MUSÉE D'ART ET D'ARCHÉOLOGIE DE L'UNIVERSITÉ
DE MADAGASCAR

TRAVAUX ET DOCUMENTS
XV

Pierre VERIN
Conrad P. KOTTAK
Peter GORLIN

THE GLOTTOCHRONOLOGY OF MALAGASY SPEECH COMMUNITIES

Reprinted from Oceanic Linguistics, Vol VIII, N° 1

MUSÉE - 17, rue du Dr Villette, Isoraka, Tananarive
1975

Reprinted from Oceanic Linguistics, Volume VIII, No. 1

PIERRE VÉRIN
CONRAD P. KOTTAK
PETER GORLIN

The glottochronology of Malagasy speech communities[1]

1. MADAGASCAR IS THE WORLD'S FOURTH LARGEST ISLAND. Located in the western Indian Ocean, Madagascar lies only 382 kilometers from the East African mainland. Few English-speaking scholars know much about Madagascar, but the history of its peopling and settlement is subject to dispute and alternative interpretations among scholars in France and Madagascar. The primary dilemma for the Malagasy historian has to do with the fact that the Malagasy share many attributes with populations of the nearby East African coast. Many Malagasy institutions and physical types could be quite at home in Tanzania or in Kenya. The native languages of Madagascar are not African, however, but are members of the Indonesian subgroup of the Malayo-Polynesian family. Biologically the populations of Madagascar show their debt to a gene pool consisting of contributions from Africa and Indonesia. However, in contrast to the wide range of phenotypes encountered throughout the island, the languages spoken in Madagascar are relatively homogeneous. Certainly there can be no doubt that they are Indonesian in origin.

Similarities between Malagasy and langu-

ages spoken in Indonesia have been recognized since Frederick de Houtman's publication in 1603 of word lists comparing a Malagasy language (northern Betsimisaraka of the east coast) and Malayan. A few years later, Father Luis Mariano noted the existence of two very different languages in Madagascar. On the northwest coast of Madagascar, Swahili (Buki) was spoken, while throughout the rest of the island (in the interior and on the other coasts) the people spoke a language which is definitely ancestral to Malagasy speech communities of the present day. According to Mariano, the Malagasy language was peculiar to the majority of the natives and completely different from Swahili. Mariano also recognized the similarity of Malagasy to Malay, "which proves almost definitely that the first inhabitants of Madagascar came from the ports of Malacca" (Mariano, in A. and G. Grandidier 1903–13: [2] 22).

Subsequently, the work of van der Tuuk (1864) established beyond doubt the relationship between Malagasy and other Indonesian languages. Dempwolff's reconstruction of Proto-Indonesian (1934–38) makes use of Malagasy (i.e., Merina, now the national language of the Malagasy Republic, which Dempwolff, following early travelers in Madagascar, calls "Hova").

2. PLACEMENT OF MALAGASY WITHIN THE INDONESIAN SUBGROUP. While Malagasy undoubtedly belongs to the Western Indonesian subgroup (the Hesperonesian of Dyen 1965) of the Malayo-Polynesian family, the problem of ascertaining the Indonesian language most closely related to Malagasy has not been resolved. Dahl (1951) has made a convincing case for a close relationship between Malagasy and Maanyan of Southeast Borneo. Dyen (1953) has confirmed this affinity by a glottochronological comparison of Malagasy with Maanyan, Malayan, Ngaju-Dayak and Toba-Batak of Sumatra. Malagasy (the Merina speech community) was found to share 45 percent of its basic vocabulary with Maanyan, 28 percent with Malayan, 26 percent with Ngaju-Dayak, and 20 percent with Toba-Batak (Dyen 1953:589–590).

Despite the impressive evidence of Dahl and Dyen we still cannot be certain that Maanyan is, in fact, Malagasy's closest Indonesian collateral.[2] In the first place, in Dyen's glottochronological study only four other Indonesian languages were compared to Malagasy. In the second, only one of several Malagasy speech communities, that is Merina, has been considered in the works of Dempwolff and Dyen. Furthermore, most of Dahl's hypotheses are also based on evidence from Merina alone.

In the present article, we shall not attempt to deny a certain obvious

homogeneity of dialects spoken in Madagascar. However, we shall be concerned primarily with demonstrating a considerable linguistic diversity which has hitherto received little attention. One result of our argument will be to suggest that Dahl's conclusions may have to be modified in the light of the data and analyses included in this paper.

3. AIMS OF THE PRESENT STUDY. The contributions anticipated by the present study are three: (1) a reassessment of existing studies of Malagasy dialectology with the aid of comparisons in basic vocabulary; (2) a clarification of the culture history of Madagascar. When viewed in combination with the findings of archaeology and ethnology, a subgrouping of Malagasy speech communities will be of significant aid to the scholar and layman interested in the origins and socio-cultural diversification of the Malagasy population; (3) a presentation of material which will assist in placing Malagasy dialects within the framework of the Indonesian subgroup.

Let us state immediately that this paper is an application of glottochronology to problems of linguistic and cultural unity and diversity within Madagascar. The authors include one archaeologist and two social anthropologists interested in the historical and present-day interrelationships of the different populations who inhabit Madagascar, and particularly in the processes whereby these socio-cultural groups have adapted to their different island ecological niches. In other words, our general aim is the elucidation of the genetic relationships which provide some of the bases for elements of cultural unity and diversity observed in Madagascar today.

It is for this purpose that we have chosen glottochronology as a major analytic tool. Although we are not primarily linguists, we are well aware of the many criticisms to which glottochronology has been subjected in the past decade, and some caveats will be discussed below. We shall also indicate in greater detail the reasons why we have chosen to use this much debated technique. For the present, however, let us say simply that we are following Gudschinsky's invitation (1956:622) to try a new linguistic tool of value to students of the past. This, in other words, is not primarily a challenge to, but an application of glottochronology.

4. DIFFERENCES AMONG MALAGASY SPEECH COMMUNITIES AND CRITERIA FOR SELECTION OF "DIALECTS."[3] Speech communities in Madagascar are assumed here to coincide in large measure with cultural groups. However, there is no agreement about the number of distinct culture

groups in Madagascar. The official census classification of the Malagasy government recognizes the existence of twenty *ethnies* ("cultures").[4] These units represent a tremendous range in population size and pre-colonial political complexity. On the one hand, the 1,570,000 Merina who inhabit the northern central highlands were able to extend their political domination from Tananarive, the current national capital, to a large part of the island before French occupation in 1896. At the opposite pole, one finds the loosely organized confederation of clans which constitute the 22,300 Tambahoaka of the Mananjary region of the southeastern coast.[5]

Basic vocabulary lists were not collected for all of the hitherto recognized *ethnies* or cultures of Madagascar. Our aim in choosing certain speech communities to be compared through glottochronology was that of trying to obtain a geographical distribution which would accurately reflect existing linguistic and other kinds of cultural differences and similarities. For this reason, some traditionally recognized dialects, for example, Betsileo, have been treated twice, in accordance with what seemed to us to be significant differences in speech patterns within the total area inhabited by people who call themselves Betsileo (the southern highlands). The "northern Betsileo" of the Ambositra region speak a dialect whose basic vocabulary is more similar to the Merina of Tananarive, the capital, than to the Betsileo spoken to the south around Fianarantsoa, the capital of the traditional Betsileo homeland, or Anabalavao, in the extreme south of the Betsileo territory.

5 ORIGIN OF DATA AND INFORMANTS. The authors deliberately decided to obtain data from informants and to avoid the use of dictionaries. Vocabularies were collected with the aid of a bilingual list composed of French and Merina terms. For words such as *fat* (no. 32), it was necessary to specify whether the referent was cooked or uncooked fat, since most dialects involved make this distinction. Since there are several words used for *claw* throughout Madagascar, we specified that the meaning was to pertain to cats or birds of prey. Other problems concerning the translation are discussed in detail below under each meaning.

Many word lists were collected by the authors themselves, but others were sent to missionaries and schoolteachers.[6]

The map in Figure 1 shows the origin of the informants and geographical placement of the word lists. Our abbreviations are those devised by Molet (1953) which are currently well accepted to refer to Malagasy dialects and socio-cultural groups.[7] Listed below are the

eighteen idiolects considered in this study, along with informants' names, data-gatherers for each idiolect, provenance of dialect speaker, and place where the glottochronological lists were collected. The map in Figure 1 locates the home areas of the speech communities represented by the word lists.

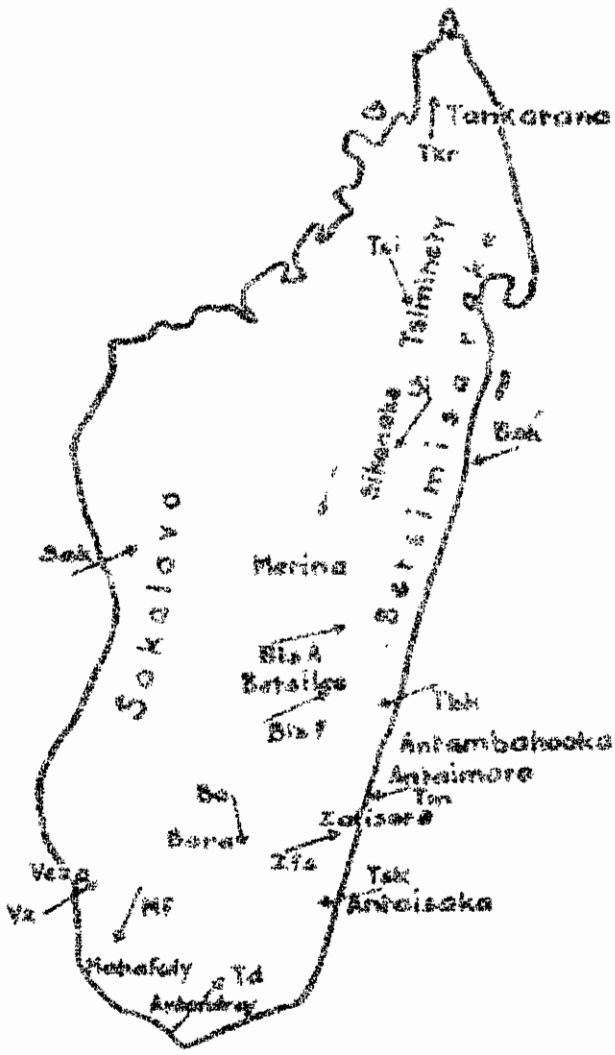


FIGURE 1: Geographical distribution of Malagasy dialects.

Data used by scholars and for reconstruction of word lists.

1. *Merina* (Mr): crosschecked in Tananarive with several informants. Had the words been extracted from nineteenth-century dictionaries, a closer relationship to Sihanaka (Si) and Betsileo of Ambositra (BtsA) would have been found to obtain.
2. *Tsimihety* (Tsi): list prepared by Frederick Kelly from Vohipeno for F. Nassey, crosschecked with list obtained by Joseph Rabe from a Tsimihety migrant to Ambalavao.
3. *Betsimisaraka* (Bsk): gathered by J. T. Hardyman from three students from the Vavatenina area in residence at Imerimandroso: Randrianasivo from Anjahambe, Toto Michel from Voloboangy, and Ratsiatsika from Ambohibe.
4. *Sihanaka* (Si): gathered by J. T. Hardyman from three illiterate peasants from villages of the Imerimandroso area: Rakotozandry and Totomanahirana from Ambohitrampirano and Rakavavy from Ambohazarimasina.
5. *Betsileo, Ambositra* (BtsA): established by Vérin with information from Jeanette Razafindravao from Andina near Ambositra, and crosschecked with informants from the same area.
6. *Betsileo, Fianarantsoa* (BtsF): two lists have been made separately, one by Kottak with several informants, the other by Vérin with information provided by Philibert Rakotozafy. This allowed a careful crosscheck.
7. *Antambahoaka* (Tbk): list written by Mlle. Celestine from Vohimasina, studying in Mananjary. Transmitted by F. du Noyer and Rakotondrazaka.
8. *Antaisaka* (Tsk): list obtained from local informants by F. Jourdan and Norbert Rolmmary.
9. *Zafisoro* (Zfs): collected in Ambalavao by Kottak from Jean Marcel
10. *Tsimihety* (Tsi): written by Mosesy Rajaonadaka from Mandritsara.
11. *Antaimoro* (Tsr): collected in Tananarive by Vérin with Marie-Therese Vahiloly from Ivory (Ambilobe). For a few words the informant checked with her relatives.
12. *Vezo* (Vz): collected by Vérin from Metalahy in the village of Anakao. Crosschecked on the spot with other informants.
13. *Masikoro* (Mi): collected by Vérin from Mlle. Elias in Ampanihy.
14. and 15. *Sakalava 1 and 2* (Sak): two lists used. One provided by MM. Flavien Bora, attaché de Cabinet au Ministère des Affaires culturelles, and Joseph Theminuche (Morondava). The other collected in

Ambalavao by Kottak from Justin Michafaly from Belavaro-Tsaribihina (north of Morondava).

16. *Bara* (Ba): collected by Kottak from Tsirongaty from Antsoabohoke-Fiarena, sous-préfecture Ivohibe. This is the Eastern Bara dialect. Data from the Ranobira area (west) would have probably reflected more resemblance with Vezo and Mahafaly.

17 and 18. *Antandroy 1 and 2* (Td): several lists collected by Kottak, including one with Falika dit Masinore in Ambalavao. Informants from Antanimoro area.

6. PHONOLOGY AND ORTHOGRAPHY. Although we are aware that a great deal of work remains to be done on Malagasy phonology, we have been obliged to take into account the writing conventions currently in use in Madagascar. Merina dialect, now the official language of the Malagasy Republic, has been written since 1820. We have therefore employed the official orthography of the Merina dialect, making use of certain modifications necessary to describe "provincial" speech patterns. Since it was not always possible for the authors to reach the areas included in this study, the use of the official orthography enabled us to obtain lists for some speech communities through correspondence with literate informants.

Dez, following Faublée, gives the following paradigm of consonant phonemes of Merina transcribed with the official orthography (see Table 1).

With the exception of Merina (Me), all Malagasy dialects have, in addition to most of the phonemes indicated in Table 1, a velar or sometimes palatal nasal consonant [ŋ], which, since the Malagasy governmental decree no. 62-404 of August, 1962, must be transcribed

TABLE 1. Malagasy consonant phonemes, Merina dialect

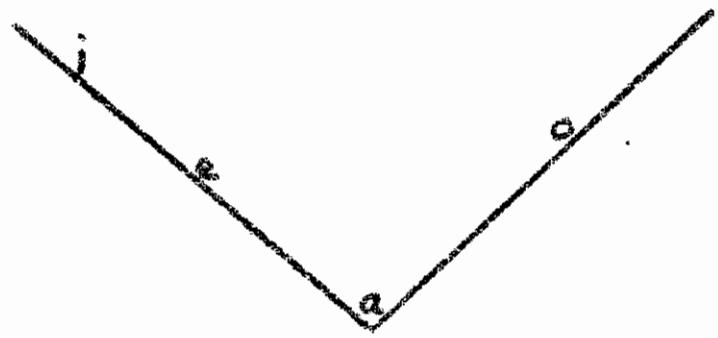
	Velar	Palatal	Alveolar	Dental	Labial
Voiced					
Stops	g	j	dr	d	b
Nasals	ng	nj	ndr	nd	mb
Continuants		z	z	l	v
Voiceless					
Stops	k	s	tr	t	p
Nasals	nk	nts	ntr	n	mp
Continuants	h				f

(Source: Dez 1965:511.)

as *ŋ* in Madagascar. This phone attains phonemic status in most Malagasy speech communities, where /n/ contrasts with /ŋ/ in minimal pairs, for example, *araŋa* (prawn) and *araŋa* (rain). Except for this additional phoneme, the consonant phoneme systems of Merina and all other Malagasy speech communities are uniform. Merina consonant phonemes have regular correspondences in other Malagasy speech communities: for example, [s] in Me is equivalent to [š] in Betsimisaraka (Bsk) and other areas, [ts] in Me is equivalent to [t͡š] in Betsileo (Bts), Tanala (Ta), and Bara (Ba). However, there is never, within any Malagasy speech community, a contrast of /ts/ and /t͡š/ or /s/ and /š/. These slight phonological differences among Malagasy speech communities therefore may be neglected as long as we are dealing only with lexical comparisons.

