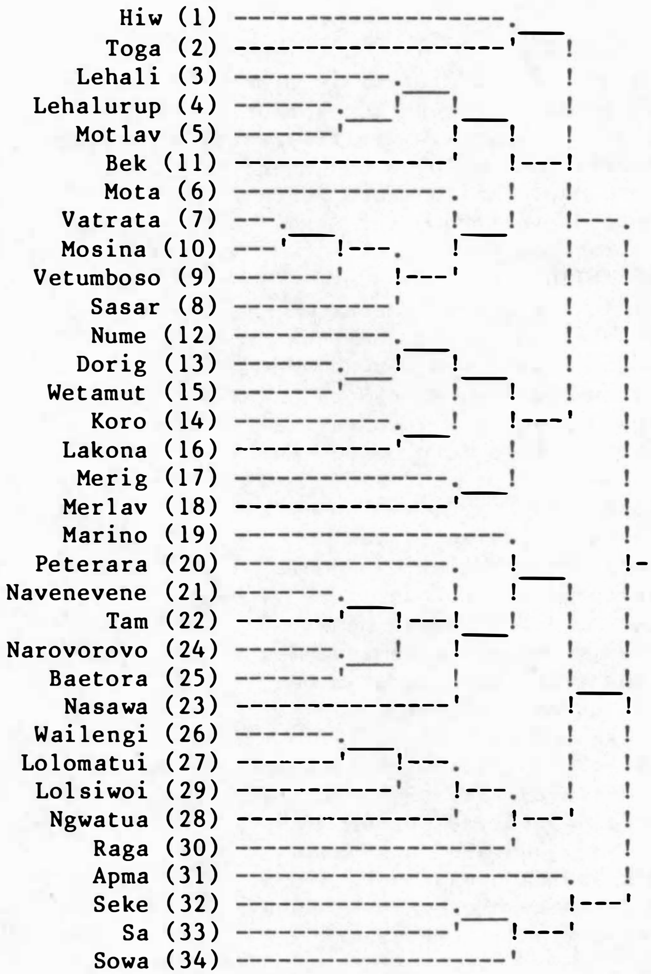
[41]

Phylogeny of 34 communalects of the Northeast New Hebrides reconstituted from χ^2/N measurements taken before removing near-empty rows and columns from tables of possible sound correspondences (affixes stripped)



[42]

Phylogeny of 34 communalects of the Northeast New Hebrides reconstituted from χ^2 measurements taken after removing near-empty rows and columns from tables of possible sound correspondences (affixes stripped)



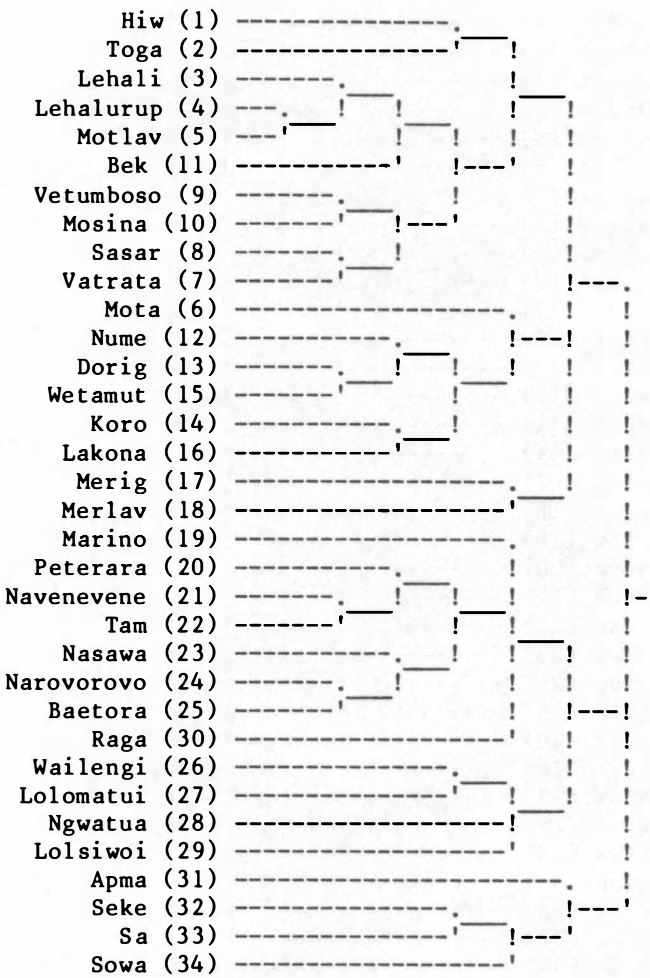
[43]

Phylogeny of 34 communalects of the Northeast New Hebrides reconstituted from χ^2 measurements taken before removing near-empty rows and columns from tables of possible sound correspondences (affixes stripped)

154:BLACK
miiix7 mududut Miut milie mlexlex
silsilixa korkor kurkur korkor kurkur
qu9qu9 wiriwirix wiwirix werweriu w7rwirix
wewerix polo polo osooso maeto
maeto maeto maito maeto maeto
maeto maeto meto maeat meto
me proh mamlek mee
155:WHITE
iui lulu lul qaqa qaxqax
aqaxa qaqaxa qaxqax qax qax
qauqax wewend wen wenwen vWet
awet wenwen wenwen si9ara oxWoxWo
ovovo ovovo vWas voas wasixi
mavute mevute lotu movuot maita
dap asav palpal adao
156:RED
mie mi m3 lawlaw lawlaw
memea memee lawlaw mame meme
laulau me mme me mme
meme meme meme memea memea
memea memea memea memea memea
memea memea kara memea memea
meme arvan 9i9i ame
157:YELLOW
9o9o e9o 9oi 9oi9oi 9oi9oi
a9oa9o a9a9 a9a9 a9a9 a9a9
9or9or a9a9or a9 t0rt0ro 9a9or
a9 a9a9 a9a9 a9oaxa a9oxa
a9oxa a9oxa a9oxa a9oxa a9oxa
a9oka a9oka o9oro a9oak a9oxa
sesede msoh litlit seset

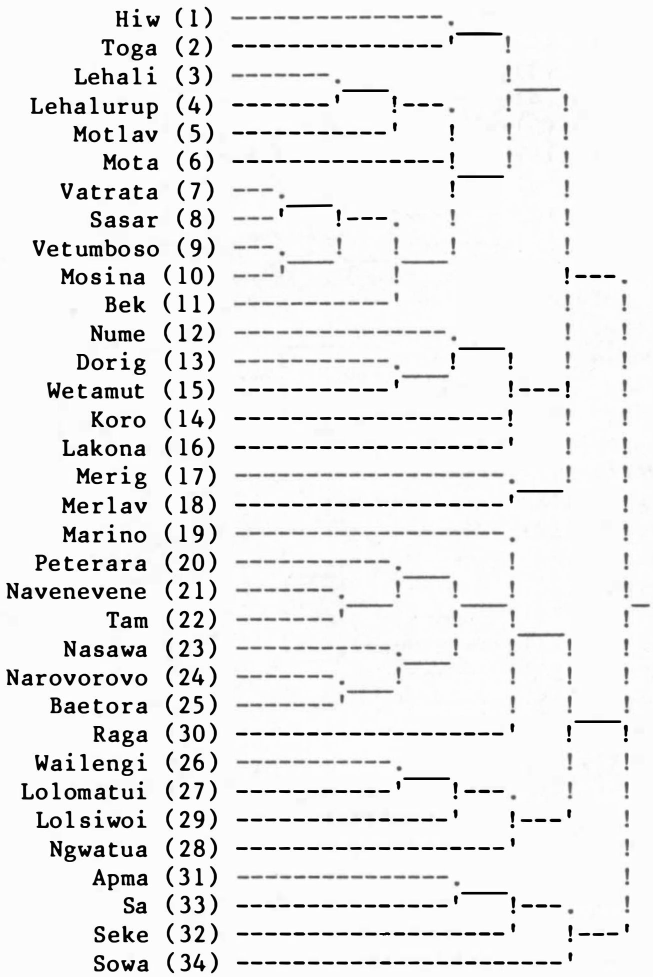
[44]

Four words in 34 languages and dialects of
the Northern New Hebrides, stripped of probable affixes,
and transliterated in a near-phonemic notation



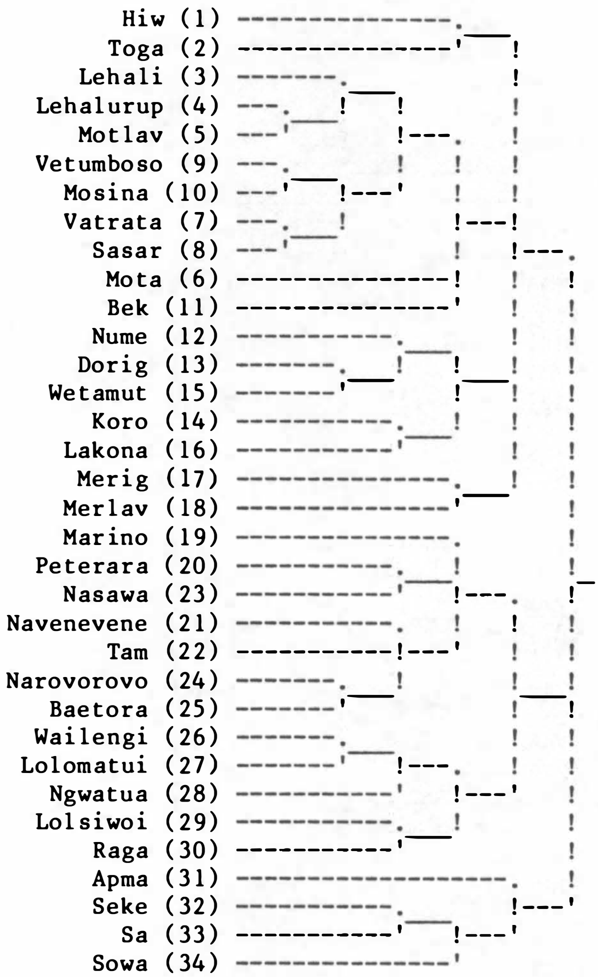
[45]

Phylogeny of 34 communalects of the Northeast New Hebrides reconstituted from $(\chi^2\text{-b})/N$ measurements taken after removing near-empty rows and columns from tables of possible sound correspondences (affixes stripped, near-phonemic, single-symbol transliteration)



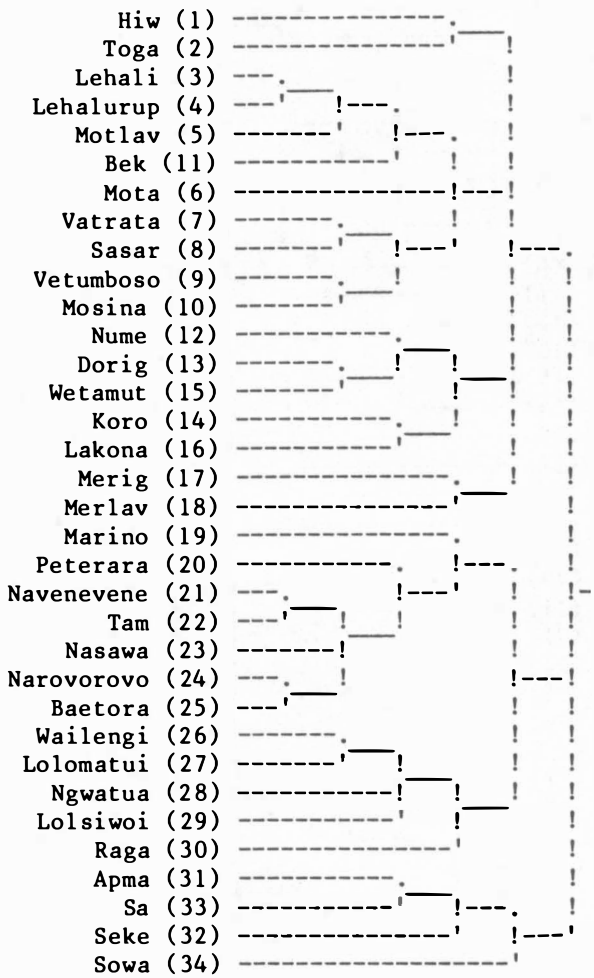
[46]

Phylogeny of 34 communalects of the Northeast New Hebrides reconstituted from $(\chi^2\text{-}b)/N$ measurements taken before removing near-empty rows and columns from tables of possible sound correspondences (affixes stripped, near-phonemic, single-symbol transliteration)



[47]

Phylogeny of 34 communalects of the Northeast New Hebrides reconstituted from χ^2-b measurements taken after removing near-empty rows and columns from tables of possible sound correspondences (affixes stripped, near-phonemic, single-symbol transliteration)



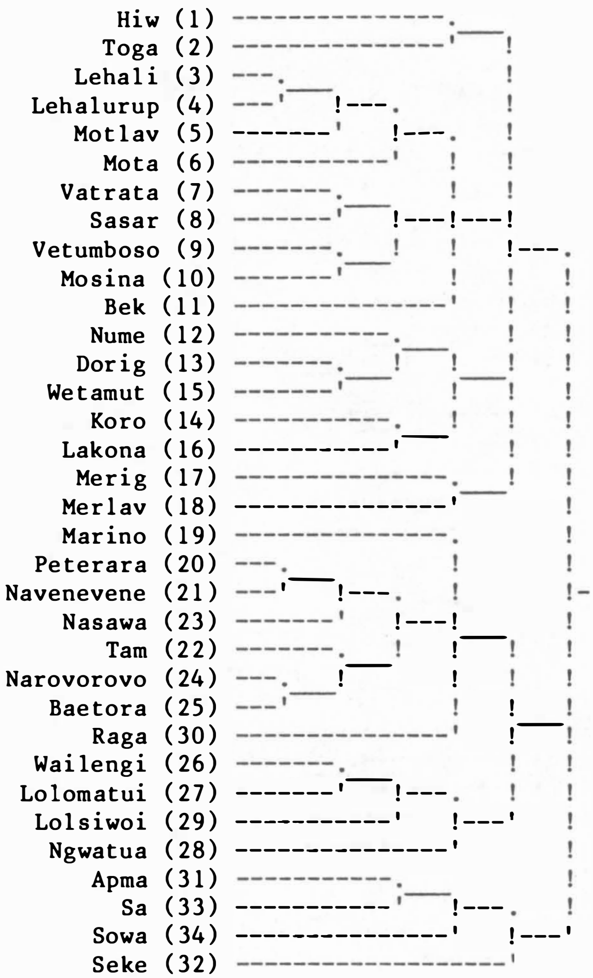
[48]