Semiconsonant [i] in other dialects often corresponds to [z] in Me. It has been an accepted convention since Malagasy orthography was established in 1820 to write final [i] as *y*. Following the orthography, it is standard for literate Malagasy to write *z* initially and medially, and *y* terminally, to represent approximately similar sounds which are, in fact, allophones of the same phoneme, and allographs of the same grapheme. We have respected this writing convention except when final [i] is replaced by [z] in dialects other than Me.

Dez (1963:515) summarizes the vowel system of Merina as shown in the diagram below. The vowel *o* is pronounced most commonly as [u] (English zoo), but may actually be a closed [o*] in northern Betsimisaraka (Bsk) and Antankarana (Tk). Following the decree of 1962 mentioned above, it is represented by *o* when these northern dialects are written.



7. GLOTTOCHRONOLOGICAL PROCEDURE. In collecting basic vocabulary for Malagasy speech communities, we have used Swadesh's 100-item

test list of diagnostic terms (reproduced in Hymes 1960: 6). In accordance with the use of this list, we have assumed a retention rate of 86 percent per thousand years (Hymes 1960: 9).

The 100 meanings have been collected for sixteen Malagasy speech communities (see list above, section 5). Two additional lists were constructed taking into account the existence of dual vocabularies in the Sakalava and Tandroy dialects. Dual vocabularies are commonly encountered in western and southern Madagascar. The dual aspect of the vocabulary consists of the use of different terms for a single meaning in accordance with two conditions. First, special terms are used to reflect relative placement in a social hierarchy. For example, in the Tandroy dialect, the normal Malagasy word *loha*, for head, is replaced by *ambone* wherever reference is made to the head of a chief, member of a ruling lineage, or elder. Special terms are particularly numerous when reference is being made to body parts and functions. While the use of these terms seems to be diminishing, some special vocabulary words have entered everyday parlance, now devoid of much of its former hierarchical framework. In all cases where there are special terms correlated with high status, there are alternatives for reference to commoners. Had special terms alone been considered in the selection of forms for the basic vocabulary lists, the estimate of divergence time on the basis of glottochronology would have been too great. Even if coins had been tossed to decide between common and status terms, there would have been a skewing toward a greater time depth. This suggests a real problem for glottochronology. How does one handle special terms for which commonly used alternatives are available? Both sets of terms are integrally involved in the Malagasy verbal perception of hierarchical social relations. Our solution was to include two lists for the dialects where such distinctions exist. Since the common terms are also current, we decided to use only common terms in our estimates of divergence times.

A second avenue for the introduction of special terms into Malagasy vocabulary in certain speech communities has to do with taboo and replacement. For example, when a powerful chief or noble dies, a common term which happened to form part of his name (Malagasy names are usually composite and descriptive) may henceforth be interdicted, and another term is used to refer to the same meaning. For example, when a chief called "the man who lived on the hill" dies, the word for *hill* may be tabooed and another expression proclaimed officially to be used to refer to that meaning. It is very probable that the effect of taboo on the use of words which have been part of names of deceased chiefs plays a great role in word change in the western, Sakalava area of Madagascar. This is a result of the fact that there have

been highly evolved and culturally complex kin terms in this region since the eighteenth century. Taboo is more difficult to control than dual vocabulary, and we are unable to say to what extent it has affected the results of our study. However, since there were survivals of the tabooed forms in some areas, it was also possible for us to reconstruct ten lists for the Sakalava dialect. Time estimates were then based on the list of original forms. Like other culture elements, taboo words spread from centers. Like other ideological strata to state systems, taboos are most forceful at the center or capital of the state administration. Some areas, which were under only the nominal political control of the Sakalava sovereigns, no doubt adopted taboos to the same extent as communities nearer the geographical locus of kingly power.

It is perhaps obvious to add, but necessary to the understanding of Malagasy dialectology, that neither status vocabulary nor taboo is apt to be associated with languages of societies which are not socioeconomically stratified. The authors suggest that stratification is a necessary, but, as is confirmed by the absence of special vocabulary and taboo in the dialect of the Merina empire, insufficient condition for the existence of status vocabulary and taboo terms.

Lists were presented to informants with meanings given in French and Merina (national Malagasy). Informants were asked for the most common term in their speech communities for a given meaning. In those cases in which informants were unwilling or unable to make judgments about relative frequency of use, or argued that two or more words were used with equal frequency for the same meaning, the authors have tossed coins to choose the single term which enters the list used for comparison.

Insofar as possible, we have followed Gunschinsky's instruction (1956:615) to register as noncognate the forms which are similar because one language has borrowed them from the other or because both have borrowed from a common source. In many instances it is impossible to say whether cognates have been borrowed by one Malagasy dialect from another. In an environment in which contacts between different communities took place through trade and warfare, there must have been word borrowing. One can only assume, if there is no discordant phonological evidence to the contrary, that basic cognates are shared because of common inheritance rather than as loan words. Because of the relative homogeneity of speech patterns throughout Madagascar, we could not be sure that all internal borrowings have been eliminated. It was possible, however, for us to discard those terms which can be demonstrated to have been borrowed from Swahili, for example, certain words for *fish*, *day* and *house*.[4]

It was inevitable that some of the meanings on the 100-word list would present problems to Malagasy informants. The specific problems encountered in trying to elicit Malagasy equivalents for the 100 meanings are discussed in the following section, which is a detailed examination of the data: the different Malagasy expressions for the meanings of the basic vocabulary list.

8. THE WORD LISTS. This section discusses the forms encountered in the eighteen lists according to meaning on Swadesh's revised (100 word) basic vocabulary list. Problems associated with each meaning, including the criteria employed for judging cognation, are also presented to the reader, who may, if he questions our interpretation, make his own judgments using the data we have gathered.

Under each meaning entry below we shall indicate the equivalents in the various Malagasy dialects. Malagasy words will be broken up into subgroups, and justifications for these subgroups will be provided.

1. 'I': **A**: *aho* (Me, BtsA, BtsF, Ba), *zaho* (Bsk, Tkr, Mf, Sak_1, Sak_2, Td_1, Td_2), *izaho* (Tm, Si, Tbk, Tsi, Vz), *iaho* (Tsk, Zfs).
Me also has *iz,aho*, but it is used for emphasis. There is a tendency for *y* (*ia*) to change toward *z* (*iza*) in many dialects (Dez 1963: 5[9]).

2. 'you': **A**: *hianao* (Me), *anao* (Tm, Tsk, Si, BtsA, Tbk, Tsi, Tkr), *anao* (BtsF, Zfs), *ianao* (Tsk, Td_1, Td_2), *hanao* (Ba); **B**: *riha* (Vz, Mf), *iha* (Sak_1, Sak_2).
hika is also known inland as a very colloquial form among the southern Betsileo. Initial *h* in Me *hianao* is hardly heard and sometimes omitted.

3. 'we': The inclusive form has been selected. **A**: *isika* (Me, BtsF), *itsika* (Tm, Si, BtsA), *antsika* (Bsk), *sika* (Tsk), *atsika* (Zfs, Tsi), *atika* (Tkr, Vz), *tika* (Mf), *tsika* (Sak_1, Sak_2, Ba), *itika* (Td_1, Td_2); **B**: *alsen-a* (Tbk).

4. 'this': All Malagasy dialects have a large set of demonstratives whose use depends on distance and effective visibility. The following connote maximum proximity for a visible object. **A**: *ity* (Me, Si, BtsA, Tbk, Mf, Td_1, Td_2), *itiky* (Tm, Tsk), *ty* (Zfs), *itsto* (BtsF), *tō* (Bsk), *itoy* (Ba), *toy* (Tsi, Sak_1, Sak_2); **B**: *io* (Tkr, Vz).
Iroa was given as equivalent in Vz to *io*, and to Mf *ity*. Tossing a coin was necessary for choice. Me and many others also have *io*, but informants agreed that *ity* connoted the closest object. *Ito* is known in Me as obsolete.

5. 'that': The following connote maximum distance for a visible object. A: *izy* (Me, Tm, Si, Tbk), *ty* (BtsA, Vz, Tkr), *itsa* (Bsk), *tsay* (BtsF, Vz, Mf, Td_1, Td_2), *ray* (Sak$_1$, Sak$_2$, Ba), *iriky* (Tsk), *roá* (Tsi).

6. 'who': A: *ia* (BtsF, Vz, Mf, Sak$_1$, Sak$_2$, Ba, Td_1, Td_2), *iza* (Me, Tm, Si, BtsA, Tbk, Zfs), *za* (Tsk), *azay* (Tkr), *zay* (Bsk, Tsi).
The merging of all these words in the same group is justified by the etymology of *zay*, which according to Ferrand (1903:208 n.) is derived from *Iza ho avy* (or *Iza avy*) 'who is coming?' For the correspondence *iz/y* see word 1 and also Ferrand (*Equivalence du z merina avec i,j provinciaux*; Ferrand 1903:21).

7. 'what': A: *inona* (Me, Si, BtsA), *ino* (Tm, Bsk, BtsF, Tbk, Tsk, Zfs, Tsi, Vz, Mf, Sak$_1$, Sak$_2$, Ba, Td_1, Td_2), *zino* (Tkr).
Final *na* disappears in western and southern dialects (Dez 1963: 591).

8. 'not': A: *tsy* (Me, Tm, Bsk, Si, BtsA, Tbk, Tsk, Zfs, Tsi, Vz, Mf, Sak$_1$, Sak$_2$, Ba), *tsy* (Td_1, Td_2), *tsa* (BtsF), *tsia* (Tkr).

9. 'all': A: *rehetra* (Me, Si, BtsA); B: *aby* (BtsF, Vz, Ba), *izy aby* (Tsi, Tsk, Zfs), *ziaby* (Tsi), *iaby* (Mf, Sak$_1$, Sak$_2$, Td_1, Td_2), *jaby* (Bsk), *diaby* (Tkr); C: *izy marobe* (Tbk).
Izy is a personal pronoun of the plural (third person). *Aby* is known in the nineteenth-century dictionaries of Merina. In Merina *izy maybe* would mean 'many of them'.

10. 'many': A: *maro* (Zfs, Tkr, Vz, Mf, Sak$_1$, Sak$_2$, Ba, Td_1, Td_2), *marobe* (Tsk), *marobetsaka* (Tm); B: *fatratra* (Tbk), *fontsy* (Tsi); C: *be* (BtsF), *betsaka* (Me, Bsk, Si, BtsA).
Maro is widely known in Merina but *betsaka* has a more general meaning and use. Similarly, southern Betsileo also has *rehana* but *be* seems more common. *T/on* is an infix in *fontsy*. It can be reasonably put together with *fotratra*, which may contain a reduplication. Corresponding to final *tra* non-Merina dialects often have a more variable vowel at the end of *tr* or another consonant (Dez 1963: 593). *Feotcha* is also known in Merina to connote the strength of a feeling. The case of *marobetsaka* was difficult, for the word could be either in the A group or in the C group. Tossing a coin was necessary.

11. 'one': A: *iray* (Me), *ray* (BtsF, Zfs), *iraike* (Bsk, Si), *raika* (Tm, BtsA, Tbk, Td_1, Td_2), *raiky* (Tsk, Ba), *raike* (Vz, Mf, Sak$_1$, Sak$_2$), *araiky* (Tsi), *araika* (Tkr).

12. 'two': **A**: *roa* (Me, Tm, Si, BtsA, BtsF), *roay* (Bsk), *ray* (Tbk, Tsk, Zfs, Ba), *aroy* (Tsi, Tkr), *roe* (Vz, Mf, Sak_1, Sak_2, Td_1, Td_2).

13. 'big': **A**: *lehibe* (Me, BtsA, BtsF, Tbk), *lihibe* (Ba), *be* (Tkr, Vz, Td_1, Td_2), *zakobe* (Tm), *bevata* (Mf, Sak_1, Sak_2); **B**: *maventy* (Bsk), *vaventy* (Tsk); **C**: *Ngeda* (Si), *zeda* (Tsi); **D**: *makadiry* (Zfs).
The problem here is to select what is the most general or common term. In addition to *lehibe*, Merina has *vaventy*, *ngeza*, and *makadiry*.

14. 'long': **A**: *lava* (all dialects).

15. 'small': **A**: *kely* (Me, Tm, Si, BtsA, BtsF, Vz, Mf), *kely* (Bsk, Tsi, Tkr), *kele* (Sak_1, Sak_2), *krde* (Td_1, Td_2), *kedy* (Ba), *kidy* (Tsk, Zfs), **B**: *bitaka* (Tbk).
Madinika is also used in the south and east and in the uplands. Merina has it as well as *hitika* 'tiny', but *kely* has a more general meaning.
For *k/h* changes see Ferrand (1903:17); for *l/d* changes see Dez (1963:589); voiceless *e* instead of final *y* is frequently encountered in the west and southwest.

16. 'woman': **A**: *vehivavy* (Me, BtsA), *viavy* (Tm, Bsk, Tbk, Tsk, Zfs, Tkr), *vaivay* (Si, Tsi); **B**: *ampela* (Vz, Mf, Sak_1, Sak_2, Td_1, Td_2), *epela* (BtsF, Ba).
This is one of the most clear-cut double distributions between east and west.

17. 'man': **A**: *lehilahy* (Me, BtsA, BtsF, Tkr, Vz, Mf, Ba), *leilahy* (Si), *lelahy* (Tm, Tsk, Zfs), *lalahy* (Bsk, Tsi), *lilahy* (Tbk), *lahilahy* (Td_1, Td_2), *lahy* (Sak_1, Sak_2).
This set displays an astounding variation in the first part of a word made with the reduplicated root *lahy* (male). *Vavy* denotes female in all dialects.

18. 'person': **A**: *olona* (Me, Bsk, Si, Tbk), *olo* (Tsi), *olo* (Tm, BtsF, Tsk, Zfs, Tkr, Sak_1, Sak_2, Ba); **B**: *ndaty* (Vz, Mf, Td_1, Td_2).
It is possible that a velar *n* in *olona* was not heard or written in several cases. On changes of *na* see Dez (1963:591–592); especially for the northern dialects "la vocalisation de la finale dépend du timbre de la voyelle qui la précède immédiatement" (p. 591), e.g., Tsi *orollo* (nose). Also for the disappearance of *na* "Les finales nasales de l'Indonésien Commun ont disparu dans l'Ouest et le Sud" (p. 592).