Phylogeny of 34 communalects of the Northeast New Hebrides reconstituted from χ^2-b measurements taken before removing near-empty rows and columns from tables of possible sound correspondences (affixes stripped, near-phonemic, single-symbol transliteration)

```
        Hiw (1) -----------.
       Toga (2) -----------'   !
     Lehali (3) ---.           !
  Lehalurup (4) ---'   !---.   !---.
     Motlav (5) -------'   !   !   !
  Vetumboso (9) ---.       !   !   !
    Mosina (10) ---'   !---!---'   !
    Vatrata (7) ---.   !   !       !
      Sasar (8) ---'       !       !
       Bek (11) -----------'       !
       Mota (6) ---------------.   !   !
      Nume (12) -------.       !---!   !
     Dorig (13) ---.   !   !   !   !   !
   Wetamut (15) ---'       !       !   !
      Koro (14) -------.   !       !   !
    Lakona (16) -------'           !   !
     Merig (17) ---------------.   !   !
    Merlav (18) ---------------'       !
    Marino (19) ---------------.       !
  Peterara (20) -------.       !       !
Navenevene (21) ---.   !   !   !       !-
       Tam (22) ---'       !---!       !
    Nasawa (23) -----------!   !   !   !
Narovorovo (24) -------.   !   !   !   !
   Baetora (25) -------'       !   !---!
      Raga (30) ---------------'   !   !
  Wailengi (26) -----------.       !   !
 Lolomatui (27) -----------'   !   !   !
   Ngwatua (28) ---------------!       !
  Lolsiwoi (29) ---------------'       !
      Apma (31) -------------------.   !
      Seke (32) -----------.       !---'
        Sa (33) -----------'   !---'
      Sowa (34) ---------------'
```

[49]

Phylogeny of 34 communalects of the Northeast New Hebrides reconstituted from χ^2/N measurements taken after removing near-empty rows and columns from tables of possible sound correspondences (affixes stripped, near-phonemic, single-symbol transliteration)



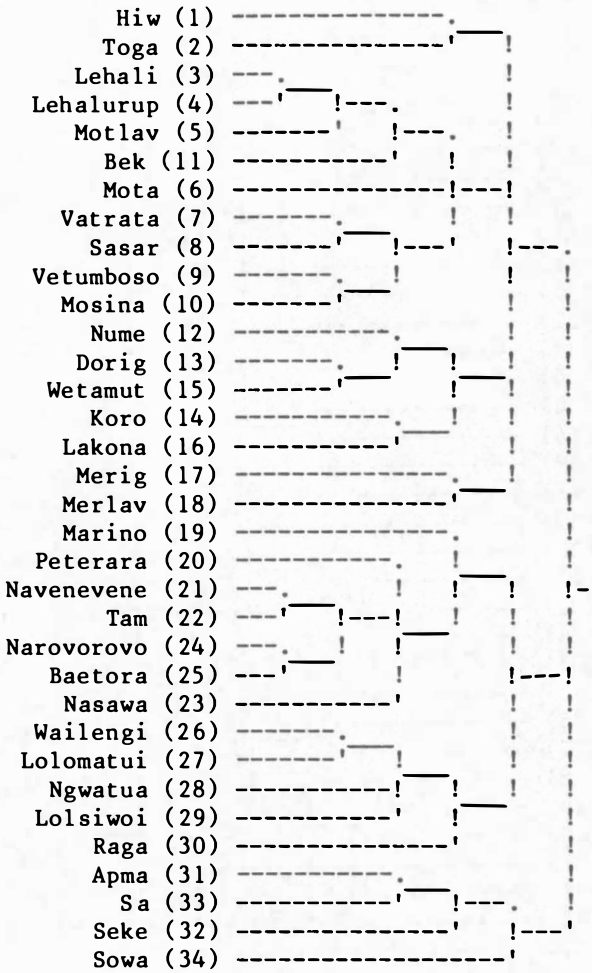
[50]

Phylogeny of 34 communalects of the Northeast New Hebrides
reconstituted from χ^2/N measurements taken after removing
near-empty rows and columns from tables of
possible sound correspondences
(affixes stripped, near-phonemic, single-symbol transliteration)

```
        Hiw (1) -------------------.___
       Toga (2) -------------------'   !
     Lehali (3) -----------.___        !
  Lehalurup (4) -------.___!   !       !
     Motlav (5) -------'       !---.   !
  Vetumboso (9) -------.___    !   !   !
    Mosina (10) -------'   !---'   !   !
    Vatrata (7) -------.___!       !---!
      Sasar (8) -------'           !   !---.
       Mota (6) -------------------!   !   !
       Bek (11) -------------------'   !   !
      Nume (12) ---------------.___    !   !
     Dorig (13) -----------.___!   !___!   !
   Wetamut (15) -----------'       !   !   !
      Koro (14) ---------------.___!   !   !
    Lakona (16) ---------------'       !   !
     Merig (17) -------------------.___!   !
    Merlav (18) -------------------'       !
    Marino (19) -------------------.       !
  Peterara (20) ---------------.___!       !
Navenevene (21) -----------.___!   !       !_
       Tam (22) -------.___!       !---.   !
Narovorovo (24) ---.___!           !   !   !
   Baetora (25) ---'               !   !___!
    Nasawa (23) -------------------'   !   !
  Wailengi (26) ---------------.___    !   !
 Lolomatui (27) ---------------'   !---'   !
   Ngwatua (28) ---------------.___!       !
  Lolsiwoi (29) -----------.___!           !
      Raga (30) -----------'               !
      Apma (31) -----------------------.   !
      Seke (32) ---------------.___    !---'
        Sa (33) ---------------'   !---'
      Sowa (34) -------------------'
```

[51]

Phylogeny of 34 communalects of the Northeast New Hebrides reconstituted from χ^2 measurements taken after removing near-empty rows and columns from tables of possible sound correspondences (affixes stripped, near-phonemic, single-symbol transliteration)



[52]

Phylogeny of 34 communalects of the Northeast New Hebrides
reconstituted from χ^2 measurements taken before removing
near-empty rows and columns from tables of
possible sound correspondences
(affixes stripped, near-phonemic, single-symbol transliteration)

	Hukua										
Valpei	895	Valpei									
Nokuku	798	778	Nokuku								
Tasmate	718	728	715	Tasmate							
Wusi1	515	458	503	649	Wusi1						
Wusi2	538	540	522	630	933	Wusi2					
Kerepua	536	527	503	649	884	877	Kerepua				
Vunapu	710	716	738	755	555	556	530	Vunapu			
Piamatsina	698	699	679	748	544	556	535	777	Piamatsina		
Tolomako	573	525	576	678	577	610	594	626	676	Tolomako	
Malmariv	533	470	523	610	638	683	633	557	581	631	Malmariv
Nonona	530	457	509	660	801	856	801	538	549	599	Nonona
Navut	529	454	497	620	659	680	620	545	563	575	Navut
Lametin	490	424	468	596	651	673	629	491	527	600	Lametin
Morouas	463	404	457	521	482	519	497	472	489	514	Morouas
Batunlamak	458	461	470	523	511	513	511	475	492	543	Batunlamak
Matae	549	465	524	648	684	711	674	559	564	587	Matae
Akei	491	409	482	578	566	581	570	478	511	484	Akei
Fortsenal	497	446	500	549	543	590	566	489	517	536	Fortsenal
Penantsiro	491	419	494	559	542	576	551	492	531	507	Penantsiro
Tasiriki	518	491	503	572	570	579	576	506	541	536	Tasiriki
Wailapa	521	500	533	573	567	584	555	505	533	539	Wailapa
Roria	475	409	457	500	478	516	491	449	460	458	Roria
Amblong	428	424	428	500	470	484	456	433	456	485	Amblong
Narango	476	469	461	523	523	516	511	475	486	520	Narango
Polonombau	373	360	358	382	391	391	404	351	363	408	Polonombau
Butmas	395	379	384	433	419	431	427	382	388	440	Butmas
Tur	381	370	364	403	392	409	415	366	360	418	Tur
Nambel	468	397	437	518	477	503	518	462	474	478	Nambel
Tambotalo	472	388	436	493	434	475	477	446	458	448	Tambotalo
Sakao	359	322	341	384	340	374	355	364	369	408	Sakao
Lorediakar	370	333	340	362	356	380	384	339	327	365	Lorediakar
SB1	398	341	370	379	374	417	406	371	354	392	SB1
SB2	411	340	376	390	357	408	395	373	356	384	SB2
Mafea	528	454	494	555	473	494	506	520	531	512	Mafea
Tutuba	563	466	544	587	438	468	473	541	547	483	Tutuba
Aore	569	495	548	579	483	510	512	552	558	500	Aore
Malo-North	503	436	515	544	431	475	466	508	520	451	Malo-North
Malo-South	465	453	458	522	500	514	521	493	534	500	Malo-South
Tangoa	500	439	488	523	476	503	500	481	514	507	Tangoa
Araki	558	538	549	570	558	554	552	555	579	534	Araki
	Huk	Val	Nok	Tas	Wus	Wus	Ker	Vun	Pia	Tol	

[53a]

Proportions of shared cognates x1000 for 41 communalects
of Espiritu Santo
(according to Tryon 1976)

(continued next page)

	Malmariv										
Nonona	679	Nonona									
Navut	776	718	Navut								
Lametin	749	694	738	Lametin							
Morouas	591	533	567	562	Morouas						
Batunlamak	601	525	582	582	839	Batunlamak					
Matae	756	735	818	699	565	580	Matae				
Akei	643	586	707	579	577	574	675	Akei			
Fortsenal	681	584	637	620	724	678	637	674	Fortsenal		
Penantsiro	648	576	665	589	633	613	655	778	739	Penantsiro	
Tasiriki	679	602	690	585	642	629	690	834	738	898	Tasiriki
Wailapa	657	578	647	613	695	652	656	737	733	802	Wailapa
Roria	562	526	509	603	571	552	502	454	561	498	Roria
Amblong	538	479	513	519	763	802	506	533	625	578	Amblong
Narango	580	497	537	540	750	740	537	537	628	609	Narango
Polonombau	460	388	402	394	453	444	409	406	432	428	Polonombau
Butmas	456	411	426	432	488	503	420	417	470	448	Butmas
Tur	453	390	426	436	474	469	408	424	465	444	Tur
Nambel	528	487	519	535	762	726	513	511	622	586	Nambel
Tambotalo	484	446	488	468	577	598	476	491	522	507	Tambotalo
Sakao	379	351	344	371	391	407	352	335	389	359	Sakao
Lorediakar	367	355	363	379	413	421	372	354	373	376	Lorediakar
SB1	380	377	382	391	446	457	385	378	398	414	SB1
SB2	381	360	382	380	438	469	374	373	393	415	SB2
Mafea	509	472	491	484	568	609	500	529	556	552	Mafea
Tutuba	498	455	507	476	553	581	500	502	521	537	Tutuba
Aore	518	505	502	497	574	602	522	517	537	534	Aore
Malo-North	459	439	464	455	491	561	464	469	454	509	Malo-North
Malo-South	549	477	523	523	593	611	535	558	529	581	Malo-South
Tangoa	549	480	534	500	636	669	534	559	623	639	Tangoa
Araki	624	570	612	566	632	646	627	646	671	727	Araki
	Mal	Non	Nav	Lam	Mor	Bat	Mat	Ake	For	Pen	

[53b]

Proportions of shared cognates x1000 for 41 communalects
of Espiritu Santo
(according to Tryon 1976)

(continued next page)

	Tasiriki										
Wailapa	802	Wailapa									
Roria	518	548	Roria								
Amblong	580	603	535	Amblong							
Narango	606	650	538	684	Narango						
Polonombau	406	428	409	417	471	Polonombau					
Butmas	441	424	414	470	527	754	Butmas				
Tur	430	418	401	429	497	702	849	Tur			
Nambel	588	610	504	693	828	555	610	578	Nambel		
Tambotalo	572	582	439	587	618	438	485	465	607	Tambotalo	
Sakao	377	368	374	379	451	417	497	449	412	376	Sakao
Lorediakar	369	392	372	409	472	497	556	562	459	396	Lorediakar
SB1	412	414	362	430	509	491	590	627	479	424	SB1
SB2	429	424	355	458	532	515	629	647	482	419	SB2
Mafea	590	602	455	572	636	452	497	464	587	629	Mafea
Tutuba	571	574	431	538	587	388	424	404	531	574	Tutuba
Aore	584	607	454	565	619	420	464	442	576	641	Aore
Malo-North	563	594	390	521	592	382	390	369	495	533	Malo-North
Malo-South	634	639	465	606	606	425	431	417	570	600	Malo-South
Tangoa	661	724	484	624	737	436	444	425	604	570	Tangoa
Araki	741	780	511	577	648	420	441	417	594	612	Araki
	Tas	Wai	Ror	Amb	Nar	Pol	But	Tur	Nam	Tam	

	Sakao										
Lorediakar	475	Lorediakar									
SB1	442	805	SB1								
SB2	414	770	874	SB2							
Mafea	414	453	471	471	Mafea						
Tutuba	353	364	393	392	733	Tutuba					
Aore	389	405	433	421	726	768	Aore				
Malo-North	358	353	380	372	599	653	704	Malo-North			
Malo-South	419	381	417	424	671	714	768	965	Malo-South		
Tangoa	392	422	440	441	645	607	662	670	720	Tangoa	
Araki	396	380	414	427	669	710	728	706	724	791	Araki
	Sak	Lor	SB1	SB2	Maf	Tut	Aor	Mal	Mal	Tan	

[53c]

Proportions of shared cognates x1000 for 41 communalects
of Espiritu Santo
(according to Tryon 1976)

(concluded)

	Hukua										
Valpei	956	Valpei									
Nokuku	977	957	Nokuku								
Tasmate	865	813	894	Tasmate							
Wusi1	314	352	343	600	Wusi1						
Wusi2	318	260	338	626	985	Wusi2					
Kerepua	329	293	369	609	986	990	Kerepua				
Vunapu	937	909	956	946	422	427	461	Vunapu			
Piamatsina	895	864	944	957	471	463	491	979	Piamatsina		
Tolomako	660	709	709	887	758	712	719	804	829	Tolomako	
Malmariv	249	261	306	570	787	752	752	380	465	732	Malmariv
Nonona	310	337	353	615	973	964	956	452	494	773	Nonona
Navut	251	248	323	566	784	778	784	381	461	731	Navut
Lametin	187	195	250	508	815	811	801	350	408	701	Lametin
Morouas	-163	-107	-96	-18	146	87	111	-93	0	149	Morouas
Batunlamak	-125	-163	-90	-4	89	81	79	-73	20	103	Batunlamak
Matae	294	313	358	604	837	822	813	432	516	769	Matae
Akei	149	174	215	379	564	542	537	264	351	551	Akei
Fortsenal	55	64	113	293	479	426	431	157	256	455	Fortsenal
Penantsiro	76	98	136	278	439	395	409	161	252	415	Penantsiro
Tasiriki	129	69	199	335	427	424	415	215	312	400	Tasiriki
Wailapa	141	85	186	315	381	363	388	223	326	393	Wailapa
Roria	119	212	201	387	597	536	543	260	346	645	Roria
Amblong	-153	-170	-104	-68	34	8	50	-96	-16	75	Amblong
Narango	-221	-249	-174	-164	-84	-84	-57	-184	-104	-62	Narango
Polonombau	-573	-494	-585	-654	-383	-403	-398	-607	-637	-537	Polonombau
Butmas	-602	-513	-628	-738	-453	-489	-462	-650	-694	-613	Butmas
Tur	-592	-524	-616	-719	-427	-459	-452	-652	-685	-605	Tur
Nambel	-358	-238	-285	-279	-53	-126	-122	-301	-235	-66	Nambel
Tambotalo	-99	-10	-13	-23	20	-71	-25	-18	59	71	Tambotalo
Sakao	-604	-416	-575	-732	-463	-573	-511	-593	-613	-555	Sakao
Lorediakar	-552	-483	-580	-737	-478	-493	-496	-621	-691	-664	Lorediakar
SB1	-539	-444	-563	-702	-458	-498	-481	-605	-666	-649	SB1
SB2	-544	-414	-553	-712	-464	-522	-498	-598	-657	-634	SB2
Mafea	66	114	150	76	-53	-109	-66	118	182	65	Mafea
Tutuba	295	345	361	308	86	29	74	368	430	300	Tutuba
Aore	224	251	287	251	45	-5	45	285	347	219	Aore
Malo-North	201	249	240	214	77	7	63	261	328	228	Malo-North
Malo-South	86	21	145	138	27	10	32	126	180	92	Malo-South
Tangoa	9	51	86	93	83	28	59	88	167	143	Tangoa
Araki	234	168	299	369	248	253	270	293	386	360	Araki
	Huk	Val	Nok	Tas	Wus	Wus	Ker	Vun	Pia	Tol	