19. 'fish': **A**: *hazandrano* (Me, BtsF); **B**: *trondro* (BtsA); **C**: *laoka* (Bsk, Tbk, Tkr), *lako* (Tsi), *laoka-drano* (Tm, Si); **D**: *kena* (Zfs); **E**: *fia* (Tsk, Vz, Mf, Sak_1, Sak_2, Ba, Td_1, Td_2).
On the distribution of the words for fish see Hebert (1961). Me has both *hazandrano* and *trondro*; Antaisaka has *fia* and *nily*. Tossing up was therefore necessary. In Imerina *laoka* connoted the food to accompany rice; *kena*, cognate to Zafisoro *kena*, means 'meat'. Sihanaka *laoka-drano* reflects an intermediary stage of this evolution toward a less restricted meaning. The loss of the original terms in the interior is so complete that *hazandrano* is made of the word *hara*, general term for game, and *rano* 'water'. Originally *trondro* designated a particular species of fish. The word *lako* in Tsi may come from a metathesis of *laoka* (*ao* is a short diphthong) or of *loka* (see Tkr and Bsk, the closest dialects).

20. 'bird': **A**: *vorona* (Me, Bsk, Si, BtsA), *voroña* (Tm, BtsF, Tbk, Tkr), *voroño* (Tsi), *voro* (Tsk, Zfs, Vz, Mf, Sak_1, Sak_2, Ba, Td_1, Td_2).
For discussion of the evolution from Proto-Indonesian of the nasalized final see Dez and word *ekuna*. We find here the four possibilities which Dez indicates (1965:592).

21. 'dog': **A**: *alika* (Me, Si, Sak_1, Sak_2); **B**: *amboa* (Tm, BtsA, BtsF, Tsk, Zfs, Tkr, Vz, Mf, Ba, Td_1, Td_2); **C**: *kiva* (Bsk, Tbk); **D**: *fandroaka* (Tsi).
Amboa is known among the Merina in certain cases. Since it has a derogatory meaning it can, for instance, be used to designate beautiful babies without any fear of harming them by attracting unnecessary attention from malevolent spirits. Tsi *fandroaka*, which literally means 'which is usually rejected', obviously comes from a taboo rule. *Amboa* is definitely borrowed from Bantu.

22. *Louse*: **A**: *hao* (all dialects except Tsi which has *ó*).
Under (19) we have already noted this tendency of Tsimihety to contract the diphthong *ao* into *o*.

23. 'tree': **A**: *hazo* (Me, Si, BtsA, BtsF, Tsk, Sak_1, Sak_2, Ba), *kakazo* (Tm, Bsk, Tbk, Zfs, Tsi, Tkr), *fotokazo* (Mf); **B**: *hatá* (Vz), *hatay* (Td_1, Td_2).
For the equivalence *h/k* see Ferrand (1903: 17). But in *kakazo* there is a short reduplication and in *fotokazo* a compound (*fototra*: trunk of a tree).

24. 'seed': **A**: *voa* (Me, Tm, Si, BtsA, BtsF, Tbk, Tsk, Tsi, Vz, Mf, Sak_1, Sak_2), *voañ* (Zfs), *oñy* (Tkr); **B**: *nily* (Bsk, Ba, Td_1, Td_2).
Voa is also known in Tandroy but was eliminated by tossing up.

25. 'leaf': **A**: *ravina* (Me, **Bsk**, Si, BtsA, Tbk, Vz), *ravina* (Tm, **BtsF**), *ravany* (Tsi), *ravy* (Tsk, Zfs, Mf, Ba), *rave* (Sak_1, Sak_2, Td_1, Td_2), *ravin-kazo* (Tkr).
Within a single group of cognates we get the small differences which we noted for *vorona* (20). The final *i* in Tsimihety follows the vowel of the preceding syllable. Final *i* in some western dialects is replaced by voiceless *e* (*i inaccentué*). The Tkr word adds *kazo* (tree.)

26. 'root': **A**: *fakany* (Me, Si); **B**: *vahetra* (Tm, **Bsk**, BtsA, Tsk, Tsi, Tkr), *vahany* (Tbk, Ba), *vahan* (Td_1, Td_2), *baha* (Mf), *vahakazo* (BtsF), *vahanikala* (Vz), *vahatse* (Sak_1, Sak_2); **C**: *vodi-kazo* (Zfs).
It should be possible to lump together A and B. But Dez after Dempwolff (personal communication) thinks they should be kept separate, for they derive from two Indonesian forms. We think there was an influence of A on B, for possibilities other than *tra* are "*tr, tch, ts, tz*, suivis d'une vocalisation également variable" (Dez 1963:593), not *na* or *ny*. As in the preceding and the following cases (BtsF, Vz), tree is also added in the Zafisoro word, the first part of which means 'the base of the tree'.

27. 'bark': **A**: *hodi-kazo* (Me, Tm, Bsk, Si, BtsF, Mf, Ba), *hody* (Tsk), *hodina* (Zfs), *hoditra-kazo* (BtsA), *hoditra-kakazo* (Tbk), *hodi-kakazo* (Tsi), *hoditrin-kakazo* (Tkr), *holitsa-taha* (Vz), *holitsazo* (Td_1, Td_2), *holi-kazo* (Sak_1, Sak_2).
Hoditra is standard for skin throughout Madagascar (see 28). Therefore, adding tree was necessary here to avoid confusion. Tsk and Zfs can dispense with this addition because they have devised a new word for skin. The formation of compound words is well known in Me but very imperfectly known in dialect. Often one gets 'skin of the tree' and sometimes 'tree-skin'. Me also has *hoditry hazo*; an abridged form—*hoditrazo*—is also known in Bsk. For equivalences *l/d*, *tra/ts*, and other consonants see Dez (1963:589, 593) and Ferrand (1903:15 ff., equivalences du Merina).

28. 'skin': **A**: *hoditra* (Me, Tm, Bsk, Si, BtsA, BtsF, Tbk), *hoditry* (Tsi, Tkr), *hoditse* (Ba), *holitse* (Vz, Mf), *holitshe* (Td_1, Td_2); **B**: *daro* (Tsk, Zfs, Sak_1, Sak_2).
For equivalences of *tra* see references under (27).

29. 'flesh': **A**: *nofo* (Me, Si, BtsA, Tsk, Ba), *nofotry* (Tm), *nofotro* (BtsF, Tsi, Tkr), *nofotra* (Zfs), *minofo* (Bsk), *nofotse* (Vz, Mf, Sak_1, Sak_2), *nofony* (Td_1, Td_2); **B**: *isy* (Tbk).

Although there is no doubt that the words sorted into A are cognates, there is an extremely wide range of finals.

30. 'blood': **A**: *ra* (Me, Tm, Bsk, Si, BtsA, BtsF, Tbk, Tsk, Zfs, Vz, Ba); **B**: *ha* (Tsi, Tkr, Mf, Sak_1, Sak_2, Td_1, Td_2).
One of the most interesting distributions between east and west, with the curious exception of Vezo.

31. 'bone': **A**: *taolana* (Me, Bsk, Si, BtsA, Vz, Td_1, Td_2), *taolaña* (BtsF, Tkr), *tdlaña* (Tm, Tbk), *tóla* (Zfs, Ba), *taolo* (Tsk, Mf, Sak_1, Sak_2), *tabolana* (Tsi).
Variation *b/ao* is often encountered, but insertion of *h* between *a* and *o* in Tsi is more curious. Final nasal in Proto-Indonesian evolves into *ña/na* or disappears.

32. 'grease': **A**: *menaka* (Me, Bsk, Si, BtsA, BtsF, Tbk, Zfs, Vz, Mf), *menaky* (Tm, Tsk); **B**: *soñka* (Tkr), *soñky* (Tsi, Sak_1, Sak_2), *soñky* (Ba), *tsoñky* (Td_1, Td_2).
Another east/west distribution, again with Vezo as an exception. Final *ka* has equivalences of the type *k* + variable vowel. *h/di* in Ba can be expected, but Merina has *soñka* (Huile de pied de boeuf; Malzac 1880:579) and *sodika* (action d'enlever la graisse; ibid., p. 572). For *s/ts* see Ferrand (1903:20).

33. 'egg': **A**: *atody* (Me, Tm, Bsk, Si, BtsA, BtsF, Tbk, Tsk, Zfs, Tsi, Ba), *antady* (Tkr), *atoly* (Vz, Mf, Sak_1, Sak_2, Td_1, Td_2).
A very homogeneous word. The change *d/l* shows the distribution described by Dez: "Le groupe des dialectes qui ont conservé les syllabes l et n de l'Indonésien Commun comprend: l'antankarana, le sakalava du Boina et du Menabe, le bara, le vezo, le mahafaly, l'antandroy, c'est à dire essentiellement les dialectes de l'Ouest et du Sud . . ." (1963:580); except that for us Bara and Antankarana are included in the east group, probably because our informants come from the eastern part of these areas. See also Ferrand *oly/ody* (1903:16).

34. 'horn': **A**: *tandroka* (Me, Tm, Bsk, Si, BtsA, BtsF, Tbk, Zfs), *tandroky* (Tsk), *tanaroka* (Tkr); **B**: *amponda* (Tsi); **C**: *tsifa* (Vz, Mf, Sak_1, Sak_2, Ba), *tsifany* (Td_1, Td_2).
Interesting distribution between east and west with a single exception (Tsi). The final vowel of the last syllable of A shifts only in Tkr and Tsk. The correspondence *k/ky* is also noted for other dialects from the south and east. Is it due to the influence of Merina *r*

35. 'tail': **A**: *rambe* (Me, Si, BtsA, BtsF, Zfs, Sak_1, Sak_2), *rambona* (Tm, Tbk_1, *rambohany* (Tsk); **B**: *ohy* (Bsk, Ts_1, Tkr, Mf, Ba), *ohina* (Vz), *ohiny* (Td_1, Td_2).
Sak_1, Sak_2, and Tsk (hohany) had both A and B. It was necessary to toss a coin, which resulted in all three being put in the A group. We can therefore consider that A is a word typical of the center and the south and east and B of the rest of the island.

36. 'feather': **A**. *volomborona* (Me, Si, Tsk), *volomborota* (Tm, BtsF, Tbk, Ts_1), *volomboro* (Mf, Sak_1, Sak_2, Ba, Td_1, Td_2), *volovolony* (Bsk), *volovolo* (Tkr), *volona* (Vz), *volon'ny vorona* (BtsA), *volo* + name of bird (Zfs).
For discussion of the second part of the word (bird) see 20. Note the contrast in Vz *volona* (36) and *volo* (37). Final *na* has disappeared in most cases (except Vz and Bsk) or is transformed in the compounding process. For BtsA there is the same problem of compounding as occurred in 27. Bsk and Tkr display a long reduplication form.

37. 'hair': **A**: *volo* (Me, Tm, Tsk, Vz), *volondoha* (Si, BtsA, BtsF, Zfs, Mf, Sak_1, Sak_2, Ba, Td_1, Td_2), *vorondoha* (Bsk); **B**: *randrana* (Tbk); **C**: *maramaraña* (Ts_1); **D**: *jañona* (Tkr).
The addition of *loha* (head) is optional in all cases. As in English, the word in Malagasy for hair applies to the head and to the body. Merina or Taimoro can also specify *volondoha*. *Randrana* has in Merina and elsewhere the meaning of plait of hair.

38. 'head': **A**: *loha* (Me, Tm, Bsk, Si, BtsA, BtsF, Tbk, Tsk, Zfs, Tkr, Vz, Mf, Td_1), *talondoho* (Ts_1); **B**: *kabeso* (Sak_2, Ba); **C**: *fañanteke* (Sak_1); **D**: *ambone* (Td_2).
This is a choice word for dialects with dual vocabulary. *Ambone* means what is above. In Bara, where a dual vocabulary also existed (though to a lesser extent), *loha* is known as an insulting word (used for slaves in the past).

39. 'ear': **A**: *sofina* (Me, Si), *sofiña* (BtsA, BtsF), *sofiny* (Td_1), *sofify* (Ts_1, Tk_1), *sofy* (Vz, Mf, Sak_1, Sak_2, Ba); **B**: *tadiny* (Bsk, Zfs), *tadiñ* (Tm, Tbk, Tsk); **C**: *ravembia* (Td_2).
In *sofina* and *tadiny* the nasalized final shows normal variation, such as in *ravina* (25). *Tadiny*, which comes from the standard Proto-Indonesian word for ear, has been restricted in many dialects to the meaning of ear's hole (Malzac 1888: 595 for Me).

40. 'eye': **A**: *maso* (Me, Tm, Bsk, Si, BtsA, BtsF, Tbk, Tsk, Zfs, Ts_1, Tk_1, Vz, Sak_1, Ba, Td_1); **B**: *fanjo* (Mf); **C**: *fijere* (Sak_2); **D**: *fihaino*

(Td_2). A word remarkably unique and affected only by taboo or dual vocabulary.

41. 'nose': A: *orona* (Me, Bsk, Si, BtsA), *oron* (Zfs), *oroñ* (Tkr), *oroña* (Tm, BtsF, Tbk, Td_1), *oroño* (Tsi), *oro* (Tsk, Vz, Mf, Sak_1, Sak_2, Ba); B: *fantsoana* (Td_2).
A word replaced only in one case by the effect of dual vocabulary, but the nasalized final shows the same wide variations as *rorona* (35).

42. 'mouth': A: *vava* (Me, Tm, Bsk, Si, BtsA, BtsF, Tbk, Tsk, Zfs, Tsi, Tkr, Vz, Mf, Sak_1, Ba, Td_1); B: *fitsamike* (Sak_2); C: *falie* (Td_2).
A very homogeneous word throughout the island, affected only in two cases of dual vocabulary. Words which do not have finals of the *ka, tra, na* type show fewer variations (see also *lava*, 14).

43. 'tooth': A: *nify* (Me, Tm, Bsk, Si, BtsA, BtsF, Tbk, Tsk, Zfs, Tsi, Tkr, Ba, Td_1); B: *niky* (Vz), *hy* (Mf, Sak_1); C: *salaka* (Sak_2); D: *fihitsika* (Td_2).
Besides C and D which belong to dual vocabulary, we have a west group (B), the word of which means the gums in the east-center group.

44. 'tongue': A: *lela* (Me, Tm, Bsk, Si, BtsA, BtsF, Tbk, Tsk, Zfs, Tsi, Tkr, Vz, Mf, Sak_1, Sak_2, Ba, Td_1); B: *fameleke* (Td_2).
A stable word affected only in one case by dual vocabulary.

45. 'foot': A: *tongotra* (Me, Si, BtsA, BtsF), *tongoke* (Vz), *hongotra* (Bsk), *hongotro* (Tsi), *ongotra* (Tbk); B: *raadro* (Tm), *rangotry* (Tsk), *rongotra* (Zfs); C: *nity* (Tkr); D: *fañitsake* (Sak_2); E: *tomboke* (Mf, Sak_1), *tomboky* (Ba), *tomboke* (Td_1); F: *fandia* (Td_2).
The definition of the parts of the leg or what has to do with it varies greatly (see also 46 and 47). *Rongotra* may be used for foot, claw, and knee. D and F are of course words of dual vocabulary made from the words to tread on (D) and to walk (F). E has influenced A in the Vz word. *Raadro* (Tm) has a correspondence in Me where the word *ranjo* connotes the leg. B group may derive from a combination of *ranjo* + *ongotra*. *Ongotra* may also have had a role in a former group beginning with *tombo* (*ton* +) which led to splitting up with derived forms such as *tongotra* or unmodified forms of the present E group. As in *fakany* and *valiatra*, we have preferred to keep them apart.