[54a]

Linear-correlation measures x1000 for 41 communalects
of Espiritu Santo
(calculated from proportions of shared cognates)

(continued next page)

	Malmariv										
Nonona	865	Nonona									
Navut	973	861	Navut								
Lametin	960	892	940	Lametin							
Morouas	435	170	392	421	Morouas						
Batunlamak	380	150	341	361	969	Batunlamak					
Matae	968	905	987	948	353	302	Matae				
Akei	839	645	835	766	582	533	832	Akei			
Fortsenal	778	549	768	729	782	749	737	898	Fortsenal		
Penantsiro	743	506	741	656	672	633	719	963	920	Penantsiro	
Tasiriki	719	498	739	666	633	610	709	955	895	979	Tasiriki
Wailapa	693	465	697	611	706	720	667	896	916	944	Wailapa
Roria	820	640	775	797	718	682	767	759	862	721	Roria
Amblong	321	87	290	317	962	984	256	481	703	581	Amblong
Narango	158	-19	158	182	878	905	121	393	607	499	Narango
Polonombau	-394	-415	-372	-315	29	3	-425	-282	-177	-220	Polonombau
Butmas	-474	-490	-483	-441	-50	-111	-518	-384	-311	-330	Butmas
Tur	-461	-463	-459	-433	-72	-102	-486	-372	-308	-310	Tur
Nambel	185	-44	118	141	851	844	84	351	560	432	Nambel
Tambotalo	223	25	186	183	778	801	183	465	586	573	Tambotalo
Sakao	-512	-538	-534	-463	45	-11	-564	-411	-317	-316	Sakao
Lorediakar	-600	-531	-566	-538	-183	-196	-596	-482	-445	-411	Lorediakar
SB1	-573	-519	-541	-523	-177	-193	-565	-443	-420	-386	SB1
SB2	-572	-527	-547	-523	-121	-171	-561	-418	-386	-358	SB2
Mafea	105	-48	112	57	581	603	102	347	394	439	Mafea
Tutuba	224	90	203	156	463	517	227	412	409	446	Tutuba
Aore	159	25	169	107	457	517	163	356	363	422	Aore
Malo-North	189	54	166	117	438	451	188	371	359	410	Malo-North
Malo-South	164	65	192	137	459	553	176	377	409	450	Malo-South
Tangoa	334	113	312	287	760	773	302	597	655	678	Tangoa
Araki	497	298	508	434	623	647	491	726	707	755	Araki
	Mal	Non	Nav	Lam	Mor	Bat	Mat	Ake	For	Pen	

[54b]

Linear-correlation measures x1000 for 41 communalects
of Espiritu Santo
(calculated from proportions of shared cognates)

(continued next page)

	Tasiriki										
Wailapa	960	Wailapa									
Roria	660	712	Roria								
Amblong	576	687	622	Amblong							
Narango	494	605	485	933	Narango						
Polonombau	-278	-318	-256	51	143	Polonombau					
Butmas	-415	-412	-345	-49	39	968	Butmas				
Tur	-383	-402	-342	-31	43	957	976	Tur			
Nambel	374	476	528	857	866	300	214	207	Nambel		
Tambotalo	520	640	473	810	834	59	-50	-69	762	Tambotalo	
Sakao	-440	-379	-284	93	110	787	711	777	377	135	Sakao
Lorediakar	-474	-500	-506	-136	-31	617	699	751	75	-143	Lorediakar
SB1	-452	-457	-472	-113	-39	639	672	691	96	-125	SB1
SB2	-453	-447	-441	-116	-40	659	674	722	155	-68	SB2
Mafea	451	560	271	643	656	-115	-204	-193	536	904	Mafea
Tutuba	491	600	331	540	526	-284	-368	-387	388	822	Tutuba
Aore	450	554	274	544	545	-260	-358	-375	379	811	Aore
Malo-North	436	526	284	491	463	-260	-328	-343	352	750	Malo-North
Malo-South	484	593	240	536	594	-250	-315	-338	375	759	Malo-South
Tangoa	685	789	512	780	759	-162	-238	-248	663	906	Tangoa
Araki	807	878	538	655	610	-371	-490	-479	414	756	Araki
	Tas	Wai	Ror	Amb	Nar	Pol	But	Tur	Nam	Tam	

	Sakao										
Lorediakar	671	Lorediakar									
SB1	711	982	SB1								
SB2	802	971	985	SB2							
Mafea	1	-225	-183	-149	Mafea						
Tutuba	-141	-384	-350	-318	925	Tutuba					
Aore	-151	-371	-338	-293	942	980	Aore				
Malo-North	-70	-322	-289	-251	855	890	913	Malo-North			
Malo-South	-223	-326	-312	-306	819	831	863	933	Malo-South		
Tangoa	-39	-286	-244	-206	850	827	815	791	816	Tangoa	
Araki	-401	-521	-487	-480	787	810	800	763	813	892	Araki
	Sak	Lor	SB1	SB2	Maf	Tut	Aor	Mal	Mal	Tan	

[54c]

Linear-correlation measures x1000 for 41 communalects
of Espiritu Santo
(calculated from proportions of shared cognates)

(concluded)

	Hukua										
Valpei	340	Valpei									
Nokuku	209	229	Nokuku								
Tasmate	177	170	151	Tasmate							
Wusi1	78	87	73	128	Wusi1						
Wusi2	108	111	92	147	335	Wusi2					
Kerepua	74	87	71	140	281	280	Kerepua				
Vunapu	144	168	144	188	76	82	58	Vunapu			
Piamatsina	150	133	126	162	78	80	57	170	Piamatsina		
Tolomako	60	69	61	93	95	79	97	75	105	Tolomako	
Malmariv	99	75	89	124	150	169	121	77	82	95	Malmariv
Nonona	61	59	55	136	245	262	222	62	69	101	Nonona
Navut	83	74	59	113	155	145	127	69	85	66	Navut
Lametin	77	72	62	101	124	115	112	77	87	83	Lametin
Morouas	68	65	55	80	70	72	75	49	62	63	Morouas
Batunlamak	80	68	64	81	62	76	47	35	50	34	Batunlamak
Matae	98	70	84	154	167	205	178	94	80	93	Matae
Akei	90	58	50	89	96	96	101	50	66	54	Akei
Fortsenal	78	75	68	95	89	101	91	57	60	61	Fortsenal
Penantsiro	66	60	57	82	82	98	100	50	57	43	Penantsiro
Tasiriki	117	109	90	122	115	137	105	74	81	48	Tasiriki
Wailapa	117	99	90	112	127	121	101	78	63	62	Wailapa
Roria	72	55	60	72	60	82	65	39	48	41	Roria
Amblong	53	49	61	55	61	58	61	34	56	37	Amblong
Narango	83	76	63	78	76	86	57	55	38	36	Narango
Polonombau	63	45	40	33	33	32	17	23	25	14	Polonombau
Butmas	61	42	53	56	49	37	23	35	32	24	Butmas
Tur	38	36	34	29	27	28	23	20	12	15	Tur
Nambel	71	76	64	93	76	70	84	60	50	39	Nambel
Tambotalo	78	64	57	82	50	67	49	42	45	51	Tambotalo
Sakao	16	19	13	10	11	8	5	5	9	19	Sakao
Lorediakar	49	54	25	34	40	31	6	20	23	19	Lorediakar
SB1	41	33	37	27	29	14	20	12	20	26	SB1
SB2	68	46	55	46	36	42	33	27	25	26	SB2
Mafea	110	92	80	112	83	65	101	95	77	62	Mafea
Tutuba	91	80	89	104	58	50	75	98	82	53	Tutuba
Aore	93	77	88	93	55	68	59	80	79	56	Aore
Malo-North	119	84	109	92	45	78	74	86	111	55	Malo-North
Malo-South	51	56	49	72	51	24	59	35	109	61	Malo-South
Tangoa	79	81	70	79	77	75	56	53	82	64	Tangoa
Araki	112	124	80	82	90	89	76	73	79	60	Araki
	Huk	Val	Nok	Tas	Wus	Wus	Ker	Vun	Pia	Tol	

[55a]

$(\chi^2\text{-b})/N$ measurements x1000 for 41 communalects of Espiritu Santo
(taken after the removal of near-empty rows and columns
from tables of possible sound correspondences calculated
on unphonemicized lists)

(continued next page)

	Malmariv										
Nonona	173	Nonona									
Navut	310	191	Navut								
Lametin	285	136	236	Lametin							
Morouas	115	103	113	96	Morouas						
Batunlamak	75	57	61	79	156	Batunlamak					
Matae	269	218	353	222	112	68	Matae				
Akei	169	121	208	143	95	59	201	Akei			
Fortsenal	188	104	153	144	191	102	163	175	Fortsenal		
Penantsiro	143	110	174	132	119	65	165	239	196	Penantsiro	
Tasiriki	137	135	147	98	103	99	186	252	184	257	Tasiriki
Wailapa	141	115	157	145	159	99	158	181	167	198	Wailapa
Roria	103	73	86	109	92	71	72	51	115	52	Roria
Amblong	89	64	81	81	174	174	53	70	129	94	Amblong
Narango	80	53	62	72	110	120	58	56	95	80	Narango
Polonombau	54	37	40	45	33	40	25	29	42	37	Polonombau
Butmas	50	33	27	51	48	67	48	39	51	39	Butmas
Tur	48	20	38	46	39	44	37	36	58	45	Tur
Nambel	109	100	105	115	202	129	92	95	149	100	Nambel
Tambotalo	59	46	64	55	78	77	57	72	84	69	Tambotalo
Sakao	21	26	18	27	19	7	13	19	27	22	Sakao
Lorediakar	37	26	25	40	32	27	14	18	27	24	Lorediakar
SB1	39	41	46	46	44	40	36	38	42	39	SB1
SB2	42	34	44	45	37	63	31	38	35	35	SB2
Mafea	105	73	83	85	92	89	87	71	97	78	Mafea
Tutuba	66	47	57	78	90	95	60	58	68	70	Tutuba
Aore	69	63	54	58	92	78	68	70	75	71	Aore
Malo-North	53	51	54	62	82	98	61	72	71	81	Malo-North
Malo-South	56	24	45	12	73	110	36	50	54	62	Malo-South
Tangoa	99	78	90	103	146	113	98	98	130	130	Tangoa
Araki	98	96	93	83	103	89	119	125	112	130	Araki
	Mal	Non	Nav	Lam	Mor	Bat	Mat	Ake	For	Pen	

[55b]

$(\chi^2\text{-b})/N$ measurements x1000 for 41 communalects of Espiritu Santo
(taken after the removal of near-empty rows and columns
from tables of possible sound correspondences calculated
on unphonemicized lists)

(continued next page)

	Tas	Wai	Ror	Amb	Nar	Pol	But	Tur	Nam	Tam	
	Tasiriki										
Wailapa	220	Wailapa									
Roria	54	85	Roria								
Amblong	81	102	67	Amblong							
Narango	108	119	57	116	Narango						
Polonombau	43	60	51	43	56	Polonombau					
Butmas	37	52	70	45	60	174	Butmas				
Tur	40	41	52	41	53	142	233	Tur			
Nambel	93	140	95	178	198	73	107	103	Nambel		
Tambotalo	68	96	72	89	85	78	61	47	97	Tambotalo	
Sakao	10	20	31	14	16	25	29	19	27	9	Sakao
Lorediakar	24	44	36	43	39	62	91	88	50	44	Lorediakar
SB1	28	45	54	51	38	81	101	141	71	35	SB1
SB2	47	61	54	60	69	87	146	167	73	34	SB2
Mafea	77	116	68	116	90	54	55	44	113	92	Mafea
Tutuba	74	92	60	92	98	27	55	47	76	102	Tutuba
Aore	74	109	51	90	70	54	56	49	89	119	Aore
Malo-North	103	162	53	92	108	39	45	33	94	71	Malo-North
Malo-South	98	79	39	70	70	19	16	25	30	62	Malo-South
Tangoa	133	183	94	127	136	30	36	32	139	106	Tangoa
Araki	176	180	68	87	114	66	49	43	133	116	Araki

	Sak	Lor	SB1	SB2	Maf	Tut	Aor	Mal	Mal	Tan	
	Sakao										
Lorediakar	20	Lorediakar									
SB1	23	276	SB1								
SB2	29	237	395	SB2							
Mafea	19	45	63	58	Mafea						
Tutuba	21	24	53	55	178	Tutuba					
Aore	20	34	49	45	199	217	Aore				
Malo-North	16	31	39	41	114	145	162	Malo-North			
Malo-South	0	28	26	44	80	108	115	270	Malo-South		
Tangoa	27	54	61	62	114	115	142	194	76	Tangoa	
Araki	29	19	17	36	107	126	121	192	127	41	Araki

[55c]

$(\chi^2-b)/N$ measurements x1000 for 41 communalects of Espiritu Santo
(taken after the removal of near-empty rows and columns
from tables of possible sound correspondences calculated
on unphonemicized lists)

(concluded)