46. 'claw': A: *rongotra* (Me, Si, Zfs), *rongotry* (Tm, Bsk, Tsk), *ron-*

poho (Tsi), *rangoka* (Sak$_1$, Sak$_2$, Ba, Td$_1$, Td$_2$); B: *angoha* (Tbk); C: *koha* (BtsA, BtsF, Mf); *angôfo* (Tkr); D: *tsimalahoka* (Vz).
A and B are kept separate as in 46. The final of A group shows a wide range of regular variations. In *angôfo*, a shift of the following consonant is a prefixation of a type also found in words such as *akoho* (hen), *akanga* (guinea fowl). See also *anghombe* in Flacourt (1658) and now *ombe* in Besileo, elsewhere in the uplands *aomby* (cv). *H* can alternate with *ng* in such cases and be another possibility to *f* in other contexts. In many other places *koho* is the standard word for nail.

47 'knee': A: *lohalika* (Me, Si, BtsA, BtsF), *lohalike* (Sak$_1$, Sak$_2$), *lohalitsy* (Tkr, Ba), *lohalitra* (Bsk, Tbk), *lohalitsy* (Tsi); B: *tsipoko* (Tm), *tipoko* (Zfs), *pokopoko* (Tsk); C: *angotra* (Vz, Mf, Td$_1$, Td$_2$).
The correspondences of final *ka* with finals of the *tra* type are not extremely common (see in 45 Me *tongotra* and Vz *tongoka*). In this type of situation we discover a greater complexity than previous studies by Ferrand (1902: 12) and Dez (1963: 593) had described.

48 'hand': A: *tanana* (Me, Si, BtsA), *tañana* (Tm, Bsk, BtsF, Tbk, Tsk), *tsiam* (Tkr), *tañaña* (Tsi), *taña* (Zfs, Sak$_1$, Ba, Td$_1$), *tanga* (Mf), *sambotanâ* (Vz); B: *fandrambo* (Sak$_2$); C: *fika* (Td$_2$).
The word derived from Proto-Indonesian is replaced only in two cases by words of dual vocabulary. It is a control case to follow the transformations of an intervocalic nasal and a nasalized final. Prefixation of *ramba* in Vz (tail) suggests that the word *tanâ* may have, at least in this dialect, designated the arm as well as the hand.

49 'belly': A: *kibo* (Me, Bsk, Si, Tbk), *koko* (Tsi); B: *troka* (BtsA, BtsF, Td$_1$), *troky* (Tm, Tsk, Zfs, Ba), *trike* (Vz, Mf, Sak$_1$), *troko* (Tkr); C: *saroto* (Sak$_2$); D: *fitsafoa* (Td$_2$).
Besides two innovations due to dual vocabulary, we note a double distribution between A (center-east) and B (north, west, south and east). It is suggested that the Tsi *koko* is derived from *kibo* by metathesis and assimilation. See also *ambozoho* (50) and *amboko* (52). Variations of final *ka* are extremely regular.

50 'neck': A: *vozona* (Me, Bsk, Si, BtsA, Td$_1$, Td$_2$), *vozoño* (Tm, BtsF, Tbk, Tkr), *ambozoño* (Tsi), *vozo* (Tsk, Zfs, Vz, Mf, Sak$_1$, Sak$_2$, Ba).
For final *na*, compare with *varona* (20 and 36), *ravina* (25), *elona* (18), etc. Tsi has the fossil prefixation of the *angôfo* type (see under 46).

31. 'breasts': **A**: *nono* (Me, Tm, Bsk, Si, BtsA, BtsF, Tbk, Tsk, Zfs, Tkr, Vz, Mf, Sak_1, Sak_2, Ba, Td_1, Td_2); **B**: *somondrara* (Tsi).
B is known all over the island with a more restricted meaning. In Tbk it means 'breasts of young girls'. See also Malzac dictionary, p. 382: "*Somondrara*: dont les seins se développent, se dit des jeunes filles." Besides the Tm exception, the word is remarkably stable.

32. 'heart': **A**: *fo* (Me, Tm, Bsk, Si, BtsA, BtsF, Tbk, Tsk, Tkr, Vz, Mf, Sak_1, Sak_2, Ba, Td_1, Td_2), *fon* (Zfs); **B**: *amboka* (Tsi).
Tsi cannot be included in A, for regular change does not give *f/mb* but *f/mp*. See 49.

33. 'liver': **A**: *ay* (Me, Bsk, Si, BtsA, BtsF, Tbk, Tsk, Zfs, Tsi, Ba), *ai* (Tkr, Vz, Mf, Sak_1, Sak_2, Td_1, Td_2).
A very stable word. In the whole west voiceless *e* replaces voiceless *i*.

34. 'drink': **A**: *misotro* (Me, Si, BtsA, BtsF); **B**: *minona* (Bsk, Tbk), *misotsa* (Tm), *minon* (Zfs), *mino* (Tsk, Vz, Mf, Sak_1, Ba, Td_1, Td_2); **C**: *migiaga* (Tsi), *migiaka* (Tkr); **D**: *mitele* (Sak_2)
A is an innovation on the plateau (*sotro* means spoon), for B must have been the most widespread word (for evolution of nasalized final, compare with 50A). Until recently Me had kept *minona* in certain restricted contexts (absorption of the poisonous ordeal of *tangena*). *Migaka* is known also today in Imerina with the meaning of drinking moderately. *Mitele* is an innovation due to dual vocabulary. Me cognate *mitelina* means 'swallow'.

35. 'eat': **A**: *mihinana* (Me, BtsA, BtsF), *mihana* (Tsk, Zfs, Tkr, Mf, Sak_1, Ba); **B**: *mihumuha* (Tm), *hamaha* (Tsi), *homana* (Si, Tbk, Tsk), *hómana* (Bsk), *homa* (Vz, Td_1); **C**: *mitramike* (Sak_2); **D**: *milama* (Td_2).
If one pulls out C and D due to dual vocabulary we have two words probably cognate in the protolanguage in Indonesia. *Mi* predicative prefix of the active form of A and of one case in B, when removed, leaves *hinana* and *homana*, which may be derived from *hana* with infix *in* and *om*. On account of this separation already present in the protolanguage (according to Dempwolff) we have kept A and B separate. *Homana* is known in Me in specialized contexts. The nasalized final is of the *na/fa* or zero type, a normal range of variation; the suppression occurring in west and south?

36. 'bite': **A**: *mamaksa* (Me, Si), *mañakitra* (Tm, Tbk, Tsi), *mañaktra* (BtsA), *manakitra* (Bsk), *mañakitry* (Tsi), *mañahitsa* (BtsF), *mañahitsy*

(Bs), *maŝeheise* (Vz, Mf, Sak_1), *maŝeitra* (Zfs), *moñaitry* (Tkr); **B**: *mandramatse* (Sak_2); **C**: *mitifatse* (Td_1, Td_2).
Besides shifts due to taboo or dual vocabulary, there is one group with multiple variants due to: uneven evolution of nasal following prefix *ma*; shortening in *a* of diphthong *ai* or the reverse; equivalence of *h/k*; modifications according to the dialects of the final *tsefira* type.

57. 'to see': **A**: *mahita* (Me, Tm, Bsk, Si, BtsA, BtsF, Tbk, Tsk, Zfs, Tsi, Tkr, Vz, Sak_1, Sak_2, Ba); **B**: *manenta* (Mf), *maŝenta* (Td_1); **C**: *mahavajo* (Td_2).
There is often a tendency to confuse 'to see' and 'to look at'. *Manenty* is known among the Sakalava but *mahita* is more common. *Mahavajo* is due to dual vocabulary.

58. 'to hear': **A**: *mihaino* (Me, BtsA, Tsk, Td_2), *miheno* (Tm, Si, Zfs), *mitahino* (Bsk), *mitekino* (BtsF), *miteno* (Tbk); **B**: *mitandreny* (Tsi), *tandriñy* (Tkr); **C**: *mijonjy* (Sak_1, Sak_2); **D**: *nahare* (Vz, Mf, Ba, Td_1).
There is no clear-cut distinction between 'hear' and 'listen to'. *Mahare* or *mandre* means 'listen to' in Me, and *mitandrina* 'watch out'. In group A we find the *ai/e* change encountered in 56A. Existence of *te* or *to* could come from paronymic attraction with another word such as *mitandrina*.

59. 'to know' (things): **A**: *mahay* (all dialects except Zfs: *hay* and Tkr: *he*). *Ma* is a predicative prefix.

60. 'sleep': **A**: *matory* (Me, Tm, Si, BtsA, Tsk, Zfs, Ba), *matoro* (Tks); **B**: *manary* (Bsk, BtsF, Tbk, Tsi); **C**: *mirovo* (Vz, Mf, Sak_1, Sak_2, Td_1, Td_2).
There is a definite distribution between C (west) and A (center and east). The existence of B group is due to the confusion between sleep and lie. Some dialects do not have a distinct word for these two notions (Bsk, Tbk); others have a slight distinction created by long (BtsF) or short (Tsi) reduplications (see under 67). In other Malayo-Polynesian languages, for example, Tahitian, *moe* (sleep) has been replaced by *ta'oto* (to lie), which now means to sleep. In A Tkr *matoro* shows once more the instability of final vowels which in northern dialects follow the vowel of the preceding syllable.

61. 'die': **A**: *maty* (Me, Tm, Bsk, Si, BtsA, BtsF, Tbk, Tsk, Zfs, Tsi, Tkr, Ba, Td_1, Td_2), *mate* (Vz, Mf, Sak_1); **B**: *mirofotse* (Sak_2).

Except in one case owing to dual vocabulary the word is very homogeneous, with the slight variation *j*/voiceless *s*.

52. 'to kill': **A**: *mamono* (Me, Tm, Bsk, Si, BtsA, BtsF, Tbk, Tsk, Zfs, Tsi, Tkr, Mf, Sak_1, Ba, Td_1, Td_2); **B**: *mamango* (Vz); **C**: *mandevina* (Tsk_2)
There are no reasons to include B in A, in spite of an apparent resemblance. C is an innovation of dual vocabulary.

53. 'to swim': **A**: *milomano* (Me, Si, BtsA, Tsk_1, *milomotse* (Tm, Bsk, BtsF, Zfs, Tsi, Tkr, Sak_1, Sak_2), *lomano* (Ba), *miloño* (Mf, Td_1, Td_2), *milano* (Vz), *mandeto* (Tbk).
The root word *lano* or *leño* is supplemented by prefixes *mi* and *ma* + nasal and infix *om*. Two forms exist in certain dialects, such as in Tm *milaono* and *milomano*.

64. 'to fly': **A**: *mandrina* (Me, BtsA); **B**: *manembana* (Bsk, Si), *mañembaña* (Tsi), *manembotse* (Tm), *mañembon* (Zfs), *mañembo* (Tsk); **C**: *mitsidrika* (BtsF), *mitilily* (Tkr), *mitilin* (Td_1, Td_2), *mitsily* (Mf, Sak_1, Sak_2, Ba); **D**: *miriña* (Tbk), *maniriky* (Vz)
In B we get *n* or *ñ* indifferently but we cannot account for *bo* versus *ba*. The words, however, do not seem to derive from roots as different as *ambony* (above) and *ambany* (below). C and D have rare consonant or final correspondences *t/t*, *ka/ky*.

65. 'to walk': **A**: *mandeha* (all dialects except B); **B**: *manitsaka* (Sak_2).
The only innovation due to dual vocabulary has a cognate in Me, where the word *manitsaka* means 'to tread on'.

66. 'to come': **A**: *avy* (All dialects).

67. 'to lie': **A**: *mandry* (Me, Tm, Bsk, Si, BtsA, Tbk, Tsk, Zfs, Tkr, Ba, Td_1, Td_2), *mandrimandry* (BtsF), *mandrimandry* (Tsi), *mandre* (Vz, Mf, Sak_1, Sak_2).
Tm sometimes also uses *mandrimandry* and Tbk *mivalampatra*. With a few variations due to reduplication or change *e/i*, the word is very uniform; for interference with 'to sleep' see under 60.

68. 'to sit': **A**: *mipetraka* (Me, Bsk, Si, Tkr), *mipetrake* (Sak_1); **B**: *mitsetra* (BtsA), *mitsetse* (BtsF), *mitsetre* (Tbk), *manitsy* (Tsi); **C**: *mitoboka* (Tm), *mitoboky* (Tsk, Zfs), *mitoboky* (Ba), *mitoboke* (Vz, Mf, Td_1, Td_2); **D**: *mitsiake* (Sak_2).
In A, C, and D finals *ka/ki/ke* follow the normal area distribution, as do *tra/tri/tse* in B. B is distinguished more by prefixes *mi* or *ma* + nasal (as in C3) and the infix *an* in the Tbk word. C group covers the south of the island. Besides D, which is due to dual vocabulary, a

word of A or B group is selected in each dialect but usually both are known. *Mitoetra* is known in Me, but is less common, and *mijoetraka* is used in Ambositra, but less often than *mitoetra*.

69. 'stand': A: *mitsangana* (Me, Tm, Bsk, BtsA, Tbk, Vz), *mitsangan* (Zfs), *mitsangaña* (BtsF, Tsi), *mitsanga* (Tsk, Tkr, Sak_1, Sak_2, Ba, Td_1, Td_2, Mf); B: *mijoro* (Si).
A is known all over the island (with the variations or disappearance of *na*), even in Si (*mitsangana*), where *mijoro* is more common. Conversely, *mijoro* and cognates exist as a less common synonym in many other dialects, especially in Me and BtsA (*mijoro*) and Tbk (*midoredoro*)

70. 'give': A: *manome* (Me, Si, BtsA, Vz, Mf), *matoma* (Tm, BtsF, Tbk, Tsk, Zfs, Tsi, Sak_1, Sak_2, Ba, Td_1, Td_2), *mashony* (Bsk), *mashony* (Tkr).
The quality of the nasal varies in *ma* + *n*. There are curious changes *o/a* (influence of the preceding syllable) and voiced *e/* voiced *i* for Tkr and Bsk words.

71. 'say': A: *milaza* (Me, Bsk, Si, BtsA, BtsF, Zfs, Ba); B: *mizaka* (Tm, Tbk, Zfs); C: *mivola* (Tsk, Tkr, Vz, Mf, Sak_1), *mivolaña* (Tsi); D: *maiamba* (Sak_2); E: *milalily* (Td_1, Td_2).
Here again, as for the verbs 'stand' and 'see', we are faced with a problem of choice of synonyms. Tbk had both *mizaka* and *milaza*, Tsk *mizaka* and *mivola*, and tossing up was necessary. In Bsk, besides *milaza*, B and C are known. Besides A, *misiny* and *manambara* (a cognate to D) are known in Me.

72. 'sun': A: *andro* (Vz, Mf, Sak_1, Sak_2), *taniandro* (Tkr), *masoandro* (Me, Tm, Si, BtsA, BtsF, Tbk, Tsk, Zfs, Td_1, Td_2); B: *masova* (Tsi), *masiva* (Tkr); C: *maheniky* (Ba).
Masoandro can be placed either in A or B, and tossing up was necessary. In the interior *andro* means only 'day'. *Masoandro* is therefore the eye of the day; cf. Malay *matahari*. However, *maso* and *mata* are not related, contrary to Dempwolff's opinion. For anthropological inferences see Hebert (1965: 90).

73. 'moon': A: *volana* (Me, Bsk, Si, BtsA, Tbk, Td_1, Td_2), *volaña* (Tm, BtsF, Tsk), *vola* (Zfs, Vz, Mf, Ba), *dasolaña* (Tsi); B: *fenjava* (Tkr); C: *boara* (Sak_1, Sak_2)
B also means 'that which is shining'. *Dasolaña* resembles Me *diavolana* 'moonshine'. For comparison with changes of final *na* see *hasana* under 55. There is also an interesting discussion by Hebert (1965: 93–96), whose conclusions resemble our own.