	Hukua										
Valpei	942	Valpei									
Nokuku	957	926	Nokuku								
Tasmate	682	705	780	Tasmate							
Wusi1	132	143	156	555	Wusi1						
Wusi2	165	179	213	587	968	Wusi2					
Kerepua	165	150	195	558	961	956	Kerepua				
Vunapu	836	770	881	873	252	312	341	Vunapu			
Piamatsina	686	748	807	819	222	272	330	890	Piamatsina		
Tolomako	397	361	457	800	688	751	716	638	631	Tolomako	
Malmariv	43	64	35	417	597	574	637	210	218	560	Malmariv
Nonona	76	89	100	492	925	914	929	233	205	690	Nonona
Navut	57	8	76	428	576	631	628	191	160	622	Navut
Lametin	69	17	86	415	546	591	555	189	148	556	Lametin
Morouas	-39	-53	-2	123	180	188	210	8	19	162	Morouas
Batunlamak	25	40	89	19	-20	-49	76	64	109	50	Batunlamak
Matae	93	126	97	468	710	672	682	230	274	630	Matae
Akei	52	76	86	353	449	489	468	159	156	410	Akei
Fortsenal	26	-22	6	242	359	369	393	70	86	329	Fortsenal
Penantsiro	66	19	20	275	399	390	388	89	114	348	Penantsiro
Tasiriki	194	161	183	382	425	432	482	239	255	435	Tasiriki
Wailapa	146	147	159	328	324	392	432	194	281	370	Wailapa
Roria	0	70	51	193	372	280	332	113	46	352	Roria
Amblong	-76	-89	-100	-29	-39	-29	5	-81	-98	-7	Amblong
Narango	86	103	141	78	13	0	100	64	99	-8	Narango
Polonombau	-188	-162	-178	-418	-334	-340	-368	-298	-452	-554	Polonombau
Butmas	-233	-207	-257	-499	-389	-358	-375	-337	-493	-601	Butmas
Tur	-256	-303	-295	-512	-349	-385	-405	-407	-518	-607	Tur
Nambel	-81	-132	-99	-113	34	67	17	-133	-183	-30	Nambel
Tambotalo	139	166	193	60	-30	-84	22	200	177	9	Tambotalo
Sakao	-205	-294	-257	-427	-245	-221	-208	-273	-411	-348	Sakao
Lorediakar	-180	-224	-175	-438	-322	-336	-316	-316	-387	-438	Lorediakar
SB1	-234	-249	-282	-458	-309	-308	-348	-351	-439	-483	SB1
SB2	-263	-217	-283	-493	-316	-375	-385	-370	-435	-517	SB2
Mafea	242	268	394	332	52	136	94	382	403	313	Mafea
Tutuba	266	259	323	217	-120	-75	-62	326	393	199	Tutuba
Aore	215	213	265	198	-96	-130	3	313	343	159	Aore
Malo-North	188	286	242	192	-47	-134	-56	247	394	186	Malo-North
Malo-South	232	134	270	122	-138	-19	-56	281	281	125	Malo-South
Tangoa	100	-4	132	159	14	76	149	148	120	132	Tangoa
Araki	319	197	358	419	168	240	302	340	357	280	Araki
	Huk	Val	Nok	Tas	Wus	Wus	Ker	Vun	Pia	Tol	

[56a]

Linear-correlation measures x1000 from $(\chi^2\text{-}b)/N$ measurements
of 41 communalects of Espiritu Santo
(taken after the removal of near-empty rows and columns
from tables of possible sound correspondences calculated
on unphonemicized lists)

(continued next page)

	Malmariv										
Nonona	753	Nonona									
Navut	948	748	Navut								
Lametin	961	726	943	Lametin							
Morouas	417	254	398	484	Morouas						
Batunlamak	110	27	102	91	763	Batunlamak					
Matae	940	830	947	894	353	40	Matae				
Akei	760	613	784	744	540	233	799	Akei			
Fortsenal	679	544	746	739	741	494	688	853	Fortsenal		
Penantsiro	685	532	694	652	593	357	716	963	871	Penantsiro	
Tasiriki	609	517	673	625	555	264	646	891	773	918	Tasiriki
Wailapa	602	503	603	563	641	504	625	839	850	879	Wailapa
Roria	602	426	527	616	671	493	555	498	625	510	Roria
Amblong	141	42	116	202	867	869	132	270	540	346	Amblong
Narango	75	92	90	141	771	809	88	293	483	359	Narango
Polonombau	-307	-377	-290	-241	-168	-43	-291	-223	-220	-236	Polonombau
Butmas	-339	-377	-293	-293	-243	-168	-408	-349	-308	-335	Butmas
Tur	-304	-347	-288	-255	-194	-81	-345	-270	-286	-277	Tur
Nambel	235	80	206	263	758	755	186	298	575	416	Nambel
Tambotalo	74	2	12	98	615	666	48	208	369	326	Tambotalo
Sakao	43	-233	23	86	151	24	-18	30	95	32	Sakao
Lorediakar	-338	-324	-288	-298	-296	-144	-340	-313	-358	-324	Lorediakar
SB1	-292	-342	-276	-261	-281	-164	-340	-300	-339	-300	SB1
SB2	-349	-347	-327	-313	-292	-233	-362	-337	-363	-322	SB2
Mafea	64	80	96	158	467	510	103	183	231	209	Mafea
Tutuba	-107	-133	-124	-131	262	445	-95	11	81	52	Tutuba
Aore	-116	-152	-96	-73	276	486	-108	15	87	98	Aore
Malo-North	-98	-136	-112	-156	319	501	-87	84	108	172	Malo-North
Malo-South	-177	-100	-126	-69	210	360	-80	86	73	130	Malo-South
Tangoa	228	110	258	269	699	681	190	438	573	483	Tangoa
Araki	274	213	333	277	490	498	293	545	525	576	Araki
	Mal	Non	Nav	Lam	Mor	Bat	Mat	Ake	For	Pen	

[56b]

Linear-correlation measures x1000 from $(\chi^2\text{-}b)/N$ measurements
of 41 communalects of Espiritu Santo
(taken after the removal of near-empty rows and columns
from tables of possible sound correspondences calculated
on unphonemicized lists)

(continued next page)

	Tasiriki										
Wailapa	868	Wailapa									
Roria	390	511	Roria								
Amblong	305	524	581	Amblong							
Narango	311	564	548	825	Narango						
Polonombau	-314	-304	67	-67	27	Polonombau					
Butmas	-390	-394	-92	-77	-20	893	Butmas				
Tur	-362	-332	13	-56	-13	890	924	Tur			
Nambel	344	491	602	783	712	153	3	18	Nambel		
Tambotalo	366	511	385	615	707	-11	-74	-68	553	Tambotalo	
Sakao	-6	77	354	136	124	431	347	513	273	229	Sakao
Lorediakar	-383	-386	-92	-146	-132	511	538	689	-81	-282	Lorediakar
SB1	-354	-352	-164	-118	-110	452	593	624	-105	-245	SB1
SB2	-412	-429	-177	-176	-209	524	520	634	-98	-236	SB2
Mafea	245	363	280	436	507	-209	-258	-244	290	752	Mafea
Tutuba	109	293	70	404	389	-97	-253	-253	266	664	Tutuba
Aore	150	291	142	411	506	-157	-212	-227	192	668	Aore
Malo-North	281	298	32	338	448	-242	-344	-308	115	628	Malo-North
Malo-South	168	363	-81	338	366	-210	-241	-281	193	436	Malo-South
Tangoa	434	630	377	700	640	-226	-287	-258	542	546	Tangoa
Araki	613	685	190	412	520	-252	-348	-357	253	515	Araki
	Tas	Wai	Ror	Amb	Nar	Pol	But	Tur	Nam	Tam	

	Sakao										
Lorediakar	340	Lorediakar									
SB1	369	963	SB1								
SB2	234	973	944	SB2							
Mafea	-42	-245	-277	-278	Mafea						
Tutuba	-152	-185	-252	-258	876	Tutuba					
Aore	-75	-211	-230	-229	833	929	Aore				
Malo-North	-231	-259	-291	-262	572	683	672	Malo-North			
Malo-South	-188	-213	-187	-245	524	647	624	767	Malo-South		
Tangoa	-12	-227	-219	-271	575	506	477	348	588	Tangoa	
Araki	-227	-466	-444	-496	514	452	496	482	592	817	Araki
	Sak	Lor	SB1	SB2	Maf	Tut	Aor	Mal	Mal	Tan	

[56c]

Linear-correlation measures x1000 from $(\chi^2-b)/N$ measurements
of 41 communalects of Espiritu Santo
(taken after the removal of near-empty rows and columns
from tables of possible sound correspondences calculated
on unphonemicized lists)

(concluded)

	Hiw	Tog	Leh	Leh	Mot	Mot	Vat	Sas	Vet	Mos	
	Hiw										
Toga	685	Toga									
Lehali	560	607	Lehali								
Lehalurup	540	584	776	Lehalurup							
Motlav	502	565	639	713	Motlav						
Mota	519	582	679	720	717	Mota					
Vatrata	512	575	676	722	708	734	Vatrata				
Sasar	511	556	686	738	686	730	845	Sasar			
Vetumboso	483	559	633	667	653	724	799	731	Vetumboso		
Mosina	506	545	636	684	667	718	796	727	884	Mosina	
Nume	483	534	568	635	622	730	652	624	690	676	Nume
Dorig	462	492	539	594	568	612	607	605	622	614	Dorig
Koro	445	469	537	587	557	593	621	587	620	610	Koro
Wetamut	478	508	563	614	599	649	613	608	665	654	Wetamut
Lakona	445	485	527	569	572	598	623	557	620	607	Lakona
Merig	498	534	592	643	612	692	632	594	635	637	Merig
Merlav	496	540	589	640	607	680	632	593	629	630	Merlav
Marino	482	518	602	613	576	637	552	537	547	558	Marino
Peterara	415	466	502	545	514	548	492	486	492	492	Peterara
Navenevene	410	444	496	519	483	527	477	464	492	473	Navenevene
Tam	457	500	582	598	560	608	551	525	557	552	Tam
Nasawa	420	452	497	519	522	556	497	484	500	508	Nasawa
Narovorovo	425	473	538	564	527	569	524	503	535	535	Narovorovo
Baetora	448	486	552	573	536	584	533	508	549	541	Baetora
Wailengi	384	402	437	461	416	468	431	446	444	439	Wailengi
Lolomatui	388	402	432	461	417	458	426	441	427	422	Lolomatui
Ngwatua	385	407	425	451	421	444	437	454	409	418	Ngwatua
Lolsiwoi	439	471	503	527	497	545	500	505	527	532	Lolsiwoi
Raga	350	374	406	444	407	427	418	405	409	400	Raga
Apma	332	356	378	384	373	393	376	369	361	361	Apma
Seke	374	371	409	434	391	428	432	415	423	433	Seke
Sa	331	333	369	393	381	393	378	379	376	382	Sa
Sowa	370	381	395	422	392	420	424	413	396	407	Sowa
Bek	545	609	719	795	799	788	809	860	706	725	Bek

[57a]

Proportions of shared cognates x1000 for 34 communalects
of the Northeast New Hebrides
(according to Tryon 1976)

(continued next page)

	Nume										
Dorig	713	Dorig									
Koro	667	725	Koro								
Wetamut	762	814	689	Wetamut							
Lakona	643	615	718	643	Lakona						
Merig	760	680	640	711	611	Merig					
Merlav	755	664	636	701	603	945	Merlav				
Marino	623	548	534	611	521	644	649	Marino			
Peterara	538	504	510	519	502	578	575	705	Peterara		
Navenevene	517	494	500	519	496	559	563	681	792	Navenevene	
Tam	603	547	560	584	555	626	636	710	850	979	Tam
Nasawa	511	497	497	500	486	556	558	681	720	805	Nasawa
Narovorovo	548	492	511	528	522	577	587	638	762	856	Narovorovo
Baetora	585	508	533	548	530	613	624	638	769	833	Baetora
Wailengi	443	419	430	428	408	484	480	574	555	568	Wailengi
Lolomatui	426	410	427	417	398	481	483	564	565	578	Lolomatui
Ngwatua	421	401	400	415	373	460	456	542	528	515	Ngwatua
Lolsiwoi	545	478	497	519	492	568	558	566	593	622	Lolsiwoi
Raga	422	393	385	416	398	462	469	547	536	533	Raga
Apma	373	359	383	361	352	398	402	471	444	441	Apma
Seke	407	399	417	387	400	432	421	468	465	459	Seke
Sa	380	380	387	374	367	407	402	450	439	417	Sa
Sowa	417	397	405	394	397	430	429	485	465	464	Sowa
Bek	691	669	664	718	619	672	652	634	599	564	Bek
	Num	Dor	Kor	Wet	Lak	Mer	Mer	Mar	Pet	Nav	

[57b]

Proportions of shared cognates x1000 for 34 communalects
of the Northeast New Hebrides
(according to Tryon 1976)

(continued next page)

	Tam										
Nasawa	891	Nasawa									
Narovorovo	882	949	Narovorovo								
Baetora	859	897	887	Baetora							
Wailengi	624	598	634	694	Wailengi						
Lolomatui	617	603	632	683	914	Lolomatui					
Ngwatua	596	566	589	610	641	636	Ngwatua				
Lolsiwoi	631	686	656	683	826	820	731	Lolsiwoi			
Raga	587	574	586	623	569	578	532	588	Raga		
Apma	467	468	453	467	418	424	407	434	521	Apma	
Seke	518	472	511	522	418	411	427	496	463	599	Seke
Sa	475	500	475	492	414	427	407	429	423	531	Sa
Sowa	528	508	509	512	425	438	430	494	511	646	Sowa
Bek	597	593	562	546	511	485	548	570	507	485	Bek
	Tam	Nas	Nar	Bae	Wai	Lol	Ngw	Lol	Rag	Apm	

	Seke			
Sa	613	Sa		
Sowa	774	604	Sowa	
Bek	521	489	515	Bek
	Sek	Sa	Sow	

[57c]

Proportions of shared cognates x1000 for 34 communalects
of the Northeast New Hebrides
(according to Tryon 1976)

(concluded)

	Hiw										
Toga	988	Toga									
Lehali	848	887	Lehali								
Lehalurup	809	863	981	Lehalurup							
Motlav	802	843	962	975	Motlav						
Mota	783	829	933	963	980	Mota					
Vatrata	745	788	904	934	951	943	Vatrata				
Sasar	749	808	913	942	962	929	979	Sasar			
Vetumboso	726	745	856	893	918	930	954	918	Vetumboso		
Mosina	705	781	871	895	922	939	963	934	996	Mosina	
Nume	656	691	780	805	841	879	803	770	847	844	Nume
Dorig	593	632	702	746	790	831	762	722	811	802	Dorig
Koro	573	631	698	745	800	835	765	739	819	811	Koro
Wetamut	611	665	726	773	808	854	784	746	811	805	Wetamut
Lakona	626	666	757	799	826	867	799	787	859	853	Lakona
Merig	572	618	669	702	731	787	661	644	721	709	Merig
Merlav	572	602	666	694	722	786	643	624	711	699	Merlav
Marino	254	300	373	390	400	420	244	268	297	271	Marino
Peterara	-45	-27	92	53	55	88	-88	-90	-35	-50	Peterara
Navenevene	-103	-62	20	-5	-5	23	-148	-151	-97	-95	Navenevene
Tam	-125	-88	-19	-24	-20	8	-150	-143	-88	-103	Tam
Nasawa	-183	-141	-44	-67	-93	-67	-211	-207	-156	-175	Nasawa
Narovorovo	-169	-150	-65	-89	-81	-55	-205	-200	-150	-165	Narovorovo
Baetora	-226	-194	-119	-132	-132	-98	-252	-242	-189	-199	Baetora
Wailengi	-287	-283	-234	-252	-267	-249	-345	-308	-310	-315	Wailengi
Lolomatui	-325	-318	-269	-297	-315	-288	-390	-351	-344	-349	Lolomatui
Ngwatua	-245	-240	-163	-189	-205	-190	-314	-264	-278	-286	Ngwatua
Lolsiwoi	-271	-274	-209	-210	-224	-196	-279	-243	-251	-267	Lolsiwoi
Raga	-487	-476	-366	-398	-393	-374	-508	-458	-481	-475	Raga
Apma	-604	-658	-562	-553	-544	-581	-566	-531	-606	-588	Apma
Seke	-523	-530	-445	-457	-414	-461	-431	-402	-462	-452	Seke
Sa	-585	-608	-488	-489	-464	-487	-481	-461	-515	-500	Sa
Sowa	-578	-619	-516	-531	-503	-536	-522	-499	-532	-526	Sowa
Bek	711	738	887	923	944	913	937	955	891	890	Bek
	Hiw	Tog	Leh	Leh	Mot	Mot	Vat	Sas	Vet	Mos	