74. 'star': **A**: *kintana* (Me, Bsk, Si, Tbk), *kitana* (BtsA), *lakintana* (Tsi), *ankintana* (Tkr); **B**: *vasiana* (Tm, BtsF), *vasiha* (Tsk), *vasia* (Zfs, Vz, Mf, Ba, Td_1, Td_2), *basia* (Sak_1, Sak_2).
In A, Tsi and Tkr have prefixed form *laki*, which Hebert connects with the word for *male* (see under 17), noting that in some places there also exists *zanakinta* with prefixed form *zana* (child) (Hebert 1965:102 ff.). Final *na* follows the ordinary changes according to areas, although *h* in *vasiha* (Tsk) is unusual. Hebert's map of the two areas *vasia/kintana* in Madagascar (south and west versus north-center-east) fits our finding although we note that the *kintana* area goes more toward the south and east. Tbk has both *kintana* and *vasena*; both *vasiha* and *kinta* are known in Tsk

75. 'water': **A**: *rano* (all dialects except Sak_1 and Sak_2); **B**: *mazy* (Sak_1 and Sak_2). B is an innovation probably due to linguistic taboo. At the western frontier of Imerina a river has the name Mazy

76. 'rain': **A**: *orana* (Me, Si, BtsA, Tbk, Td_1, Td_2), *oraña* (Tm, BtsF, Tsi), *orake* (Bsk), *ora* (Vz, Mf, Ba); **B**: *harikandro* (Tsk), *rekandro* (Zfs); **C**: *maleny* (Tkr), *mahale* (Sak_1, Sak_2)
For changes of final *na* in A see under 55 and 73. The Me cognate of B is *erika* 'drizzling rain', and *erik'andro* 'drizzling weather', an often met situation on the east coast. C is paralleled in Me by *mahalena* 'which causes you to be wet', an expression denoting replacement after the word *orana* in the Sakalava kingdom at the time it also included the Tankarana people.

77. 'stone': **A**: *vato* (all dialects)

78. 'sand': **A**: *fasika* (Me, Si, BtsA, BtsF), *fasana* (Tm), *fasina* (Bsk, Tbk), *fasiky* (Tkr), *fasony* (Zfs), *fasy* (Tsk, Mf, Ba), *fase* (Sak_1, Sak_2), *fasy* (Td_1, Td_2); **B**: *alokaka* (Tsi); **C**: *tany* (Vz).
In A we have the greatest variety of finals. For correspondences *na/ka* see under the word 'ashes' (83). For race *ka/ka* see under 47 and also Ferrand (1903:23, 26). Initially Td may have had *fasike*, which was shortened. The word for designating earth is known in Vezo (ibid., p. 79). B is known only in the northwest.

79. 'earth': **A**: *tany* (Me, Tm, Bsk, Si, BtsA, BtsF, Tbk, Tsk, Mf—also *hatany*—Tsi, Tkr, Ba), *tane* (Sak_1, Sak_2, Td_1, Td_2), *tany* (Zfs); **B**: *fasimahery* (Vz).
The only slight changes are due to nasalization of *n* in Zafisoro and [illegible] [illegible] finals. The Vz word means 'solid sand', an under-

standable innovation by a maritime people who love to say that they only "walk by canoes."

80. 'cloud': **A**: *rahona* (Me, Si, BtsA), *rahoña* (Bsk, BtsF), *rahotr* (Tbk), *raho* (Vz, Mf, Ba, Td$_1$, Td$_2$); **B**: *mika* (Tm, Tbk, Tsk, Zfs) **C**: *randro* (Tsi); **D**: *zaoño* (Tkr); **E**: *Aiboke* (Sak$_1$, Sak$_2$).
C, D, and E may reflect invention of several dialects in the Sakalava kingdom when *raho* (or another word) became tabooed (for instance, after the death of a king in whose name the word was included). *Mika* is typical of the south and east and was already noted in seventeenth-century Antanosy by Flacourt. Final *n* follows the ordinary changes according to areas except in Tbk which is very unusual. The Me cognate of Tkr *zaoño* is *zaovona* (fog).

81. 'smoke': **A**: *setroka* (Me, Si, BtsA, BtsF, Tsk), *setroky* (Tm, Ba) *setroke* (Mf, Sak$_1$, Sak$_2$), *setroko* (Tkr), *setrok* (Vz); **B**: *tsiño* (Tm) *aisina* (Tbk), *isy* (Zfs); **C**: *emboka* (Bsk), *temboka* (Tsi); **D** *mahotod* (Td$_1$, Td$_2$). For changes in A finals compare with 8 (*tandroka*), 49 (*troka*), and 81 (*setroka*). *S* group is cognate with the Vz word for sand (78C). *C emboka* in Me is translated as follows by Malzac: "encens, resines propres à brûler" (1888:118), but Me *temboka* means 'sweat'.

37. 'fire': **A**: *afo* (Me, Tm), Bsk, Si, BtsA, BtsF, Tbk, Tsk, Zfs, Vz Mf, Ba, Td$_1$, Td$_2$); **B**: *moro* (Tsi, Tkr, Sak$_1$); **C**: *mahatonak* Sak$_2$).
Except for the innovation of C, B, a borrowing from Swahili, appears in the west and northwest where Swahili influence was important.

33. 'ashes': **A**: *lavenona* (Me, Si, BtsA, Tbk, Vz), *lavenoña* (BtsF) *lakevoño* (Tsi), *lakevo* (Tsk, Zfs, Ba), *lavenoke* (Mf, Sak$_1$, Sak$_2$, Td$_1$ Td$_2$); **B**: *jofo* (Tsi), *jifo* (Bsk, Tkr).
In A we encounter the correspondence in finals *na/ka* such as in 71 (*fianna, fasika*) and 99 (*mamafmaike*). But in the south and east a metathesis occurred, *lavenoka* giving *lakevoño* or *lakevo*. *Jofo* means 'dust' in Sihanaka. See also in Me after Malzac: *mitra-jofo* "qui produit de la fumée, de la poussière" (1038:274).

84. 'burn': **A**: *mandevo* (Me, Si, BtsA, Tsk), *mañevo* (Bsk, BtsF, Tbk Zfs, Tsi, Tkr, Mf, Sak$_1$, Sak$_2$, Ba, Td$_1$, Td$_2$), *mañevo* (Tm), *manevo* (Vz)
A word common to the whole island. Differences come from the predicative form *ma* + nasal+, a root of which we do not know the initial consonant.

85. 'path': A: *lalana* (Me, Bsk, Si, BtsA, Tbk), *lalaña* (Tm, BtsF, Tsi, Td_1, Td_2), *lalan* (Zfs, Tkr), *lala* (Tsk, Vz, Mf, Sak_1, Sak_2, Ba). A very homogeneous word, modified only by the variations of *na*, *ña*, *n*, zero according to distributions by areas (west and south have *na* suppressed, Me and neighbors have a nonvelar *n*). Compare with *lanana* under 48.

86. 'mountain': A: *vohitry* (Tm), *vohitra* (Tsk, Zfs), *vohitse* (Mf, Sak_1, Sak_2), *vohitsambo* (Vz), *lohavohitra* (Td_1, Td_2), *tandrombohitra* (BtsF, Ba), *tendrombohitra* (Me, BtsA); B: *tanety* (Bsk, Si, Tsi); C: *tarirana* (Tbk); D: *bongo* (Tkr).
The problem here is to know what kind of mountain is talked about. A includes in all the words *vohitra* (or its variations *tri/tse*) and adds top (*tendro* or *loha*) or high (*abo*) in Vz. *Tanety* is known in Me with the meaning of 'hill' and *bongo* connotes a small elevation.

87. 'red': A: *mena* (all dialects).

88. 'green': A: *maitso* (Bsk, Si, Tkr, Vz, Sak_1, Sak_2, Ba, Td_1, Td_2), *maintso* (Me, BtsA, BtsF, Tsi, Mf), *metso* (Tm, Tbk, Tsk, Zfs).
The occurrence of *n* before *ts* is not significant. In fact in Me the word is officially written *maitso*, although *maintso* is heard more often. In certain dialects, *ai* is replaced by *e*, and although here it is localized in the south and east, elsewhere it seems more haphazardly distributed (see under 69).

89. 'yellow': A: *vony* (Me, Si, BtsA, Vz), *voñy* (BtsF, Tsk, Mf, Ba), *voñe* (Sak_1, Sak_2, Td_1, Td_2), *voñ* (Zfs); B: *tamotamo* (Tm), *manamotamo* (Tbk, Tkr); C: *avaka* (Bsk); D: *fandrasla* (Tsi).
In A intermediate *n* is more or less velarized and *i* may be replaced by voiceless *e* in the west or in the south. In B Me, BtsA, Tm *mavo* is a very near synonym to *vony*. *Tamo* reduplicated and with prefix in Tbk has its equivalences elsewhere. In Me it means 'saffron'.

90. 'white': A: *fotsy* (Me, Tm, Bsk, Si, BtsA, BtsF, Tbk, Tsk, Zfs), *foty* (Vz, Mf, Sak_1, Sak_2, Ba, Td_1, Td_2); B: *malandy* (Tsi, Tkr). In A the change *ts/t* is a well-known phenomenon (see Dez 1963, 569).

91. 'black': A: *mainty* (Me, Bsk, Si, BtsA, BtsF, Ba, Td_1, Td_2), *maity* (Ba), *mety* (Tsi), *mainka* (Tbk), *mity* (Tsk), *mity* (Zfs), *maitonty* (Tm), *mainte* (Mf, Sak_1, Sak_2); B: *joby* (Tkr).
In A there is a range of variation including *ai*, *e*, *i* and more or less nasalization. *Ts* alternates with *t* in the west or southwest

Other possibilities for *ai* are more numerous than in *maitso* (88) and *maina* (99).

92. 'night': A: *alina* (Me, Bsk, Si), *aliña* (Tm, BtsF), *alin* (Zfs), *alite* (BtsA, Tbk), *alitsy* (Tsi, Tkr), *haliña* (Td$_1$, Td$_2$), *aly* (Tsk, Vz, Mf, Ba), *ale* (Sak$_1$, Sak$_2$).
Final *na* exhibits variations as already described for *ravina* (25) and *mimaftadroy* (59). The replacement of *na* by *la* in two dialects is unusual and must be attributed to the influence of the preceding syllable. Initial *h* in Tandroy is unexplained.

93. 'hot': A: *mafana* (all dialects except Td$_2$); B: *matrooka* (Td$_2$).

94. 'cold': A: *mangatsiaka* (Me, Bsk, Si, Tbk, Zfs), *mangatsiaky* (Tm); B: *manara* (BtsF, Tsk); C: *mangitsy* (BtsA), *manintsy* (Tsi, Tkr, Vz, Mf, Sak$_1$, Sak$_2$), *manitsy* (Ba), *mañintsy* (Td$_1$, Td$_2$).
C is used in the western half of Madagascar. In the eastern half A and B often coexist in one dialect. Me also has *manara* for the "cold" house above a tomb.

95. 'full': A: *feno* (Me, Bsk, Si, BtsA, BtsF, Tbk, Tsk, Zfs, Tsi, Tkr, Mf, Ba, Td$_1$, Td$_2$); B: *totomorana* (Tm); C: *atseka* (Vz), *atseke* (Sak$_1$, Sak$_2$).
B and C are innovations in vocabulary localized in the south and east and in the west.

96. 'new': A: *vao* (all dialects except Mf, which has *vaovao*, a reduplication also known elsewhere).

97. 'good': A: *tsara* (Me, Tm, Bsk, Si, BtsA, Tbk, Tsk, Zfs, Tsi, Tkr); B: *soa* (BtsF, Vz, Mf, Sak$_1$, Sak$_2$, Ba, Td$_1$, Td$_2$).
BtsA and Tsk have both *tsara* and *soa*, and tossing up was necessary. B definitely belongs to the west and the southwest; *soa* is known in Me in certain contexts (*tonga soa*: arrived safely).

98. 'round': A: *boribory* (all dialects except Tbk, Tm and Tsi, which have *tavoribory*).

99. 'dry': A: *maina* (Me, Si, BtsA), *maiña* (Tm, Bsk, BtsF, Tbk, Tsk, Zfs), *maitsy* (Tsi), *maike* (Vz, Mf, Sak$_1$, Sak$_2$, Ba, Td$_1$, Td$_2$), *meky* (Tkr).
The diphthong *ai* does not vary much, but in finals we get either the *na* class or the *ka* class, as in 83 and 78.

100. 'name': A: *anarana* (Me, Si, BtsA), *adaraña* (Tsi, Tm, Bsk, BtsF, Tbk, Tkr), *añara* (Tsk, Zfs, Sak$_1$, Sak$_2$, Ba, Td$_1$, Td$_2$), *anara* (Vz, Mf).
The first *na* has no velar in the center and in the southwest. The

final na has no velar in the center and disappears in the west and the south

Analysis of results

9. THE MOST STABLE ITEMS. The most stable relationships between form and meaning for Malagasy are indicated in Table 2. The following

TABLE 2: Most stable basic vocabulary items in eighteen Malagasy speech communities

Meaning and position on 100-word list	Number of appearances	Meaning and position on 100-word list	Number of appearances
1. I	18	51. breasts	17
3. we (inclusive)	17	52. heart	17
4. this	16	53. liver	18
5. that	18	59. to know	18
6. who	12	61. to die	17§
7. what	13	62. to kill	16§
8. not	18	63. to swim	18
11. one	18	65. to walk	17§
12. two	18	66. to come	18
14. long	18	67. to lie	18
15. small	17	69. to stand	17¶
17. man	13	70. to give	18
20. bird	10	75. water	16
22. louse	16	77. stone	18
25. leaf	18	79. earth	17
27. bark	18	84. to burn	18
29. flesh	17	85. path	18
31. bone	18	87. red	18
33. egg	18	88. green	18
36. feather	18	91. black	17
38. head	14*	92. night	18
40. eye	15†	93. hot	17‡
41. nose	17‡	96. new	18
42. mouth	16†	98. round	18
44. tongue	17‡	99. dry	18
48. hand	16§	100. name	18
50. neck	18		

* There are four cases of special terms used as part of dual vocabulary. ('Special term' refers to alternatives where there is dual vocabulary resulting from such differences in speech patterns or linguistic taboo.) The common term *loha*, however, [illegible] in all eighteen dialects. When another term is used, [illegible] derogatory connotation.

† Eye, and long, have special terms.

‡ [illegible] has a special term.

§ Hand, has a special term.

¶ The exception has the common term as a less frequent synonym.

meanings were most stable in that they have a single cognate form to express them throughout all Malagasy speech communities: I, that, who, what, not, one, two, long, man, bird, louse, leaf, bark, bone, egg, feather, neck, liver, to know, to swim, to come, to lie, to give, stone, to burn, path, red, green, night, new, round, dry, name. The existence of dual vocabularies, one for addressing commoners, and the other for referring to or addressing nobility, chiefs, ranked persons, or elders, affected basic vocabulary, particularly in words for parts of the body, and also in the verbs to die, to kill, and to walk. In the Sakalava and Tandroy special vocabularies, one could not say that the chief or the elder went walking using the same expression which was employed when walking was done by an ordinary man. Similarly, because of the *mana* residing in his person, the death of an important man could not be expressed with the common term. His murder or death was a special event of significance to the society as a whole and as such was marked with a special word signifying this event. For variation among less stable forms see the preceding section.