[58a]

Linear-correlation measures x1000 for 34 communalects
of the Northeast New Hebrides
(calculated from proportions of shared cognates)

(continued next page)

	Nume										
Dorig	952	Dorig									
Koro	938	955	Koro								
Wetamut	969	985	963	Wetamut							
Lakona	945	953	971	941	Lakona						
Merig	921	849	834	877	841	Merig					
Merlav	910	844	818	866	837	997	Merlav				
Marino	443	397	389	397	432	551	564	Marino			
Peterara	118	40	76	104	119	218	255	823	Peterara		
Navenevene	60	-19	24	28	65	159	189	769	984	Navenevene	
Tam	34	-11	11	18	55	154	174	798	969	986	Tam
Nasawa	-28	-129	-75	-54	-22	69	102	671	950	961	Nasawa
Narovorovo	-41	-93	-61	-65	-31	78	99	752	932	951	Narovorovo
Baetora	-80	-117	-101	-93	-58	58	78	752	916	939	Baetora
Wailengi	-242	-309	-281	-256	-270	-124	-101	366	584	594	Wailengi
Lolomatui	-272	-345	-326	-294	-304	-157	-142	368	567	583	Lolomatui
Ngwatua	-224	-304	-277	-247	-252	-107	-84	457	664	688	Ngwatua
Lolsiwoi	-205	-226	-224	-214	-228	-81	-53	471	565	559	Lolsiwoi
Raga	-365	-436	-394	-391	-405	-217	-199	444	731	755	Raga
Apma	-571	-563	-568	-569	-564	-490	-490	-117	200	223	Apma
Seke	-456	-463	-439	-456	-442	-434	-415	-175	116	159	Seke
Sa	-476	-475	-445	-478	-464	-416	-405	-44	260	319	Sa
Sowa	-540	-524	-497	-532	-521	-474	-472	-192	141	180	Sowa
Bek	764	739	715	712	769	624	628	231	-145	-191	Bek
	Num	Dor	Kor	Wet	Lak	Mer	Mer	Mar	Pet	Nav	

[58b]

Linear-correlation measures x1000 for 34 communalects
of the Northeast New Hebrides
(calculated from proportions of shared cognates)

(continued next page)

	Tam										
Nasawa	921	Nasawa									
Narovorovo	959	978	Narovorovo								
Baetora	944	956	986	Baetora							
Wailengi	518	652	611	648	Wailengi						
Lolomatui	530	648	617	665	996	Lolomatui					
Ngwatua	573	731	680	721	939	937	Ngwatua				
Lolsiwoi	565	577	636	696	946	938	911	Lolsiwoi			
Raga	670	801	759	771	816	816	880	763	Raga		
Apma	163	276	243	209	208	224	302	126	521	Apma	
Seke	49	241	115	67	90	125	165	-78	439	914	Seke
Sa	209	308	290	238	215	207	302	155	603	942	Sa
Sowa	60	209	155	127	136	132	206	-22	427	941	Sowa
Bek	-147	-275	-212	-240	-404	-412	-385	-338	-553	-599	Bek
	Tam	Nas	Nar	Bae	Wai	Lol	Ngw	Lol	Rag	Apm	

	Seke			
Sa	926	Sa		
Sowa	965	950	Sowa	
Bek	-480	-539	-547	Bek
	Sek	Sa	Sow	

[58c]

Linear-correlation measures x1000 for 34 communalects
of the Northeast New Hebrides
(calculated from proportions of shared cognates)

(concluded)

	Hiw										
Toga	108	Toga									
Lehali	103	100	Lehali								
Lehalurup	59	74	197	Lehalurup							
Motlav	34	56	102	169	Motlav						
Mota	52	72	110	139	110	Mota					
Vatrata	67	103	149	183	149	177	Vatrata				
Sasar	77	96	143	206	124	155	333	Sasar			
Vetumboso	60	79	115	129	102	139	234	190	Vetumboso		
Mosina	66	70	121	150	109	161	264	235	292	Mosina	
Bek	67	86	151	251	209	102	248	278	104	146	Bek
Nume	75	84	110	113	105	194	188	139	153	148	Nume
Dorig	42	48	72	70	61	99	113	91	90	85	Dorig
Koro	52	57	113	108	72	108	154	118	99	117	Koro
Wetamut	53	57	89	93	83	126	122	103	107	103	Wetamut
Lakona	39	60	77	84	63	101	138	123	116	103	Lakona
Merig	60	94	127	136	118	183	190	144	140	157	Merig
Merlav	58	90	109	133	107	166	174	144	123	140	Merlav
Marino	46	55	94	84	56	136	108	82	62	78	Marino
Peterara	40	53	69	77	59	88	91	93	74	71	Peterara
Navenevene	35	55	64	84	57	112	101	100	72	87	Navenevene
Tam	35	67	92	86	65	143	120	113	80	117	Tam
Nasawa	34	44	73	79	66	97	101	87	80	77	Nasawa
Narovorovo	24	47	75	68	48	118	95	66	57	75	Narovorovo
Baetora	26	52	86	57	52	122	103	101	74	77	Baetora
Wailengi	40	55	65	56	37	70	76	70	56	59	Wailengi
Lolomatui	39	47	64	58	34	67	69	76	42	56	Lolomatui
Ngwatua	41	49	78	78	52	70	70	82	55	48	Ngwatua
Lolsiwoi	53	58	66	68	43	92	75	81	55	65	Lolsiwoi
Raga	34	32	47	56	45	62	66	55	57	52	Raga
Apma	23	22	39	56	34	43	48	49	30	44	Apma
Seke	23	31	51	46	33	50	80	60	51	55	Seke
Sa	20	21	40	53	35	45	46	52	31	42	Sa
Sowa	37	26	44	39	30	45	67	52	36	44	Sowa
	Hiw	Tog	Leh	Leh	Mot	Mot	Vat	Sas	Vet	Mos	

[59a]

(χ^2-b)/N measurements x1000 for 34 communalects
of the Northeast New Hebrides
(taken after the removal of near-empty rows and columns
from tables of possible sound correspondences calculated
on unphonemicized lists)

(continued next page)

	Bek										
Nume	144	Nume									
Dorig	58	163	Dorig								
Koro	128	178	150	Koro							
Wetamut	101	211	220	174	Wetamut						
Lakona	83	152	109	183	114	Lakona					
Merig	149	212	134	163	178	144	Merig				
Merlav	136	207	126	154	172	133	567	Merlav			
Marino	88	131	91	84	107	82	155	150	Marino		
Peterara	81	89	78	111	76	84	110	103	178	Peterara	
Navenevene	102	99	80	101	102	94	123	108	178	232	Navenevene
Tam	100	148	83	132	128	101	156	152	205	257	Tam
Nasawa	69	72	80	107	93	83	151	149	189	137	Nasawa
Narovorovo	63	108	74	80	79	72	114	118	161	153	Narovorovo
Baetora	60	126	66	94	98	102	130	116	170	189	Baetora
Wailengi	99	78	71	80	55	74	93	86	125	116	Wailengi
Lolomatui	84	80	63	61	54	64	85	75	112	104	Lolomatui
Ngwatua	81	71	54	63	65	58	95	88	100	110	Ngwatua
Lolsiwoi	103	92	68	66	58	89	115	90	96	110	Lolsiwoi
Raga	45	74	42	49	58	47	88	76	101	100	Raga
Apma	31	39	34	41	37	27	50	43	38	44	Apma
Seke	82	53	42	44	52	47	71	70	48	57	Seke
Sa	53	34	24	35	34	39	49	50	29	42	Sa
Sowa	38	48	22	39	51	40	57	63	35	50	Sowa
	Bek	Num	Dor	Kor	Wet	Lak	Mer	Mer	Mar	Pet	

[59b]

($(\chi^2$-b)/N measurements x1000 for 34 communalects
of the Northeast New Hebrides
(taken after the removal of near-empty rows and columns
from tables of possible sound correspondences calculated
on unphonemicized lists)

(continued next page)

	Navenevene										
Tam	518	Tam									
Nasawa	263	307	Nasawa								
Narovorovo	307	356	351	Narovorovo							
Baetora	300	332	238	295	Baetora						
Wailengi	130	138	115	112	163	Wailengi					
Lolomatui	128	123	90	135	145	350	Lolomatui				
Ngwatua	96	109	96	72	122	179	148	Ngwatua			
Lolsiwoi	137	143	108	100	153	236	191	155	Lolsiwoi		
Raga	123	134	121	100	162	131	123	104	104	Raga	
Apma	50	48	43	38	46	48	60	50	60	76	Apma
Seke	67	69	59	40	49	44	52	73	40	40	Seke
Sa	62	54	73	24	48	50	40	48	34	33	Sa
Sowa	50	48	51	27	39	34	36	50	52	57	Sowa
	Nav	Tam	Nas	Nar	Bae	Wai	Lol	Ngw	Lol	Rag	

	Apma			
Seke	86	Seke		
Sa	80	91	Sa	
Sowa	114	202	98	Sowa
	Apm	Sek	Sa	

[59c]

$(\chi^2-b)/N$ measurements x1000 for 34 communalects
of the Northeast New Hebrides
(taken after the removal of near-empty rows and columns
from tables of possible sound correspondences calculated
on unphonemicized lists)

(concluded)

	Hiw										
Toga	858	Toga									
Lehali	659	799	Lehali								
Lehalurup	612	709	872	Lehalurup							
Motlav	566	674	884	946	Motlav						
Mota	440	665	690	584	658	Mota					
Vatrata	571	672	791	846	826	735	Vatrata				
Sasar	521	685	826	851	898	646	944	Sasar			
Vetumboso	540	633	691	682	685	759	871	824	Vetumboso		
Mosina	497	707	738	727	741	743	912	846	968	Mosina	
Bek	519	658	837	889	865	653	777	809	706	706	Bek
Nume	372	605	595	537	597	845	622	582	630	652	Nume
Dorig	294	409	374	292	361	691	401	308	464	440	Dorig
Koro	359	569	503	480	590	779	560	543	614	549	Koro
Wetamut	277	444	427	350	438	732	487	372	475	470	Wetamut
Lakona	386	537	549	449	511	814	610	505	604	644	Lakona
Merig	260	449	366	358	415	640	396	354	394	387	Merig
Merlav	253	449	427	337	430	671	413	318	412	403	Merlav
Marino	-127	150	45	0	82	447	20	39	71	60	Marino
Peterara	-269	-24	-49	-122	-98	316	-68	-82	-103	-30	Peterara
Navenevene	-325	-88	-82	-169	-120	225	-109	-112	-125	-76	Navenevene
Tam	-273	-96	-134	-134	-104	229	-93	-106	-73	-96	Tam
Nasawa	-350	-68	-91	-156	-130	302	-105	-127	-127	-53	Nasawa
Narovorovo	-267	-55	-98	-145	-97	248	-105	-75	-74	-63	Narovorovo
Baetora	-295	-95	-195	-181	-181	221	-156	-180	-137	-96	Baetora
Wailengi	-194	-153	-225	-234	-245	-81	-275	-191	-281	-256	Wailengi
Lolomatui	-207	-121	-232	-249	-248	-75	-260	-227	-248	-257	Lolomatui
Ngwatua	-165	-69	-169	-188	-168	-6	-214	-186	-252	-195	Ngwatua
Lolsiwoi	-199	-40	-142	-180	-153	11	-179	-136	-192	-190	Lolsiwoi
Raga	-400	-200	-260	-330	-285	99	-296	-246	-264	-223	Raga
Apma	-424	-580	-437	-381	-314	-388	-349	-328	-296	-327	Apma
Seke	-181	-286	-219	-58	-20	-173	-104	-20	-114	-72	Seke
Sa	-367	-395	-248	-123	-63	-257	-122	-110	-158	-157	Sa
Sowa	-332	-358	-341	-215	-208	-302	-238	-193	-168	-173	Sowa
	Hiw	Tog	Leh	Leh	Mot	Mot	Vat	Sas	Vet	Mos	

[60a]

Linear-correlation measures x1000 from $(\chi^2-b)/N$ measurements
of 34 communalects of the Northeast New Hebrides
(taken after the removal of near-empty rows and columns
from tables of possible sound correspondences calculated
on unphonemicized lists)

(continued next page)

	Bek										
Nume	432	Nume									
Dorig	339	849	Dorig								
Koro	413	870	856	Koro							
Wetamut	297	873	939	860	Wetamut						
Lakona	496	865	823	908	847	Lakona					
Merig	305	680	549	597	614	599	Merig				
Merlav	324	665	563	607	620	616	986	Merlav			
Marino	-7	333	360	447	369	405	369	376	Marino		
Peterara	-70	178	156	187	216	270	153	167	860	Peterara	
Navenevene	-159	101	72	158	104	141	82	113	812	928	Navenevene
Tam	-126	60	155	144	117	203	120	119	829	919	Tam
Nasawa	-125	187	122	137	155	192	164	173	807	891	Nasawa
Narovorovo	-127	91	124	210	180	211	159	144	873	893	Narovorovo
Baetora	-141	55	147	156	101	156	103	131	875	925	Baetora
Wailengi	-185	-122	-75	-131	-135	-20	-49	-55	417	466	Wailengi
Lolomatui	-141	-152	-61	-90	-156	-24	-64	-51	440	478	Lolomatui
Ngwatua	-46	-90	-14	-48	-133	47	25	38	542	526	Ngwatua
Lolsiwoi	-136	-40	2	48	-31	50	5	81	617	597	Lolsiwoi
Raga	-222	-87	7	-17	-53	57	30	75	765	800	Raga
Apma	-189	-474	-408	-483	-380	-372	-262	-179	-207	-89	Apma
Seke	-160	-203	-235	-152	-145	-182	-64	-31	-177	-108	Seke
Sa	-90	-356	-316	-230	-267	-268	-101	-86	-40	2	Sa
Sowa	-132	-334	-225	-298	-285	-286	-118	-130	-282	-236	Sowa
	Bek	Num	Dor	Kor	Wet	Lak	Mer	Mer	Mar	Pet	

[60b]

Linear-correlation measures x1000 from $(\chi^2-b)/N$ measurements of 34 communalects of the Northeast New Hebrides (taken after the removal of near-empty rows and columns from tables of possible sound correspondences calculated on unphonemicized lists)

(continued next page)

	Navenevene										
Tam	985	Tam									
Nasawa	899	897	Nasawa								
Narovorovo	921	921	957	Narovorovo							
Baetora	930	934	929	938	Baetora						
Wailengi	380	339	334	415	473	Wailengi					
Lolomatui	372	372	405	350	510	970	Lolomatui				
Ngwatua	441	380	373	483	531	882	900	Ngwatua			
Lolsiwoi	476	449	447	536	593	898	926	929	Lolsiwoi		
Raga	733	709	700	797	831	711	719	791	831	Raga	
Apma	-69	-105	-68	-88	-60	89	26	141	-34	65	Apma
Seke	-89	-127	-104	-85	-135	-152	-178	-170	-117	-31	Seke
Sa	6	22	-89	150	-27	-129	-34	28	7	138	Sa
Sowa	-170	-190	-201	-173	-211	-200	-176	-117	-291	-169	Sowa
	Nav	Tam	Nas	Nar	Bae	Wai	Lol	Ngw	Lol	Rag	

	Apma			
Seke	742	Seke		
Sa	732	804	Sa	
Sowa	707	729	742	Sowa
	Apm	Sek	Sa	

[60c]

Linear-correlation measures x1000 from $(\chi^2-b)/N$ measurements
of 34 communalects of the Northeast New Hebrides
(taken after the removal of near-empty rows and columns
from tables of possible sound correspondences calculated
on unphonemicized lists)

(concluded)

	Hiw										
Toga	96	Toga									
Lehali	82	93	Lehali								
Lehalurup	43	74	228	Lehalurup							
Motlav	36	72	137	180	Motlav						
Mota	51	72	119	143	152	Mota					
Vatrata	55	103	163	215	230	196	Vatrata				
Sasar	65	108	179	241	217	181	416	Sasar			
Vetumboso	47	80	113	157	171	151	296	242	Vetumboso		
Mosina	56	93	164	203	171	197	347	338	347	Mosina	
Bek	53	58	130	278	259	108	272	324	80	139	Bek
Nume	61	86	124	128	147	207	237	182	195	196	Nume
Dorig	40	46	75	91	85	126	153	146	139	130	Dorig
Koro	45	57	110	135	99	114	182	159	119	162	Koro
Wetamut	47	56	99	114	108	136	153	153	140	158	Wetamut
Lakona	39	51	86	76	80	109	158	117	131	120	Lakona
Merig	48	88	134	151	150	189	222	172	156	190	Merig
Merlav	53	87	117	150	136	176	203	171	139	176	Merlav
Marino	42	49	92	81	62	142	105	82	62	95	Marino
Peterara	39	41	82	80	71	90	97	99	73	87	Peterara
Navenevene	24	49	66	79	67	109	95	81	73	89	Navenevene
Tam	28	55	98	77	82	148	112	98	86	109	Tam
Nasawa	15	23	82	87	75	95	107	99	89	78	Nasawa
Narovorovo	19	33	86	78	64	126	93	65	75	73	Narovorovo
Baetora	22	33	88	46	69	126	93	94	75	75	Baetora
Wailengi	25	41	69	57	44	68	82	81	44	64	Wailengi
Lolomatui	32	48	76	63	50	63	79	95	39	67	Lolomatui
Ngwatua	26	34	75	77	50	69	80	96	55	63	Ngwatua
Lolsiwoi	31	38	59	56	51	96	71	62	57	73	Lolsiwoi
Raga	25	32	55	52	48	64	79	70	56	58	Raga
Apma	28	27	56	74	50	50	67	60	39	55	Apma
Seke	6	19	63	47	41	47	93	66	53	63	Seke
Sa	21	12	41	48	53	42	55	60	25	50	Sa
Sowa	33	19	51	45	45	53	76	63	48	57	Sowa
	Hiw	Tog	Leh	Leh	Mot	Mot	Vat	Sas	Vet	Mos	

[61a]

$(\chi^2\text{-}b)/N$ measurements x1000 for 34 communalects
of the Northeast New Hebrides
(taken after the removal of near-empty rows and columns
from tables of possible sound correspondences calculated
on unphonemicized lists, affixes removed)

(continued next page)

	Bek										
Nume	149	Nume									
Dorig	71	202	Dorig								
Koro	130	208	200	Koro							
Wetamut	142	250	273	201	Wetamut						
Lakona	47	172	149	226	139	Lakona					
Merig	152	238	167	192	196	147	Merig				
Merlav	144	233	163	179	193	148	583	Merlav			
Marino	84	141	78	93	108	73	169	164	Marino		
Peterara	91	92	76	117	78	88	117	112	205	Peterara	
Navenevene	75	99	84	96	100	86	125	112	177	245	Navenevene
Tam	64	155	82	134	141	87	162	163	204	289	Tam
Nasawa	62	77	102	102	103	92	155	153	229	151	Nasawa
Narovorovo	41	106	79	66	79	60	120	121	168	168	Narovorovo
Baetora	72	132	66	91	115	84	129	120	176	196	Baetora
Wailengi	70	74	77	79	52	56	91	81	128	118	Wailengi
Lolomatui	54	76	53	61	57	50	87	76	120	110	Lolomatui
Ngwatua	53	72	59	61	72	51	94	88	124	106	Ngwatua
Lolsiwoi	66	93	42	60	83	71	110	101	103	107	Lolsiwoi
Raga	29	80	47	42	60	46	89	80	117	107	Raga
Apma	45	57	51	55	50	31	63	56	45	51	Apma
Seke	68	52	37	58	55	52	76	74	51	59	Seke
Sa	32	36	22	41	37	42	45	49	27	48	Sa
Sowa	61	56	42	48	63	47	57	60	50	53	Sowa
	Bek	Num	Dor	Kor	Wet	Lak	Mer	Mer	Mar	Pet	

[61b]

(χ^2-b)/N measurements x1000 for 34 communalects
of the Northeast New Hebrides
(taken after the removal of near-empty rows and columns
from tables of possible sound correspondences calculated
on unphonemicized lists, affixes removed)

(continued next page)

	Navenevene										
Tam	538	Tam									
Nasawa	243	313	Nasawa								
Narovorovo	329	380	405	Narovorovo							
Baetora	306	344	278	315	Baetora						
Wailengi	123	130	120	111	166	Wailengi					
Lolomatui	137	113	90	96	145	397	Lolomatui				
Ngwatua	95	114	84	82	115	192	155	Ngwatua			
Lolsiwoi	132	142	91	105	156	268	192	165	Lolsiwoi		
Raga	127	136	128	120	174	133	130	93	126	Raga	
Apma	59	63	48	45	54	56	77	54	78	91	Apma
Seke	55	69	63	43	54	43	45	63	43	52	Seke
Sa	56	48	68	24	41	47	49	51	38	35	Sa
Sowa	60	57	54	35	45	42	42	54	63	63	Sowa
	Nav	Tam	Nas	Nar	Bae	Wai	Lol	Ngw	Lol	Rag	

	Apma			
Seke	122	Seke		
Sa	127	123	Sa	
Sowa	142	268	118	Sowa
	Apm	Sek	Sa	

[61c]

$(\chi^2-b)/N$ measurements x1000 for 34 communalects
of the Northeast New Hebrides
(taken after the removal of near-empty rows and columns
from tables of possible sound correspondences calculated
on unphonemicized lists, affixes removed)

(concluded)

	Hiw										
Toga	857	Toga									
Lehali	544	778	Lehali								
Lehalurup	618	726	889	Lehalurup							
Motlav	579	724	839	948	Motlav						
Mota	454	740	747	653	726	Mota					
Vatrata	538	748	811	864	910	771	Vatrata				
Sasar	513	708	812	892	941	688	954	Sasar			
Vetumboso	497	746	778	700	751	833	887	843	Vetumboso		
Mosina	538	785	770	765	847	776	934	862	960	Mosina	
Bek	436	701	876	840	866	678	771	795	741	763	Bek
Nume	464	658	599	614	671	866	683	655	728	736	Nume
Dorig	361	522	479	447	505	698	560	460	617	614	Dorig
Koro	404	570	553	515	610	779	624	570	699	610	Koro
Wetamut	363	515	478	489	558	771	611	509	619	567	Wetamut
Lakona	358	523	445	438	479	721	539	485	638	639	Lakona
Merig	319	485	413	420	451	670	439	406	463	450	Merig
Merlav	265	468	458	390	460	682	450	371	472	446	Merlav
Marino	-229	-50	7	-65	-24	301	-87	-89	12	-70	Marino
Peterara	-324	-126	-57	-136	-117	251	-141	-180	-104	-120	Peterara
Navenevene	-356	-232	-99	-234	-183	156	-217	-232	-147	-192	Navenevene
Tam	-341	-212	-154	-189	-179	140	-196	-231	-109	-173	Tam
Nasawa	-370	-182	-69	-176	-143	261	-163	-216	-105	-116	Nasawa
Narovorovo	-362	-197	-106	-200	-154	155	-194	-195	-106	-149	Narovorovo
Baetora	-402	-236	-198	-213	-219	120	-238	-287	-173	-203	Baetora
Wailengi	-292	-249	-248	-274	-293	-154	-313	-265	-285	-285	Wailengi
Lolomatui	-294	-224	-232	-261	-288	-140	-279	-278	-250	-263	Lolomatui
Ngwatua	-301	-172	-141	-223	-216	10	-221	-252	-200	-165	Ngwatua
Lolsiwoi	-363	-241	-241	-297	-296	-98	-298	-270	-272	-292	Lolsiwoi
Raga	-484	-330	-259	-372	-342	33	-358	-362	-250	-278	Raga
Apma	-426	-470	-291	-287	-237	-360	-271	-239	-230	-243	Apma
Seke	-174	-318	-207	-60	-46	-181	-116	-67	-91	-83	Seke
Sa	-420	-389	-207	-157	-194	-320	-164	-194	-125	-189	Sa
Sowa	-410	-359	-249	-151	-154	-306	-158	-157	-155	-154	Sowa
	Hiw	Tog	Leh	Leh	Mot	Mot	Vat	Sas	Vet	Mos	

[62a]

Linear-correlation measures x1000 from $(\chi^2-b)/N$ measurements of 34 communalects of the Northeast New Hebrides (taken after the removal of near-empty rows and columns from tables of possible sound correspondences calculated on unphonemicized lists, affixes removed)

(continued next page)

	Bek										
Nume	535	Nume									
Dorig	472	883	Dorig								
Koro	494	877	886	Koro							
Wetamut	406	893	923	891	Wetamut						
Lakona	467	855	882	917	866	Lakona					
Merig	384	704	606	624	643	628	Merig				
Merlav	382	692	604	648	639	604	992	Merlav			
Marino	-72	161	209	209	207	230	328	337	Marino		
Peterara	-125	92	20	54	122	81	144	153	818	Peterara	
Navenevene	-213	13	-53	6	29	-12	56	86	768	925	Navenevene
Tam	-158	-24	56	-13	14	85	100	91	802	901	Tam
Nasawa	-147	118	23	32	80	50	147	155	756	859	Nasawa
Narovorovo	-161	-6	-2	42	76	66	119	114	847	867	Narovorovo
Baetora	-240	-50	14	-24	-14	3	77	94	860	899	Baetora
Wailengi	-240	-218	-242	-250	-180	-176	-104	-97	387	381	Wailengi
Lolomatui	-159	-239	-177	-211	-238	-191	-113	-102	365	374	Lolomatui
Ngwatua	-69	-107	-108	-114	-137	-87	14	24	509	512	Ngwatua
Lolsiwoi	-254	-158	-84	-144	-186	-161	-24	-11	566	519	Lolsiwoi
Raga	-239	-177	-150	-170	-147	-124	-10	17	747	752	Raga
Apma	-90	-401	-350	-369	-306	-268	-254	-201	-251	-167	Apma
Seke	-53	-178	-132	-146	-143	-127	-102	-82	-193	-151	Seke
Sa	-5	-356	-226	-246	-302	-250	-179	-194	-170	-159	Sa
Sowa	-98	-281	-210	-210	-241	-202	-161	-164	-288	-213	Sowa
	Bek	Num	Dor	Kor	Wet	Lak	Mer	Mer	Mar	Pet	

[62b]

Linear-correlation measures x1000 from $(\chi^2-b)/N$ measurements
of 34 communalects of the Northeast New Hebrides
(taken after the removal of near-empty rows and columns
from tables of possible sound correspondences calculated
on unphonemicized lists, affixes removed)

(continued next page)

	Navenevene										
Tam	977	Tam									
Nasawa	874	845	Nasawa								
Narovorovo	889	903	944	Narovorovo							
Baetora	921	926	899	940	Baetora						
Wailengi	338	277	253	306	408	Wailengi					
Lolomatui	265	290	278	291	408	957	Lolomatui				
Ngwatua	426	345	394	400	519	867	896	Ngwatua			
Lolsiwoi	439	396	415	429	549	847	921	916	Lolsiwoi		
Raga	723	708	699	756	824	679	659	791	788	Raga	
Apma	-120	-190	-143	-154	-148	67	-33	74	-49	-18	Apma
Seke	-98	-159	-124	-120	-168	-173	-159	-150	-117	-46	Seke
Sa	-130	-108	-207	-36	-133	-106	-62	-67	-45	65	Sa
Sowa	-184	-186	-192	-171	-203	-177	-150	-136	-213	-123	Sowa
	Nav	Tam	Nas	Nar	Bae	Wai	Lol	Ngw	Lol	Rag	

	Apma			
Seke	787	Seke		
Sa	813	795	Sa	
Sowa	738	892	830	Sowa
	Apm	Sek	Sa	

[62c]

Linear-correlation measures x1000 from $(\chi^2\text{-b})/N$ measurements
of 34 communalects of the Northeast New Hebrides
(taken after the removal of near-empty rows and columns
from tables of possible sound correspondences calculated
on unphonemicized lists, affixes removed)

(concluded)

	Hiw										
Toga	148	Toga									
Lehali	117	141	Lehali								
Lehalurup	55	103	265	Lehalurup							
Motlav	45	96	183	244	Motlav						
Mota	53	100	131	153	167	Mota					
Vatrata	82	143	192	237	232	199	Vatrata				
Sasar	113	155	238	314	292	194	482	Sasar			
Vetumboso	80	117	187	202	203	169	347	354	Vetumboso		
Mosina	80	143	214	215	201	184	375	408	447	Mosina	
Bek	66	99	186	237	299	133	290	449	152	189	Bek
Nume	86	112	138	148	157	221	239	224	227	229	Nume
Dorig	45	64	108	122	104	136	171	190	169	154	Dorig
Koro	57	89	126	163	113	128	227	216	162	204	Koro
Wetamut	61	84	123	122	122	146	182	196	189	187	Wetamut
Lakona	46	80	99	93	82	110	200	147	143	145	Lakona
Merig	72	123	167	182	181	209	243	236	195	224	Merig
Merlav	70	119	140	178	156	187	217	210	179	203	Merlav
Marino	54	67	103	91	82	152	100	113	72	106	Marino
Peterara	63	74	107	114	102	109	111	112	98	101	Peterara
Navenevene	28	71	93	81	87	110	114	117	98	100	Navenevene
Tam	30	68	103	87	97	162	122	125	105	147	Tam
Nasawa	19	40	93	92	88	103	115	135	88	100	Nasawa
Narovorovo	22	53	95	86	79	159	115	106	102	91	Narovorovo
Baetora	29	46	99	42	85	137	104	110	83	86	Baetora
Wailengi	42	66	74	67	58	85	85	107	62	81	Wailengi
Lolomatui	39	69	76	61	59	74	85	103	57	71	Lolomatui
Ngwatua	29	54	79	84	58	71	91	125	68	73	Ngwatua
Lolsiwoi	39	59	74	60	75	127	94	94	63	94	Lolsiwoi
Raga	38	45	67	55	62	67	88	88	72	73	Raga
Apma	30	41	59	84	41	51	70	63	52	61	Apma
Seke	26	36	69	63	43	54	101	89	65	81	Seke
Sa	19	32	49	53	53	33	56	70	36	49	Sa
Sowa	43	34	51	58	51	51	86	79	59	64	Sowa
	Hiw	Tog	Leh	Leh	Mot	Mot	Vat	Sas	Vet	Mos	