10. PROBLEMS IN SUBGROUPING. The lexicostatistical techniques that we have used (cf. Gudschinsky 1956) have been applied in the past primarily to languages sharing relatively few cognates out of the list of 200 key words (Swadesh 1954). This study is concerned with speech communities with a high degree of linguistic similarity and in which the number of cognates shared is in all cases higher than 50 per cent and in some cases as high as 92 percent. As a result, it can be expected that any linguistic tree produced simply on the basis of the mathematical techniques of lexicostatistics is apt to be highly inaccurate. There are many reasons for this. The original theory of glottochronology is based on the observation that a certain percentage of the words in a language out of a basic word list will be dropped in a given number of years. Empirically, this change for two languages with no linguistic contact is more or less constant, so that after a given period of time one can examine the list of shared cognates remaining in the word list and determine how long the languages have been separated. It is important to note that the percentage of words out of the word list that are independently replaced each millennium will be a constant only if the two languages have no contact with one another. Presumably the ideal situation for glottochronology would be one in which two populations sharing a common language are suddenly separated. Under such circumstances one would expect that the constant of change of cognates would operate in the ideal manner and the mathematical expectation would be fulfilled. However, if there is linguistic contact between the two popu-

lexicon, the length of time that is necessary before the languages are sufficiently separate that no linguistic contact takes place becomes an unspecified variable. There is at present no fully accepted method of determining the rate of linguistic differentiation during this period.

The primary problem of lexicostatistical analysis of dialects lies in the assumption of independence. The retention rate is supposed to operate separately on each word list. Hymes (1960:19) cites three principal factors that would prevent independent dialect divergence. These are: contact, which increases the likelihood that words retained and dropped will be the same; drift, or the similarity of change resulting from similar internal structure of the language; and dregs, which refers to heterogeneity in a word list with respect to the probability of retention. If one were to use the standard formula to calculate divergence time given the percentage of shared cognates, each of the above factors would operate to produce an underestimate in time depth.

Even aside from the problems resulting from application of the procedure of glottochronology to dialects, a number of other sources of error remain. Among the deficiencies of the "pure" lexicostatistical approach we note that certain words of the word list may be tabooed and will drop out of the word list more readily (cf. Ehrendorf 1951). Furthermore, though Swadesh (1956:305) discounts its importance, borrowing can theoretically influence the basic word list. In addition, after languages have been separated for many millennia, it becomes increasingly difficult to ascertain which words are cognates and which are not. Cross (1962:488) feels that past even 2,000 years this problem becomes insurmountable. This problem becomes particularly important since each individual cognate will contribute more and more in determining the period of time the languages have been separate. Finally, the problem of preparing a word list which is comparable for all the languages considered has introduced a source of error (Hymes 1960:8). Though the above problems affect the use of lexicostatistics generally, analysis of dialects may be expected to suffer most of the sources of error above and the other sources of error listed earlier in addition.

Given the fact that we expect the genetic tree to be only the roughest kind of approximation of reality, it might be asked why we wish to prepare one by lexicostatistical techniques? First of all, there is no alternative method of producing quickly and easily a genetic tree for a number of languages. Furthermore, by using a theoretical genetic tree the linguist has a basis against which to compare his purely historical information. Considerable work may then allow him either to amend his previous model of the genetic tree or to criticize the lexicostatistically-

produced genetic tree. In either case, he has gained insight. In the first case the specific facts of the linguistic history one is trying to explicate become clearer. In the second case a contribution to future lexicostatistical study is made. For example, the recent discussion of Dyen's Austronesian language classification (Dyen 1965, Grace 1966, Dyen 1966, Hymes 1966) can be expected to produce refinements both of the anthropologists' understanding of Pacific culture history and of Dyen's method of classification.

In our analysis of dialects we are interested in the relationships of three variables. The first is time depth, a purely chronological factor indicating the total time that two populations speaking the languages in question have been separate. Separateness in this case may be defined simply as geographical discreteness, with the border between the linguistic areas to be determined empirically in each case. The second variable is the given figure for linguistic similarity measured for our purposes by the number of shared cognates for the word list as specified by Swadesh. The third variable would be some estimate for the kind of interaction between the two populations. This variable would presumably try to take into account the alteration in rate of change of the number of shared cognates that would be brought about by the fact that the two populations speak to one another. Traditional glottochronology is a method of estimating time depth only for populations that cannot communicate. Only two of the variables indicated above are considered in the customary glottochronological analysis: time depth and cognate sharing. The formula linking time depth and cognate percentage shared is

$$t = \frac{\log C}{2 \log r}$$

where r is the retention rate, C is the percentage of shared cognates and t is the time in millennia. What we are interested in doing is introducing a third variable to account for those situations in which the members of the two populations share enough similar cognates to understand one another, but in which dialect divergence is taking place and would presumably proceed up to the point where the two-variable analysis would suffice.

Swadesh has proposed (1955) that an additional variable I be designated as the average degree of separation between a pair of languages. It would have a value of 1 for complete independence and be small with increases in the degree of dependence. The formula would read

$$t = \frac{\log C}{I^2 2 \log r}$$

Similar is the attempt by Hattori (1958) to correct for lack of independence for dialects diverging from Proto-Japanese. Hattori calculated a value for, in effect, Swadesh's r as 0.7. The new formula, assuming that Hattori's modification applies for all cases in which the diverging dialects remain in contact, is

$$t = \frac{\log C}{1.4 \log r}$$

Since we are unfamiliar with any other attempts to give an empirical value to r, we will assume, in the calculations for divergence time of the Malagasy dialects, that $r = 0.7$.

The time estimate is only part of the analysis. It is now necessary to discuss the method by which the genetic tree is constructed. As Grace (1966:17) points out, there are a number of ways of constructing a genetic tree given a table of cognates shared for a number of languages. In this paper a very simple procedure has been employed for producing a genetic tree because we felt that using a procedure as sophisticated as Dyen's for classifying highly related dialects would be wasted effort. Such a technique may be of value for an analysis of 320 languages with relatively low percentages of shared cognates. There is a reason to believe that to use the technique on our cognate table (Table 3) would be unnecessary labor. This is primarily because the main error in our genetic tree will be introduced not by the procedure employed in constructing a tree, but rather by the differential influence of our third variable. Sustained contact between populations can be expected to inhibit the loss of shared cognates and since the degree of contact varies among the pairs of languages, the value of t should be different for each pair. The different degrees of contact among the sixteen populations will undoubtedly affect the rate at which cognates are dropped and will introduce the major source of error. In view of this permanent source of error, it seems likely that any "improved" methods of tree construction will be of minimum value. Using the less accurate method in this case is somewhat like the use of significant figures in calculations for problems in physics: once a high source of error is introduced, it no longer pays to retain precise values for the other numbers used in the calculations.

In accordance with this, the method of grouping is as follows. In the first step we look at all languages that share 92 percent cognates with the other languages. Thus for example we note that Me shared 92 percent or more cognates with Si, and Si shares 92 percent or more cognates with Btsk. These languages have no similar sharing of cognates with any other language, and we isolate them as a group. No

TABLE 7. *Percentages of cognates shared between languages*

	1 Me	2 Tm	3 Bz	4 S.	5 BnA	6 BnF	7 Tbk	8 Tsk	9 Zk	10 Tw	11 Tkr	12 Vz	13 Mf	14 Sak_1	15 Sak_2	16 Ba	17 Td_1	18 Td_2
Me	100	81	81	92	92	86	78	80	79	68	71	68	70	70	61	77	69	61
Tm	81	100	78	79	80	78	81	88	86	67	68	71	72	69	61	76	73	65
Bz	81	78	100	84	78	77	80	78	75	77	70	69	68	62	54	74	69	61
S.	92	79	84	100	86	84	79	75	77	72	69	67	68	66	59	72	66	58
BnA	92	80	78	86	100	88	77	81	79	68	74	71	73	71	61	79	71	67
BnF	86	78	77	84	88	100	77	95	77	67	71	74	79	74	66	82	75	66
Tbk	78	81	80	79	77	77	100	74	74	69	66	64	64	69	53	67	64	56
Tsk	80	88	78	75	81	89	74	100	89	67	70	71	73	72	62	76	72	63
Zk	79	86	75	77	79	77	74	89	100	65	68	69	72	71	61	76	71	63
Tw	68	67	77	72	68	67	69	67	65	100	75	67	64	69	59	68	64	56
Tkr	71	68	70	69	74	71	66	70	69	75	100	65	68	69	59	70	67	56
Vz	68	71	69	67	71	74	64	71	69	67	65	100	84	74	64	77	69	59
Mf	70	72	68	68	73	79	64	73	72	64	68	84	100	80	68	81	69	70
Sak_1	70	69	62	66	71	74	69	72	71	69	69	74	80	100	82	79	75	66
Sak_2	61	61	54	59	61	66	53	62	61	59	59	64	68	82	100	70	67	63
Ba	77	76	74	72	79	82	67	76	76	68	70	77	81	79	70	100	85	74
Td_1	69	73	69	66	71	75	64	72	71	64	67	69	69	75	67	85	100	83
Td_2	61	65	61	58	67	66	56	63	63	56	56	69	70	66	63	74	83	100

other language shares 92 percent or more cognates with any other language. The next step is at the 84 percent cognate level. At this step one member of the Me-Si-Bts.A group shared 84 percent or more cognates with Bts.P and Bsk, so these were all grouped together at this stage. Similarly, the Tm-Tsk-Zfs group was formed by linking together the languages with 84 percent or more shared cognates. The same is done for the 76 percent cognate level at which point all the languages but Tkr constitute one group. At the 68 percent level all the languages are grouped together (see Chart 1). Eight percent was chosen as the interval percentage separating the levels because, when the data were analyzed, 8 percent proved to be the largest interval that achieved a sufficient degree of separation of the languages to prove useful. With

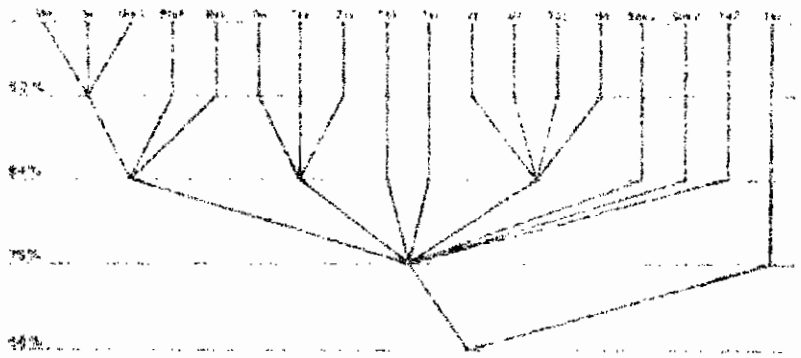


CHART 1: Branchings of Malagasy dialects based on cognate percentages at arbitrarily determined levels: levels at 8% intervals.

larger differences between the levels the groups produced were too large to be useful. When the interval was smaller too great a degree of accuracy was implied. In our discussion of the subgroupings, other slightly variant intervals will be used to produce additional genetic trees for comparison.

VI. THE PROBLEM OF TIME DEPTH. We have followed the procedure recommended by Hymes (1960:29) of taking the greatest indicated divergence time for any two Malagasy speech communities as the best estimate for the divergence time for the subgroup (the Malagasy languages) as a whole. The smallest percentage of shared cognates for any two Malagasy speech communities is that of Sakalava 2 (west coast) and Tanalana (east coast). These two groups share 52 percent cognates. This would indicate an approximate divergence time of

2108 B.P. ($\pm$ 304 for the 68 percent confidence interval) using Gudschinsky's procedure and 3087 B.P. ($\pm$ 435 for the 68 percent confidence interval) with Hattori's modification. We consider both figures to be unreliable for the following reasons. First of all, such a contrast of Sakalava 2 and Tanalabeaka involves one speech community's (Sakalava 2's) special vocabulary. Secondly, Sakalava (both 1 and 2) appear to have been influenced by taboo to a greater degree than the other languages we are concerned with. As noted earlier, the effect of taboo, as with special vocabularies, would tend to stretch the time-depth estimates. Finally, although this factor was in no sense determinative, we could not help being aware that such an estimate as the one using Hattori's modification (3087 B.P. $\pm$ 435) is considerably earlier than estimates based on other techniques of the time at which the ancestral Malagasy diverged from other Indonesian groups.

In view of these considerations, we have chosen to base our calculation on the smallest percentage of shared cognates of two Malagasy dialects neither of which is the special vocabulary list from a group with dual vocabularies. In this case the Tsimihety (Tsi) dialect of the northern interior of the island shares only 61 percent of its basic words as cognate with one other Malagasy speech community—Vezo (Vz), a group of marine fishermen located on the northwest coast near the city of Tuléar. Using the figure of 61 percent for the shared cognates, we come up with two 68 percent confidence level estimates for time depth: 1639 B.P. $\pm$ 235, and 2641 B.P. $\pm$ 364.

We must now decide which of these estimates is the more credible. The problem is a difficult one, and there can be no acceptable solution without at least some ideas of which procedure is the more likely to produce an accurate estimate. The calculation is carried out at the 61 percent level of cognates shared. This appears to be sufficiently low that at least some of the retarding factors will no longer be working on the rate of divergence. There is, however, no evidence which would suggest that these geographically separated populations have been in direct contact. On the other hand, were the populations in contact communication would still be possible with 61 percent cognate sharing. Furthermore, the effect of drift and drags may still be expected to influence the rate of retention. We therefore think it justified to conclude that the actual divergence date is somewhere between the two estimates we have, and if it leans one way or the other, it is more likely to be closer to the estimate produced with Hattori's modification.

At this point we note that using the 90 percent confidence intervals for the two estimates, we come up with an area of overlap which runs from 91 B.C. to A.D. 226. In accordance with the above statements we

thus conclude that the actual time of first occupation of the island took place at some time in the first century A.D. and, with less probability, as early as the fourth century B.C. or as late as the fourth century A.D.

This conclusion does not offer much support for Dahl's hypothesis that the ancestral group which eventually diversified to form the present-day Malagasy population left Indonesia at some time around A.D. 400. Because our result appears to differ by at least 550 years from Dahl's, the basis for his estimate deserves examination here.

Dahl has offered the hypothesis that the Proto-Malagasy, the Indonesians who left Borneo eventually to settle Madagascar, emigrated from their homeland at approximately A.D. 400. He infers this on the basis of evidence relating to the dating of the beginning of Indian influence in Borneo. Contacts between India and Borneo are assumed to have begun around the first or second centuries, A.D. (Dyen 1953: 300). There are certain Indian loanwords in Malagasy. Since there is no evidence to suggest that there has been significant direct contact between India and Madagascar, the Proto-Malagasy must have left Indonesia after Indian influence had begun to be felt. Furthermore, all the Sanskrit loanwords in Malagasy are, with a single exception, found in other Indonesian languages. Therefore, the contact of the Proto-Malagasy and Indians must have occurred while the Proto-Malagasy were still in the Indonesian, and probably the Borneo, region. The oldest inscription testifying to Indian influence on Borneo has been dated at ca. A.D. 400. Because of the relatively few words of Indian origin in Malagasy languages, when compared to other Indonesian speech communities, however, Dahl assumes, and the present authors agree, that the Proto-Malagasy must have left Indonesia sometime around the beginning of Indian influence. Thus the date for the exodus of the Indonesian component of the Malagasy population from Borneo (or some other Indonesian location where they and the ancestors of the Maanyan had been in contact with Indians) must have been, according to Dahl, around A.D. 400.

Our estimate for the beginning of divergence of the Malagasy dialects is seen to be ca. A.D. 50, some 350 years earlier than Dahl supposes. It is interesting to note that this falls precisely on the estimate produced by Dyen (1953: 301) for the separation of Merina (Mc) and Maanyan (an Indonesian collateral). Unfortunately, we shall have to discard this pleasing coincidence of time depths as merely a chance result, because we feel that Dyen's estimate is too recent. This is primarily because (1) at least some time must have been needed for the Proto-Malagasy to journey as traders from Indonesia along the coasts of India and northeast Africa to Madagascar, settle the island, and spread

out geographically in numbers sufficient to establish the basis of different speech communities and (2) Dyen's calculations assume an 81 percent retention rate for the 200-word list, with 45 percent cognates shared by Merina and Maanyan. Had Dyen found 45 percent cognates shared on a 100-word list and then used an 86 percent per millennium retention rate, he would have found a 68 percent confidence interval estimate of 331 B.C. to 1027 B.C. using Gudschinsky's formula. Any attention to Hattori's modification, of course, would have thrown the estimate much earlier. In fact the overlap of the 90 percent confidence intervals for the Gudschinsky and the Hattori estimate is 997 B.C. to 1251 B.C., fully a thousand years earlier than Dyen's present estimate. We see then that Dyen's estimates support our argument for a date of departure earlier than A.D. 400. In addition, we have reasons for believing that the true time of departure must have been somewhat earlier than the time during which the Malagasy languages began to diverge. As a result, we conclude that the departure of the first Proto-Malagasy readers may be perhaps on the order of five hundred years earlier than Dahl believes.

The five hundred year discrepancy between our estimate and Dahl's can be explained if (1) the glottochronological analysis we have adduced is deceptive and the less likely estimate obtains, (2) Hindu influence arrived earlier than is supposed or diffused faster, (3) the population ancestral to both Malagasy and Maanyan lived in an area of Indonesia other than Borneo where Hinduization occurred earlier than in Borneo, after which the Maanyan settled Borneo and the Malagasy Madagascar, (4) considerable dialect divergence occurred especially rapidly in the widespread Proto-Malagasy trading community, and the island of Madagascar was settled by a speech community with dialect divergence already begun, (5) any combination of these factors. Lacking any specific evidence of this type that would lead us to believe that our estimate is completely incorrect, we conclude that the results of our glottochronological study suggest that the population which gave rise to the sixteen speech communities included in our sample was a single people at a time depth from about 150 B.C. to about A.D. 225 and almost certainly not later than A.D. 400. It would seem, therefore, that acceptance of Dahl's estimate should be held in abeyance pending further research and that for the present it should be considered as more likely too recent a date than too extreme. Further information that would help resolve the disagreement includes (1) the discovery of more cognates from among the Malagasy and Indonesian languages than Merina and Maanyan and (2) the finding of inscriptions testifying to Indonesian influence on Borneo older than ca. A.D. 400. The former, of

course, would support Dahl, while our estimate would receive support from the later discovery.

IV. SOUTHERN INDONESIA AND MADAGASCAR. If the indicated divergence time for Proto-Malagasy and Proto-Maanyan is recalculated for a Hadwen list at 1.86 percent retention rate, then even the 99 percent confidence interval produces an indicated time depth for divergence of more recent than 107 B.C. Assuming that the time of first substantial settlement of Madagascar was A.D. 0, we can conclude that there were at least 150 years during which the Malagasy were one speech community. It should be noted that such a conclusion is independent of any inaccuracies of analysis in this paper. If we are wrong about both Dyen and Dahl, and their time estimates are eventually correct, we still have a difference of about 260 years during which the Malagasy were a single speech community, even longer if we are correct about Dyen's estimate.

Where was the time spent during the period following the separation from the Proto-Maanyan and prior to the splitting internal to the Malagasy subgroup? We assume (following Dahl 1965) that the Proto-Malagasy were a trading population who participated in a west Indian, and possibly even Pacific Ocean trade network which tied Indonesia to points east and west. To the west, the Proto-Malagasy traders journeyed through Indian, Arabian, and East African ports and eventually arrived on the coasts of Madagascar. Along the East African ports of call, the predominantly male traders took wives, and their children, tradespeople and brought contributions from African and Indonesian gene pools which have given the present-day Malagasy population its tremendous phenotypical diversity. The beginning of the divergence of Malagasy speech communities occurred through settlement at different areas along the trade route in Madagascar itself. One can imagine Proto-Malagasy stopping over at different ports between Indonesia and Madagascar. Some remained closer to the Indonesian homeland, while others moved further south or to Madagascar itself. Some descendants of the stopovers, diversifying groups, may have even migrated to Madagascar later, probably when Arab domination of East African ports forced Indonesian traders out. More will be said about this following the discussion of Malagasy subgroups.

V. SUBGROUPING. Because of the arbitrariness of the subgrouping procedure and choices in percentages of shared cognates at which classification is made, our discussion of subgrouping based solely on numeri-

cal criteria of the shared cognate matrix can be tentative at best. Still, it is possible to make a number of reasonable generalizations and conclusions based on the matrix.

Chart 1 suggests that the Malagasy dialects broke up into a relatively discrete language, Proto-Tsimihetsy (Tsh), and a language ancestral to all the other languages in our collection. In addition, we can isolate a group of languages which have resulted from a relatively high degree of linguistic fragmentation. These are Me, Si, BtsA, BtsF, Bsk, Tm, Tsa, Zfs, Vz, Mf, Td_1, and Bo. The other languages, Tbk, Tsi, Sak_1, Sak_2, and Tdy_2 are, of course, closely related to the group just isolated but at a somewhat more remote and indeterminate position.

Given the arbitrariness of procedure mentioned above, it seems advisable to try to check the degree of variation of the genealogical chart given different sets of levels. Let us examine levels separated by 5 percent, but beginning this time with 94 percent. Chart 2 summarizes

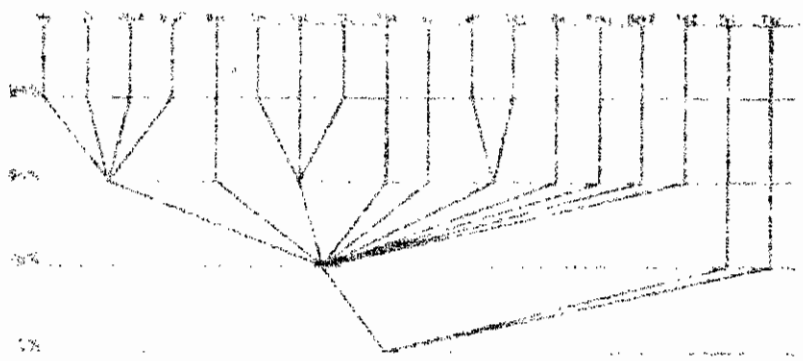


CHART 2. *Proto-age of Malagasy dialects based on cognate percentages as arbitrarily determined levels, levels at 94%, followed by 5% intervals.*

or two from those within the classification. Once again we observe that Tkr is isolated from the other languages. However, this time Tsimihetsy (Tsh) is similarly isolated. The tree continues our statements about a group in which linguistic consolidation has been more intense, but in this case the group includes only Me, Si, BtsA, BtsF, Tm, Tbk, Zfs, Mf, and Td_1. The languages Bsk, Vz, and Bo are seen to be slightly less well connected to the larger group.

Chart 2 raises the question of the relationship of Tkr, Tsi, and the other Malagasy languages. We believe that there are good reasons, based purely on the cognate matrix, for considering both Tkr and Tsi as relatively discrete. First of all, we note that Tsi shares at most 77 per-

cent of its word list with any other Malagasy language. The language with which it shares 77 per cent cognates is Bsk, a language which itself shares relatively few cognates with the other languages in our list (58 percent with Sih). Tkr is even more remotely related in that it shares no more than 73 percent cognates with any other language of the set. The language with which it shares 75 percent of the word list is Tsi, indicating a link between these two languages closer than that of Tkr with any of the other dialects. Tsi is thus a linking dialect between Tkr and the other languages.

The extent to which Tkr and Tsi are distinct can be gauged to a degree by searching the list for the language with the next lowest number of maximum shared cognates with other languages of the set. This proves to be Tbk, which shares at most 81 percent of its word list with Tsi. That Tbk is not as discrete as either Tsi or Tkr is indicated by (1) the fact that its maximum percentage of shared cognates with any other languages is fully four percentage points higher than that of Tsi, and six percentage points higher than that of Tkr, and (2) the language with which it shares 81 percent of the word list is Tsi, a language well connected with the others of the set, sharing, for example, 83 percent of its word list with Tsk and Zfs.

Despite this, even Tbk proves to be isolable, because a search for the language with the next lowest number of maximum shared cognates fails to turn up any more than three percentage points higher: we find Bsk and Vzt, both of which share no more than 86 percent of their word lists with other languages of the set. This ignores, of course, the existence of Sakalava, which is made indeterminate by the presence of taboo words. The relative isolability of Tbk here serves to re-emphasize the extent to which Tsi and Tkr are discrete. The attempt to further isolate languages by this procedure of searching for minimum maximum (minimax) percentages of shared cognates proves fruitless, because at this point it is apparent that we are discussing a very highly interrelated group of dialects.

Still another genealogical tree can be constructed by taking intervals of 6 percent. This tree, depicted in Chart 3, makes the isolation of Tbk explicit, as well as isolating Tkr and Tsi once again. As on the earlier charts, Mr, Sj, BtsA, BtsP, Tsv, Tsk, and Zfs are identified as speech communities whose ancestral dialects have undergone a relatively high degree of diversification and fragmentation when compared with other languages of the set. Mf and Tdr, are shown to be possibly less closely linked with the languages that had undergone the intensive diversification.

Our analysis of the subgroupings of the Malagasy dialects has been

based on two different procedures. First, we examined the lower parts of the genealogical trees to see whether any languages were isolated from the bulk of the languages. When certain languages were separated out, the matrix of shared cognate percentages was examined to check further the reasonableness of the trees. Secondly, beginning from the top ends of the trees, we sought out language groups that appear to have descended from single languages that had undergone a relatively

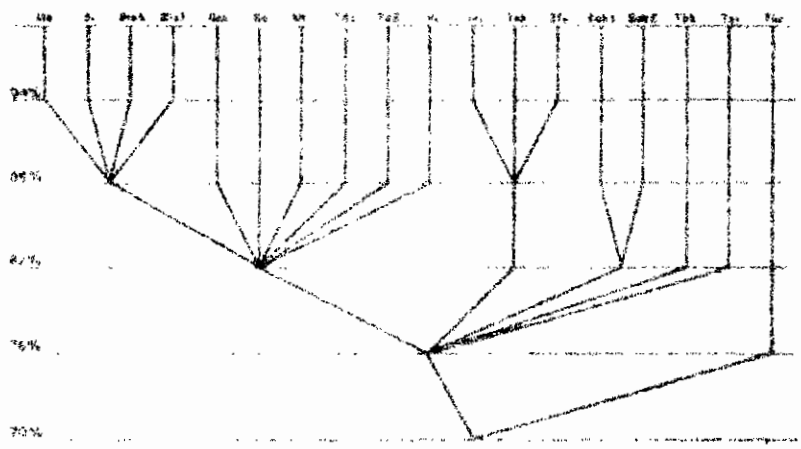


CHART 3. Branchings of Malagasy dialects based on cognate percentages at arbitrarily determined levels (levels at 5% intervals).

high degree of linguistic diversification in Madagascar. Based on this analysis, we have come to the following conclusions:

1. Tkr is quite isolated from the other languages of the Malagasy Republic. Its closest collateral is Tsi.

2. Tsi is also distinct, but somewhat less so than Tkr.

3. An additional language found to be relatively apart from the rest of the dialects is Tsk.

4. Me, Si, BtsA, BtsF, Tnr, Tsk, and Zfs are all speech communities descended from languages which have undergone considerable linguistic diversification.

5. Associated with the languages listed under (4), but somewhat less certainly related to the highly diversifying protolanguages, are Mf and Tdl_1.

6. Similar to the languages listed in (5), but more remotely associated, are Bsk, Vz, and Ba.

7. Still further removed from the group of languages we consider as

"recently diversified" are Sak_1, Sak_2 and of course Tsi and Tbk (see (2) and (3)).

8. Our languages that appear descended from highly diversifying ancestors do not all constitute one group. In fact, they are divided into two groups that are evident on all the genealogical trees. These are Me-Si-BtsA-(Bsk?) and Tm-Tsk-Zfs.

9. Chart 1 also denies an additional protolanguage that diverged from the Me-Si-Bsk-Bsk$_1$ or Tsk-Zfs group before this latter group began splintering. This is the Vz-Mf-Tdr$_1$-(Tdr$_2$)-Ba dialect family.

It seems only reasonable to be highly cautious in an attempt to make estimates for the time depths of the languages spoken within Madagascar. We shall, in fact, attempt to determine time depths for only a few of the language groupings and shall attempt to place other dialects into the picture by inference from the genealogical trees. As with our determination of the time depth of the differentiation of the entire Malagasy dialect group, we follow Hymes (1960, 29) in basing our estimated time depth of a group as a whole on the calculations for the least related languages in the group.

Calculating purely by Hacnel's formula we find that both groups mentioned in (8) above began diversifying at around the same time. The Me-Si-BtsA group has a time depth as indicated by its lowest linked pair (Bs percent for Si-BtsA) at A.D. 1167 to A.D. 1441 (A.D. 1254 ± 187) for the 98 percent confidence interval. Bsk' split from this group slightly earlier and Bsk perhaps slightly earlier still. The Tm-Tsk-Zfs group has a time depth indicated by its lowest linked pair (68 percent for Tm-Zfs) as A.D. 1191 to A.D. 1535 (A.D. 1363 ± 172 for the 68 percent confidence interval). In an attempt to find the most likely time for divergence of these two groups, we seek the pair of languages with the lowest shared cognates from among them. This proves to be Bsk and Zfs, with only 45 percent sharing of cognates. The estimated time depth of the divergence of the two subgroups we are concerned with is therefore A.D. 540-870 (A.D. 705 ± 265 for the 98 percent confidence interval). Furthermore, we might inquire about the divergence time of the Vz-Mf-Tdr$_1$-(Tdr$_2$)-Ba group. Continuing with the method used above, we find the linked pair with the lowest shared cognates to be Vz and Ba, with 27 percent common cognates on the word list. This indicates divergence sometime from A.D. 470 to A.D. 980 (A.D. 729 ± 262 for the 68 percent confidence interval). This means that Chart 1 is deceptive in that it implies that Vz, Mf, Tdr$_1$, and Ba were a single language more recently than were Me, Si, BtsA, Tm, Tsk, and Zfs. In fact, closer analysis now shows them all to have begun diverging at

about the same time. During the second part of the first millennium A.D. In this respect Chart 3 is more tentative.