[63a]

$(\chi^2-b)/N$ measurements x1000 for 34 communalects
of the Northeast New Hebrides
(taken after the removal of near-empty rows and columns
from tables of possible sound correspondences calculated
on near-phonemicized lists, affixes removed)

(continued next page)

	Bek										
Nume	171	Nume									
Dorig	149	224	Dorig								
Koro	154	255	241	Koro							
Wetamut	192	291	334	241	Wetamut						
Lakona	58	188	174	274	157	Lakona					
Merig	139	287	190	238	232	167	Merig				
Merlav	155	265	179	215	224	152	683	Merlav			
Marino	117	143	82	110	123	75	179	179	Marino		
Peterara	85	104	86	148	98	108	152	145	220	Peterara	
Navenevene	90	113	99	124	117	110	150	140	192	295	Navenevene
Tam	92	158	103	142	151	93	166	171	218	322	Tam
Nasawa	87	87	118	118	114	108	177	168	242	203	Nasawa
Narovorovo	72	140	83	97	116	79	142	142	205	240	Narovorovo
Baetora	79	144	82	110	134	88	138	129	192	232	Baetora
Wailengi	87	83	73	89	64	57	102	97	132	137	Wailengi
Lolomatui	55	87	59	76	58	49	105	92	122	124	Lolomatui
Ngwatua	70	79	70	70	83	58	108	97	131	110	Ngwatua
Lolsiwoi	107	124	56	74	102	86	132	124	126	125	Lolsiwoi
Raga	49	88	53	55	74	56	104	98	124	126	Raga
Apma	28	60	48	62	59	38	69	63	46	56	Apma
Seke	24	61	55	69	61	69	88	86	46	55	Seke
Sa	50	41	34	49	43	36	50	49	41	48	Sa
Sowa	25	60	51	51	57	51	61	60	56	58	Sowa
	Bek	Num	Dor	Kor	Wet	Lak	Mer	Mer	Mar	Pet	

[63b]

$(\chi^2\text{-}b)/N$ measurements x1000 for 34 communalects
of the Northeast New Hebrides
(taken after the removal of near-empty rows and columns
from tables of possible sound correspondences calculated
on near-phonemicized lists, affixes removed)

(continued next page)

	Navenevene										
Tam	543	Tam									
Nasawa	275	297	Nasawa								
Narovorovo	376	437	445	Narovorovo							
Baetora	317	351	301	369	Baetora						
Wailengi	148	147	137	128	186	Wailengi					
Lolomatui	134	138	129	126	171	432	Lolomatui				
Ngwatua	101	126	96	102	134	173	170	Ngwatua			
Lolsiwoi	160	169	137	165	182	282	241	201	Lolsiwoi		
Raga	150	153	116	154	182	150	149	104	140	Raga	
Apma	70	66	64	45	57	74	61	59	81	97	Apma
Seke	54	73	67	38	56	51	60	69	35	47	Seke
Sa	51	59	75	41	51	44	36	52	42	36	Sa
Sowa	61	80	55	67	46	40	40	61	61	62	Sowa
	Nav	Tam	Nas	Nar	Bae	Wai	Lol	Ngw	Lol	Rag	

	Apma			
Seke	137	Seke		
Sa	120	117	Sa	
Sowa	148	266	137	Sowa
	Apm	Sek	Sa	

[63c]

$(\chi^2\text{-}b)/N$ measurements x1000 for 34 communalects
of the Northeast New Hebrides
(taken after the removal of near-empty rows and columns
from tables of possible sound correspondences calculated
on near-phonemicized lists, affixes removed)

(concluded)

	Hiw										
Toga	902	Toga									
Lehali	631	813	Lehali								
Lehalurup	660	783	932	Lehalurup							
Motlav	601	737	918	938	Motlav						
Mota	467	637	713	677	714	Mota					
Vatrata	603	758	866	883	888	714	Vatrata				
Sasar	507	728	877	882	930	650	939	Sasar			
Vetumboso	582	786	826	815	786	741	927	854	Vetumboso		
Mosina	626	762	815	839	826	733	942	853	983	Mosina	
Bek	555	711	839	886	904	704	832	863	804	798	Bek
Nume	425	625	618	629	632	841	706	621	707	712	Nume
Dorig	375	512	514	549	560	669	641	531	616	635	Dorig
Koro	409	561	571	556	602	764	666	565	703	636	Koro
Wetamut	343	475	495	547	560	761	628	538	598	594	Wetamut
Lakona	376	484	453	510	457	680	566	488	627	632	Lakona
Merig	314	463	397	450	418	661	443	360	464	453	Merig
Merlav	292	437	442	414	447	685	444	369	439	440	Merlav
Marino	-204	-79	-40	-54	43	367	-61	-104	-29	-75	Marino
Peterara	-307	-160	-90	-171	-91	304	-143	-178	-121	-99	Peterara
Navenevene	-315	-260	-163	-212	-135	246	-191	-230	-161	-142	Navenevene
Tam	-320	-235	-167	-203	-132	181	-174	-220	-125	-190	Tam
Nasawa	-348	-231	-133	-176	-103	312	-151	-220	-121	-135	Nasawa
Narovorovo	-341	-262	-167	-227	-118	177	-198	-215	-167	-141	Narovorovo
Baetora	-389	-293	-274	-248	-207	168	-254	-298	-213	-211	Baetora
Wailengi	-264	-244	-282	-305	-231	-82	-281	-286	-282	-275	Wailengi
Lolomatui	-226	-240	-287	-291	-251	-53	-286	-295	-265	-250	Lolomatui
Ngwatua	-192	-184	-174	-193	-91	146	-165	-256	-167	-121	Ngwatua
Lolsiwoi	-313	-233	-281	-274	-204	25	-263	-247	-236	-271	Lolsiwoi
Raga	-371	-307	-303	-330	-258	140	-314	-350	-260	-249	Raga
Apma	-324	-465	-368	-337	-294	-417	-302	-328	-240	-259	Apma
Seke	-90	-202	-201	-54	-111	-220	-76	-130	-8	-41	Seke
Sa	-260	-399	-244	-93	-163	-331	-142	-212	-144	-171	Sa
Sowa	-294	-336	-274	-186	-255	-324	-196	-248	-148	-150	Sowa
	Hiw	Tog	Leh	Leh	Mot	Mot	Vat	Sas	Vet	Mos	

[64a]

Linear-correlation measures x1000 from $(\chi^2-b)/N$ measurements
of 34 communalects of the Northeast New Hebrides
(taken after the removal of near-empty rows and columns
from tables of possible sound correspondences calculated
on near-phonemicized lists, affixes removed)

(continued next page)

	Bek										
Nume	600	Nume									
Dorig	550	895	Dorig								
Koro	552	877	894	Koro							
Wetamut	494	893	950	874	Wetamut						
Lakona	533	831	828	931	829	Lakona					
Merig	425	698	583	617	615	589	Merig				
Merlav	349	704	576	632	605	581	989	Merlav			
Marino	-29	163	156	169	178	186	306	319	Marino		
Peterara	-66	86	13	37	91	99	136	167	856	Peterara	
Navenevene	-142	17	-49	12	36	12	51	90	814	948	Navenevene
Tam	-130	-13	-13	0	14	84	83	83	815	934	Tam
Nasawa	-89	102	-29	43	80	43	111	148	813	888	Nasawa
Narovorovo	-122	-35	-18	13	30	57	85	96	868	910	Narovorovo
Baetora	-173	-64	-44	-47	-32	4	46	77	875	908	Baetora
Wailengi	-191	-202	-259	-251	-205	-192	-98	-87	396	360	Wailengi
Lolomatui	-138	-221	-234	-228	-212	-172	-99	-65	408	375	Lolomatui
Ngwatua	7	-57	-136	-100	-104	-83	19	69	503	478	Ngwatua
Lolsiwoi	-198	-127	-111	-120	-145	-172	-4	26	624	546	Lolsiwoi
Raga	-190	-135	-194	-143	-140	-127	2	35	737	770	Raga
Apma	-275	-405	-301	-365	-375	-208	-231	-185	-252	-203	Apma
Seke	-112	-147	-81	-94	-146	-50	-72	-67	-288	-260	Seke
Sa	-160	-360	-222	-277	-294	-216	-204	-186	-222	-184	Sa
Sowa	-210	-314	-238	-239	-277	-170	-190	-178	-340	-239	Sowa
	Bek	Num	Dor	Kor	Wet	Lak	Mer	Mer	Mar	Pet	

[64b]

Linear-correlation measures x1000 from $(\chi^2-b)/N$ measurements
of 34 communalects of the Northeast New Hebrides
(taken after the removal of near-empty rows and columns
from tables of possible sound correspondences calculated
on near-phonemicized lists, affixes removed)

(continued next page)

	Navenevene	Tam	Nasawa	Narovorovo	Baetora	Wailengi	Lolomatui	Ngwatua	Lolsiwoi	Raga	
Tam	985										
Nasawa	876	885									
Narovorovo	917	904	934								
Baetora	923	920	912	950							
Wailengi	342	327	335	369	455						
Lolomatui	348	322	336	352	468	984					
Ngwatua	466	380	476	449	542	853	835				
Lolsiwoi	498	485	511	498	645	843	882	876			
Raga	749	745	777	744	847	709	702	806	837		
Apma	-157	-137	-176	-87	-137	-66	28	48	-154	-88	Apma
Seke	-185	-228	-220	-162	-270	-241	-245	-211	-234	-159	Seke
Sa	-92	-129	-158	-50	-159	-209	-153	-146	-226	-71	Sa
Sowa	-139	-182	-132	-193	-175	-222	-193	-183	-337	-182	Sowa
	Nav	Tam	Nas	Nar	Bae	Wai	Lol	Ngw	Lol	Rag	

	Apma	Seke	Sa	Sowa
Seke	774			
Sa	806	861		
Sowa	786	826	818	
	Apm	Sek	Sa	

[64c]

Linear-correlation measures x1000 from $(\chi^2\text{-b})/N$ measurements
of 34 communalects of the Northeast New Hebrides
(taken after the removal of near-empty rows and columns
from tables of possible sound correspondences calculated
on near-phonemicized lists, affixes removed)

(concluded)

```
     74% *                                     T
     72% *                                     T   T
     70% *
     68% *
     66% *                                  T
     64% *
T    62% *
u    60% *                              T
t    58% *                      2 TTT T
u    56% *                     2
b    54% *                 T 2TTT
a    52% *                  T
     50% *              T 3
     48% *              2T
     46% *              T2
     44% *         T  T
     42% *          T
     40% *      TTT
     38% *      T
     36% *   T T
     34% *
         *.........*.........*.........*.........*.........*
        35%       45%       55%       65%       75%       85%

                              Aore
```

[65]

Thirty-nine communalects of Espiritu Santo plotted from their proportions of cognates with Aore and Tutuba

(Graph adapted from output from GLIM © Royal Statistical Society, London)

```
    0.176 *                                       T
    0.168 *
    0.160 *
    0.152 *
    0.144 *                               T
    0.136 *
T   0.128 *                       T
u   0.120 *
t   0.112 *                           T
u   0.104 *                  T   TT
b   0.096 *             T 2
a   0.088 *                 3T  T
    0.080 *           T  TT
    0.072 *           T TT  T
    0.064 *             TT
    0.056 *        T24  2
    0.048 *         T  TT
    0.040 *
    0.032 *
    0.024 *   T  T   T
    0.016 *
          *.........*.........*.........*.........*.........*
         0.0       0.05      0.10      0.15      0.20      0.25

                                 Aore
```

[66]

Thirty-nine communalects of Espiritu Santo plotted according to their $(\chi^2-b)/N$ measurements with Aore and Tutuba (computed after the removal of near-empty rows and columns from tables of possible sound correspondences calculated on unphonemicized wordlists)

(Graph adapted from output from GLIM © Royal Statistical Society, London)

```
   0.352 *
   0.336 *                                         T
   0.320 *                                     T
   0.304 *
   0.288 *                                      T
   0.272 *
   0.256 *
T  0.240 *                           T
u  0.224 *                     T    2
t  0.208 *
u  0.192 *                       2
b  0.176 *                 T    2T
a  0.160 *                   T T
   0.144 *                     T
   0.128 *            2T  2    T
   0.112 *             22T
   0.096 *           2TT T
   0.080 *         T T
   0.064 *        TTT
   0.048 *    TT
   0.032 *
         *.........*.........*.........*.........*.........*
         0        0.08      0.16      0.24      0.32      0.40

                             Aore
```

[67]

Thirty-nine communalects of Espiritu Santo plotted according to their $(\chi^2-b)/N$ measurements with Aore and Tutuba (computed before the removal of near-empty rows and columns from tables of possible sound correspondences calculated on unphonemicized wordlists)

(Graph adapted from output from GLIM © Royal Statistical Society, London)

```
         840 *
         800 *                                     T
         760 *
         720 *
         680 *                                T
         640 *
         600 *
         560 *
T        520 *                          T
u        480 *                         T
t        440 *
u        400 *           T   T2
b        360 *          T T T  T
a        320 *          2  TT T
         280 *          32TT
         240 *       TT3TT
         200 *       T
         160 *      T
         120 *    TT T
          80 *   TTT
          40 *
             *.........*.........*.........*.........*.........*
             0        250       500       750      1000      1250

                                  Aore
```

[68]

Thirty-nine communalects of Espiritu Santo plotted according to their χ^2-b measurements with Aore and Tutuba (computed after the removal of near-empty rows and columns from tables of possible sound correspondences calculated on unphonemicized wordlists)

(Graph adapted from output from GLIM © Royal Statistical Society, London)

```
      1840 *                                   T     T
      1760 *
      1680 *
      1600 *
      1520 *                                 T
      1440 *
      1360 *                        T
   T  1280 *                         T
   u  1200 *
   t  1120 *                  T     T
   u  1040 *                     T T
   b   960 *                   T
   a   880 *                    T
       800 *                016T
```

[69]

Thirty-nine communalects of Espiritu Santo plotted according to their χ^2-b measurements with Aore and Tutuba (computed before the removal of near-empty rows and columns from tables of possible sound correspondences calculated on unphonemicized wordlists)

(Graph adapted from output from GLIM © Royal Statistical Society, London)

```
    0.230 *
    0.220 *
    0.210 *                                        T
    0.200 *
    0.190 *
    0.180 *                                 T
    0.170 *
T   0.160 *                          2
u   0.150 *                             T
t   0.140 *                T     T
u   0.130 *                 2  2T  2
b   0.120 *                 T T
a   0.110 *             TT 2 T
    0.100 *           T  22
    0.091 *           32 2 T
    0.081 *           T
    0.071 *            T T
    0.061 *
    0.051 *   T   T
    0.041 *
    0.031 *
          *.........*.........*.........*.........*.........*
         0.025     0.075     0.125     0.175     0.225     0.275

                              Aore
```

[70]

Thirty-nine communalects of Espiritu Santo plotted according to their χ^2/N measurements with Aore and Tutuba (computed after the removal of near-empty rows and columns from tables of possible sound correspondences calculated on unphonemicized wordlists)

(Graph adapted from output from GLIM © Royal Statistical Society, London)

```
   0.50 *
   0.48 *
   0.46 *                                   TT
   0.44 *
   0.42 *
   0.40 *                          T
   0.38 *                              T
T  0.36 *
u  0.34 *                 T
t  0.32 *                    T
u  0.30 *                   T2
b  0.28 *               T T
a  0.26 *               TT
   0.24 *          T  2 T  T
   0.22 *         2  2  T
   0.20 *         T  TT
   0.18 *       3TT
   0.16 *      2 T
   0.14 *   T2T
   0.12 * T
   0.10 *
        *.........*.........*.........*.........*.........*
       0.1       0.2       0.3       0.4       0.5       0.6

                              Aore
```

[71]

Thirty-nine communalects of Espiritu Santo plotted according to their χ^2/N measurements with Aore and Tutuba (computed before the removal of near-empty rows and columns from tables of possible sound correspondences calculated on unphonemicized wordlists)

(Graph adapted from output from GLIM © Royal Statistical Society, London)

```
    1000 *
     950 *                                       T
     900 *
     850 *                                  T
     800 *
     750 *
     700 *                           T
T    650 *                           T
u    600 *             T
t    550 *            TT  TT
u    500 *            T  T2T T
b    450 *           2 2 2
a    400 *          T3
     350 *       T2 2 T
     300 *      T  T
     250 *     T
     200 *  TT  T T
     150 *  T
     100 *
      50 *
       0 *
         *.........*.........*.........*.........*.........*
        125       375       625       875      1125      1375

                                 Aore
```

[72]

Thirty-nine communalects of Espiritu Santo plotted according to their χ^2 measurements with Aore and Tutuba (computed after the removal of near-empty rows and columns from tables of possible sound correspondences calculated on unphonemicized wordlists)

(Graph adapted from output from GLIM © Royal Statistical Society, London)

```
   2600 *
   2500 *
   2400 *                                       T
   2300 *                            T
   2200 *
   2100 *                               T
   2000 *
   1900 *
T  1800 *                     T
u  1700 *                 T
t  1600 *               T    2
u  1500 *                 T
b  1400 *              T
a  1300 *             T
   1200 *           3 T
   1100 *      T T T
   1000 *   2T23 2T
    900 *  2 T2
    800 * 2 2T
    700 *
    600 *
        *.........*.........*.........*.........*.........*
       750      1250      1750      2250      2750      3250

                           Aore
```

[73]

Thirty-nine communalects of Espiritu Santo plotted according to their χ^2 measurements with Aore and Tutuba (computed before the removal of near-empty rows and columns from tables of possible sound correspondences calculated on unphonemicized wordlists)

(Graph adapted from output from GLIM © Royal Statistical Society, London)

1:FIVE
five cinq cinco cinque cinco
cinci fu:nf vijf fem fem
fem pie,c' pe^t pet o:t
viisi bes, lima kvin pya't
pen'te chams chamischah finef go
tano
2:TWO
two deux dos due dois
doi zwei twee tva. to
to dwa dve^ dva ketto"
kaksi iki dud du dva
di'o itsnayn schnayim tsvei ni
mbili
3:EYE
eye oeil ojo occhio o^lho
ochi auge oog o:ga o/je
o/ye oko oko oko szem
silma: go:z mata okulo glas
ma'ti ayn ayin oig me
jicho
6:FATHER
father pe're padre padre pai
tata^ vater vader fader fader
far ojciec otec otac atya
isa: baba ajah patro atye'ts
pate'ras ab av foter chichi
baba
7:DEATH
death mort muerte morte morte
moarte tod dood do:d do/d
do/d s'mierc' smrt smrt hala'l
kuolema o:lu:m mati morto smyert
tha'natos mawt mavet toit shi
kifo
8:TO EAT
eat manger comer mangiare comer
mi^nca essen eten a:ta spise
spise jes'c' ji'sti jesti eszik
syo:da: yemek makan mang^i ye'st
tro'go yaakoul achal essen kuu
la
9:MOTHER
mother me're madre madre ma~e
mama^ mutter moeder moder moder
mor matka matka majka anya
a:iti anne induk patr'ino mat
mite'ra oum em muter haha
mama

[74]

Seven words from 26 languages transliterated for processing

```
        English (1) ---------------.
        Swedish (9) -------.        !---.
        Danish (10) -------'   !---'!
     Norwegian (11) -----------'    !    !
          Dutch (8) ---------------.!    !
         German (7) -----------.   !     !
       Yiddish (24) -----------'         !
       Rumanian (6) -------.             !
         French (2) ---.   !   !         !
        Italian (4) ---'   !   !   !     !
     Portuguese (5) -------.   !   !---. !
        Spanish (3) -------'       !   ! !-
    Esperanto (19) ---------------'    !---!
    Hungarian (15) -------------------!    !
   Indonesian (18) -------------------'    !
         Czech (13) -----------.           !
        Polish (12) -----------'   !---.   !
       Russian (20) -----------.   !   !   !
Serbo-Croatian (14) -----------'       !   !
       Finnish (16) -----------.       !   !
        Arabic (22) -------.   !   !   !   !
       Swahili (26) -------'       !   !
         Greek (21) ---------------'   !
        Hebrew (23) -------------------!
       Turkish (17) -------------------!
      Japanese (25) -------------------'
```

[75]

Phylogenetic reconstruction of 26 world languages using $(\chi^2\text{-}b)/N$ measurements taken after the removal of near-empty rows and columns from tables of possible sound correspondences

	English										
French	-12	French									
Spanish	-3	59	Spanish								
Italian	7	85	97	Italian							
Portuguese	-8	54	131	85	Portuguese						
Rumanian	5	47	38	67	26	Rumanian					
German	58	-15	-15	-5	-11	-11	German				
Dutch	90	-6	6	1	-10	-3	146	Dutch			
Swedish	52	-12	-5	-1	-4	-13	33	73	Swedish		
Danish	96	-5	-9	14	-8	-14	88	102	110	Danish	
Norwegian	103	-6	-6	12	-4	-17	96	106	206	271	Norwegian
Polish	-22	-14	-8	6	-12	-20	-6	-2	-9	0	Polish
Czech	-16	-15	-5	17	0	-3	-2	16	9	-2	Czech
Serbo-Croa	-14	-5	-10	3	-10	-18	-2	3	-8	-19	Serbo-Croa
Hungarian	-12	-3	-2	0	12	-11	-16	-12	-7	-22	Hungarian
Finnish	-12	-11	6	11	20	4	-9	1	6	0	Finnish
Turkish	12	-8	0	5	-1	-8	-5	-1	4	-5	Turkish
Indonesian	-19	1	-3	-3	-2	-10	-12	3	-13	-11	Indonesian
Esperanto	20	52	35	52	13	12	-9	4	-8	-8	Esperanto
Russian	-13	-19	-15	-9	-6	-4	-10	-10	-6	-11	Russian
Greek	-13	-20	-13	-4	-12	-10	-10	-15	-11	-14	Greek
Arabic	-9	-7	-10	1	-7	-6	-11	0	-2	-15	Arabic
Hebrew	-23	-15	-12	-4	-10	-13	-14	-8	-10	-19	Hebrew
Yiddish	49	-17	-17	1	-12	-16	145	116	24	61	Yiddish
Japanese	10	-14	-16	0	-14	-14	-4	3	17	3	Japanese
Swahili	-6	-17	-1	10	-7	-8	-13	14	-15	-8	Swahili
	Eng	Fre	Spa	Ita	Por	Rum	Ger	Dut	Swe	Dan	

[76a]

(χ^2-b)/N measurements x1000 for 26 world languages
(taken after the removal of near-empty rows and columns
from tables of possible sound correspondences calculated
on unphonemicized lists)

(continued next page)

	Norwegian										
Polish	2	Polish									
Czech	3	180	Czech								
Serbo-Croa	-12	96	134	Serbo-Croa							
Hungarian	-9	-12	-4	-14	Hungarian						
Finnish	4	3	10	1	-5	Finnish					
Turkish	3	1	0	9	-4	14	Turkish				
Indonesian	-7	-12	-16	-16	-5	-5	5	Indonesian			
Esperanto	-6	11	14	1	-3	1	-1	-18	Esperanto		
Russian	-8	57	74	62	-11	-7	-1	-16	-6	Russian	
Greek	-16	-9	-7	-17	-12	-1	-10	-13	-9	-8	Greek
Arabic	-14	-2	3	-3	-5	-1	-3	-17	-6	-14	Arabic
Hebrew	-13	-9	-9	1	-20	-1	-7	-6	4	-16	Hebrew
Yiddish	49	-6	-3	0	-17	-7	-9	-11	-12	-18	Yiddish
Japanese	11	9	27	-2	-8	15	21	-15	15	-4	Japanese
Swahili	-14	-10	-12	-8	-1	-8	3	-8	-3	-9	Swahili
	Nor	Pol	Cze	Ser	Hun	Fin	Tur	Ind	Esp	Rus	

	Greek					
Arabic	-16	Arabic				
Hebrew	-7	9	Hebrew			
Yiddish	-13	-8	-8	Yiddish		
Japanese	-8	6	8	-4	Japanese	
Swahili	-14	-9	-1	-18	8	Swahili
	Gre	Ara	Heb	Yid	Jap	

[76b]

$(\chi^2\text{-b})/N$ measurements x1000 for 26 world languages
(taken after the removal of near-empty rows and columns
from tables of possible sound correspondences calculated
on unphonemicized lists)

(concluded)

	English										
French	-88	French									
Spanish	-151	887	Spanish								
Italian	-145	941	898	Italian							
Portuguese	-190	822	954	862	Portuguese						
Rumanian	-207	896	845	905	860	Rumanian					
German	863	-245	-232	-249	-288	-266	German				
Dutch	874	-273	-308	-248	-262	-288	952	Dutch			
Swedish	868	-185	-196	-169	-209	-245	714	698	Swedish		
Danish	898	-184	-166	-196	-191	-197	736	794	973	Danish	
Norwegian	855	-199	-192	-213	-216	-205	648	755	962	947	Norwegian
Polish	-207	-169	-172	-145	-128	-128	-117	-129	-66	-148	Polish
Czech	-220	-134	-158	-201	-186	-244	-118	-207	-139	-140	Czech
Serbo-Croa	-292	-223	-161	-154	-141	-123	-142	-176	-161	-162	Serbo-Croa
Hungarian	-364	547	715	562	524	596	-462	-470	-308	-222	Hungarian
Finnish	-18	450	423	259	199	177	-174	-305	2	-75	Finnish
Turkish	-86	-54	-74	-207	-39	-24	-102	-54	83	111	Turkish
Indonesian	106	272	424	332	413	400	71	-229	42	19	Indonesian
Esperanto	-260	734	649	749	740	861	-262	-305	-207	-177	Esperanto
Russian	-279	-170	-151	-126	-200	-246	-178	-179	-133	-166	Russian
Greek	-269	160	64	-213	72	85	-310	-211	-271	-243	Greek
Arabic	-228	31	86	-82	5	65	-143	-319	-260	-206	Arabic
Hebrew	-267	62	-19	-128	-100	-55	-260	-369	-302	-301	Hebrew
Yiddish	789	-229	-215	-291	-275	-212	983	936	599	659	Yiddish
Japanese	171	-364	-338	-418	-386	-325	129	161	246	252	Japanese
Swahili	26	365	193	-47	224	257	-11	-412	-62	-160	Swahili
	Eng	Fre	Spa	Ita	Por	Rum	Ger	Dut	Swe	Dan	

[77a]

Linear-correlation measures x1000 from $(\chi^2-b)/N$ measurements
of 26 world languages
(taken after the removal of near-empty rows and columns
from tables of possible sound correspondences calculated
on unphonemicized lists)

(continued next page)

	Norwegian										
Polish	-149	Polish									
Czech	-141	975	Czech								
Serbo-Croa	-187	970	939	Serbo-Croa							
Hungarian	-426	0	-174	9	Hungarian						
Finnish	-91	208	120	162	482	Finnish					
Turkish	3	144	226	27	224	309	Turkish				
Indonesian	-94	-348	-284	-249	314	254	-211	Indonesian			
Esperanto	-220	14	-1	86	407	212	112	332	Esperanto		
Russian	-164	952	959	959	-37	272	185	-330	29	Russian	
Greek	-144	170	90	349	63	403	299	-72	54	126	Greek
Arabic	-158	260	201	311	43	438	302	141	344	319	Arabic
Hebrew	-390	142	171	9	288	338	366	-177	-76	135	Hebrew
Yiddish	658	-109	-103	-130	-452	-219	-36	27	-217	-126	Yiddish
Japanese	216	395	135	441	-73	211	395	-212	-180	371	Japanese
Swahili	-79	-103	-14	-112	151	410	366	380	305	-128	Swahili
	Nor	Pol	Cze	Ser	Hun	Fin	Tur	Ind	Esp	Rus	

	Greek					
Arabic	468	Arabic				
Hebrew	170	462	Hebrew			
Yiddish	-162	-120	-293	Yiddish		
Japanese	362	463	252	108	Japanese	
Swahili	278	484	253	78	61	Swahili
	Gre	Ara	Heb	Yid	Jap	

[77b]

Linear-correlation measures x1000 from $(\chi^2-b)/N$ measurements
of 26 world languages
(taken after the removal of near-empty rows and columns
from tables of possible sound correspondences calculated
on unphonemicized lists)

(concluded)

REFERENCES

BERGMAN, Peter M.

1968 The Concise Dictionary of 26 Languages in Simultaneous Translation. New York: Bergman Publishers.

BERGSLAND, Knut and Hans VOGT

1962 On the Validity of Glottochronology. Current Anthropology 3/2:115-153.

DYEN, Isidore, A.T. JAMES and J.W.L. COLE

1967 Language Divergence and Estimated Word Retention Rate. Language 43/1:150-171.

GUY, Jacques B.M.

1978 Proto-North New Hebridean Reconstructions. Pacific Linguistics C-61:781-850.

1980 Experimental Glottochronology: Basic Methods and Results. Pacific Linguistics B-75.

1981 Bases for New Methods in Glottochronology. Paper presented at the Third International Conference on Austronesian Linguistics.

HOEL, Paul G.

1971 Introduction to Mathematical Statistics (Fourth Edition). New York: Wiley.

SPIEGEL, Murray R.

1972 Schaum's Outline of Theory and Problems of Statistics in SI Units. New York: McGraw-Hill.

TRYON, D.T.

1976 New Hebrides Languages: an Internal Classification. Pacific Linguistics C-50.

Guy, J.B.M. *Glottochronology without cognate recognition.*
B-79, viii + 142 pages. Pacific Linguistics, The Australian National University, 1981. DOI:*10.15144/PL-B79.1*