At this point it seems worthwhile to begin to assemble the data into some coherent picture of linguistic (and hopefully nonlinguistic) prehistory in Madagascar. For reasons presented earlier, we conclude that the first occupation of the island took place sometime around the first century A.D. The population moved southeastward over the uninhabited island and almost immediately diverged into three parts: proto-Tsi, ancestral-Tsi, and the protolanguage for the remaining languages of the Malagasy Republic (cf. Chart 4). Slightly later ancestral-Tsi may have become

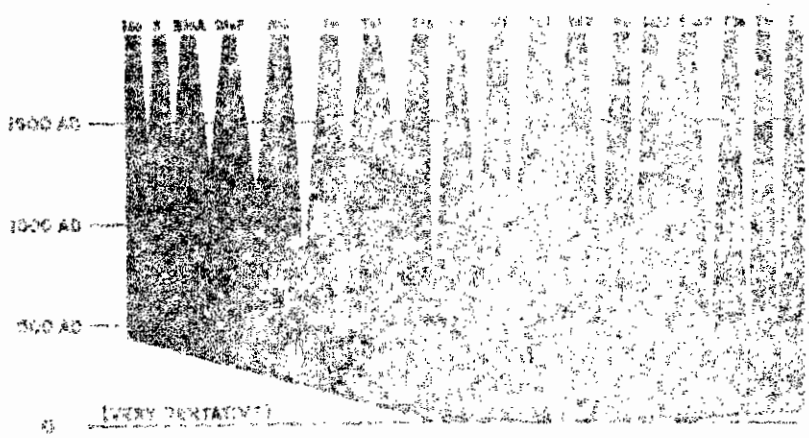


CHART 4. Suggested chronology of Malagasy dialect split

evident. Perhaps during the middle of the first millennium the proto-language divided into at least two groups, ancestral-Ne-Si-Bet-Sak-Tm-Tsk-Zfs and ancestral-Ve-Bf-Tdj-Bs. This last group began dividing during the second half of the first millennium A.D. Ancestral Sak may also have become identifiable at this time. While ancestral-Ve-Bf-Tdj-Bs was splitting into many parts, ancestral-Ne-Si-Bet-Sak-Tm-Tsk-Zfs was dividing into two parts. Further divisions in both of these latter two parts took place during the first half of the second millennium A.D.

14. GLOTTOCHRONOLOGICAL RESULTS COMPARED WITH PREVIOUS STUDIES OF MALAGASY DIALECTOLOGY. Dez (1963) has classified Malagasy speech communities into subgroups on the basis of phonology and morphology. His subgroups differ from those related through our glottochronological analysis, and we suggest, therefore, that Dez's

conclusions be reconsidered. Dez found two significant subdivisions among the Malagasy languages. The first includes the dialects of the west and south, and the second, those of the east coast and central highlands. The following are the most representative members of Dez's western-southern subgroup: Vezo (Ve), Bara (Ba), Mahafaly (Mf), Tandroy (Td), and Sakalava dialects (Sak) of Menabe and Boina (Menabe and Boina refer respectively to the southern and northern centers of the Sakalava kingdoms of the eighteenth century). The eastern-central group includes such dialects as our Me, BtsA, Si, Tra, Zfs, and Tsk. Dez regards some speech communities as intermediate between the major subgroups. Thus, he interprets northern Sakalava (which is unrepresented in our basic vocabulary lists), Tsimihety (Tsi), Tankarana (Tkr), and northern Betsimisaraka (Bsk) as members of a definable subgroup of the first group (the western-southern group). In addition, he points out that Betsileo spoken around the region of Fianarantsoa (BtsF) may be regarded as intermediate between the first and second group. Other dialects, including Tanosy of the extreme southeast, unrepresented in our lists, are also considered to be intermediate between the two major groups. While the major Malagasy subgroups isolated by Dez differ from the three we have isolated, there is much in Dez's analysis which suggests confirmation of our own results. Dez argues, for example, that even though the primary division in Madagascar seems to be between a western-southern and eastern-central group, one might also want to recognize the existence of a third group that would include the northern Malagasy dialects of Tsi and Tkr.

Our findings show that the northern dialects are indeed to be distinguished from both western-southern and eastern-central subgroups. However, not only are Tkr and Tsi sufficiently divergent to be placed apart from the western group, they are also sufficiently distinctive from one another in basic vocabulary to warrant placement in two subgroups rather than in a single subgroup. The findings of glottochronology reveal that the original diversification of Proto-Malagasy into subgroups involved three divisions, Proto-Tsi, Proto-Tkr, and proto-all the rest of the languages spoken in Madagascar today. It was later in Malagasy history that what Dez considers to be the principal division, that between eastern-central and western-southern, took place. Thus, the distinction between Proto-Tkr, Proto-Tsi, and all other Malagasy dialects which our findings suggest is not foreshadowed in Dez's work.

This is perhaps because Dez has used shared innovations in phonology only to distinguish members of the eastern-central subgroup. No shared innovations which would confirm the unity of origin of what Dez calls the western-southern group are presented. In fact, most of

the key phonological elements which Dez presents as evidence of a unity of origin for Tkr and Tsi, on the one hand, and languages such as Ba, Td, Va, and Mf, on the other, are shared retentions, from a common Proto-Indonesian parent language. Such characteristics alone cannot reveal valid subgroupings.

There is, in fact, no phonological nor morphological reason to doubt our results, which show the divergence of proto-western-southern from proto-eastern-central as having occurred later than the divergence of Proto-Tkr, Proto-Tsi, and proto-all the rest.

The principal implications of our results may now be indicated. During almost the entire time that Madagascar has been settled, there has been a division of speech patterns into three communities. One of the three groups underwent more rapid diversification than the other two, ultimately giving rise to the split which Dez makes paramount. This split, between the ancestors of the western-southern group and those of the eastern-central group, took place between A.D. 600 and A.D. 700 (A.D. 340–892 at the 68 percent confidence level). At this point only does it become meaningful to speak of a contrast between western-southern and eastern-central groups. Subsequent to this split, the ancestral western-southern group immediately began further linguistic diversification, reflecting population movements in accordance with principles to be indicated below. The members of the eastern-central group remained together longer, but eventually (ca. 1300, or A.D. 1067–1535 at the 68 percent confidence level) began to diversify into eastern coast and central highlands subgroups. The events above are the ones which have given rise to the unity and diversity observed within the Malagasy languages today.

15. CONCORDANCE OF GLOTTOCHRONOLOGICAL RESULTS WITH OTHER FACTS FOR INTERPRETATION OF MALAGASY PREHISTORY. There is nothing incompatible between our findings and interpretations of Malagasy prehistory which have been independently derived. In fact, the glottochronological findings suggest a framework for Malagasy history which seems to us more plausible than any previously suggested.

The subgroupings which have been isolated through glottochronology correspond in large measure to distinctive kinds of cultural adaptation in Madagascar. To substantiate this statement, we shall begin with recent cultural and linguistic differentiation and move back through time. The societies which belong in the central subgroup, for example, may be viewed as representatives of a single cultural adaptive radiation. Their distinctive adaptive mechanism is wet-rice cultivation through irrigation. The societies of the central highlands, which in-

clude Merina (Me), Sihanaka (Si), and Betsileo (Bts), are all distinguished from other Malagasy by the presence of this irrigation economy. The societies of the central highlands are geographically close to one another. All of them were chiefdoms or states, with the basis for this level of political organization to be found in their irrigation systems and their productive agricultural economies. Members of these three groups often claim a single origin. It is common, for example, for Betsileo to trace their ancestry either to the Tananarive region, which was the aboriginal Merina capital, or to the east coast, the country of the Tainimro (Tm). Merina also trace their origin to the east coast. The linguistic unity of this recently emerged subgroup corresponds, then, with cultural and ecological unity.

On a higher level of inclusion, the unity of the eastern-central as opposed to the western-southern subgroups has been seen from glottochronology, with the split between the ancestors of the two populations having taken place sometime around A.D. 600–700. All the members of the eastern-central group share a common adaptation consisting of the cultivation of wet rice as their principal economic resource. The eastern subdivision of this group, however, relies on rainfall rice agriculture, while irrigated rice is cultivated in the central highlands. The divergence of the central populations from the eastern-central group took place around A.D. 1300. This involved movement of the ancestors of the Merina, Sihanaka, and Betsileo from the east coast, most probably from the central and southern areas of the east coast into the interior. The ancestors of the present-day representatives of the eastern-central group were probably participants in trade networks along the Malagasy coasts (these trade networks possibly extended to the East African coast).

Some of the members of the western-southern subgroup are likely to have left the coasts for the interior at an earlier date, perhaps as early as A.D. 600 or 700. Their adaptation became one of cattle pastoralism in the Malagasy interior. They spread through the south, western interior, and the central highlands of Madagascar. They were free to radiate in this region unimpeded by competing populations for several hundred years. However, with the movement of the central members of the eastern-central subgroup to the highlands around 1300, the pastoralists were eventually pushed south and west, displaced to the most arid regions of the island by a more productive economic base, irrigated wet-rice cultivation. It is in these regions that the pastoral peoples, the Bara, Mahafaly, Tandroy, and interior Sakalava, remain today. The displaced pastoralists would be equated with the legendary Vazimba of central Madagascar. Merina and Betsileo legends speak of wars

between their ancestors and a group which inhabited the central highlands region on their arrival from the coast. Who were these Vazimba, despite all the efforts to make them Black Africans, representatives of a pre-Malagasy substratum in the island, other than collaterals descended from the original population which eventually diversified to form the western-southern and eastern-central subgroups? They were members of the ancestral western-southern group who had filled in unoccupied niches in the interior and were eventually displaced by the expanding wet-rice cultivators.

One additional correlation between the results of glottochronology and other interpretations of Malagasy prehistory may be mentioned. In Malagasy oral tradition, the Tambahoaka (Tbk), a small group located in and near the town of Mananjary on the south central east coast, traces its ancestry to Arab migrants to Madagascar. Deschamps (1965:51) has placed this migration at some time around the thirteenth century A.D. Physically, the Tambahoaka look like other inhabitants of the east coast, and the language they speak is unquestionably Malagasy. Certain of their cultural practices do suggest, however, some Arab influence. Scholars have been uncertain about how to interpret the Tambahoaka. The results of our study suggest that the Tambahoaka do occupy a relatively distinctive place in Madagascar. Eighty-one percent is the maximum of cognates shared by Tambahoaka and any other Malagasy dialect. This would suggest that the modern Tambahoaka are the outcome of a period of separate history for approximately seven hundred years. We suggest that the Tambahoaka may be representatives of a population which has been relatively isolated for at least seven hundred years from other Malagasy groups, and which, during this period of isolation, had greater contact and intermixture with Arabicized groups than other Malagasy populations of the east coast. We have no independent information which would enable us to comment on whether this isolation and contact with Arabs took place in the present-day location of the Tambahoaka. Mananjary might in the past have been a coastal Malagasy port frequented by Arabic traders. The isolation and separate history could also have begun elsewhere, before the Tambahoaka moved to their present location, possibly from the north.

The major and the oldest subgroupings of Malagasy speech communities must now be explained. It has been stated above that the earliest inhabitants of Madagascar are likely to have been traders. These people, speaking an Indonesian language, reached Madagascar by following a trade route along the East African coast, where they mixed with Africans culturally and genetically. Early in the history of

the Malagasy, the ancestral group began splitting into the three populations which eventually gave rise to three linguistic groups of equivalent status—Tankarana (Tkr), Tsimihety (Tsi), and all others.

The Proto-Tsimihety and the Proto-Tankarana settled in the north of Madagascar, while members of the third group established trade posts and settlements on the east and west coasts. The present-day Tsimihety (population: 428,000 in 1964) inhabit the northern interior of the island. They are the only Malagasy population to rely on an economic base which mixes cattle pastoralism with shifting cultivation. The Tankarana are a small group (42,000 in 1964) who inhabit the northern extremity of Madagascar, isolated from all other Malagasy groups, including their nearest neighbors, the Tsimihety, by the mountain range of Tsaratanana.

Both of these populations, if our results are accurate, have been separated from all other Malagasy for a considerable period of time. How is this separation to be explained? While we have no definitive answer to the problem, some suggestions will be made. Future archaeological investigations may place us in a better position to assess these speculations. On the one hand, it is possible that the Tsimihety and the Tankarana are relative newcomers to Madagascar, their differentiation from other Proto-Malagasy having occurred at ports of call on the East African coast. From such ports, these mixed African-Indonesians may have been pushed by the expansion of Islam. They may have been expelled from East African ports by the monopoly on trade with East Africa which Islamized Arabs established around A.D. 1000. They may simply have moved to Madagascar, settling the still unoccupied northern regions of the island, where their isolation from other Malagasy groups could have been maintained.

On the other hand, it is perhaps more likely that the Proto-Tsimihety and the Proto-Tankarana have inhabited Madagascar as long as any of the other Proto-Malagasy groups. Their occupation of the isolated northern region of the island could have been early, while members of the third group were in the process of settling the other coasts. The physical environmental limitations to contacts between northern Madagascar and the other parts of the island would have insured that this original isolation would have been maintained.[9]

The Tsimihety, in adapting to an interior ecological niche in an arid region north of the central highlands through their mixed economy, would never have been threatened by the wet-rice cultivators who pushed the pastoralists out of the central highlands, to the south. In Malagasy tradition, the Tsimihety are considered to be a fiercely independent group, one of the few in the island which was never

completely subjugated by the expanding political control of one of the Malagasy kingdoms. The only exception to the general independence of the Tsimihety was the result of a loose inclusion in the Sakalava empire at the time of its maximal expansion in the eighteenth and nineteenth centuries. Isolation from potentially competing groups seems the best explanation for the early divergence of the Tsimihety from other Malagasy populations.

Little is known of the 42,000 Tankarana. 'People of the rocks' is the literal translation of the name Tankarana into English. Several times during the eighteenth and nineteenth centuries the Tankarana were conquered, nominally, by the Sakalava, Merina, and French. Their earlier contacts with outsiders seem to have been limited to Islamic groups of traders known as Antalaotse and Iharanians (Deschamps 1965:102), who had established a few ports on the northern coast. Geographical factors which did not allow them much contact with other Malagasy would explain the long isolation of the Tankarana from other Malagasy groups.

16. SUMMARY AND CONCLUSIONS. All members of the present-day population of the Malagasy Republic share a linguistic unity of origin in the Indonesian traders who reached Madagascar along a maritime trade network. The likely date for the first settlement of Madagascar would seem to be ca. B.C./A.D. At least by this time linguistic differentiation among the ancestral population had already begun. Madagascar was colonized by three groups. One of them became the Tsimihety of the northern interior. Their isolation was maintained by the fact that the resources which they exploited were of no interest to any other Malagasy or foreign group. As an interior population, they did not compete in the coastal trade routes. A second became the Tankarana, isolated from their nearest neighbors, the Tsimihety, and from all other Malagasy groups by a mountain range, and eventually isolated from trade by a relatively unproductive interior and successive competition from Arabs, Europeans, and later other Malagasy, including the Sakalava and the Merina.

A third group which has subsequently diversified to form all the rest of the contemporary Malagasy populations established itself along the eastern and the western coasts. Sometime around A.D. 700 a western group began to lose contact with the eastern group, and differentiation into a western-southern as opposed to an eastern-central group began as some of the inhabitants of the west coast journeyed to and took up residence in the interior as cattle herders. Later, perhaps around A.D. 1300, some members of the eastern group left the east coast and moved

to the interior highlands where they began irrigated cultivation of wet rice. In the course of their adaptation to a highlands environment, through this efficient economic base, they came into conflict with pastoralists who had lived there previously. Eventually, productivity and political organization of the irrigation-based groups became sufficiently advanced so that the pastoralists could be both absorbed and pushed out of the highlands into the more arid zones, not cultivable even by Malagasy irrigation techniques, of the south and west. In the meantime, other members of the western group, Sakalava and Vezo, had remained on the coast, where the former continued to specialize in trade. After European contact and introduction of the slave trade, the Sakalava eventually became a widespread and powerful state of the eighteenth and early nineteenth centuries. The Vezo, on the southwest coast, achieved a less ambitious adaptation as marine fishermen.

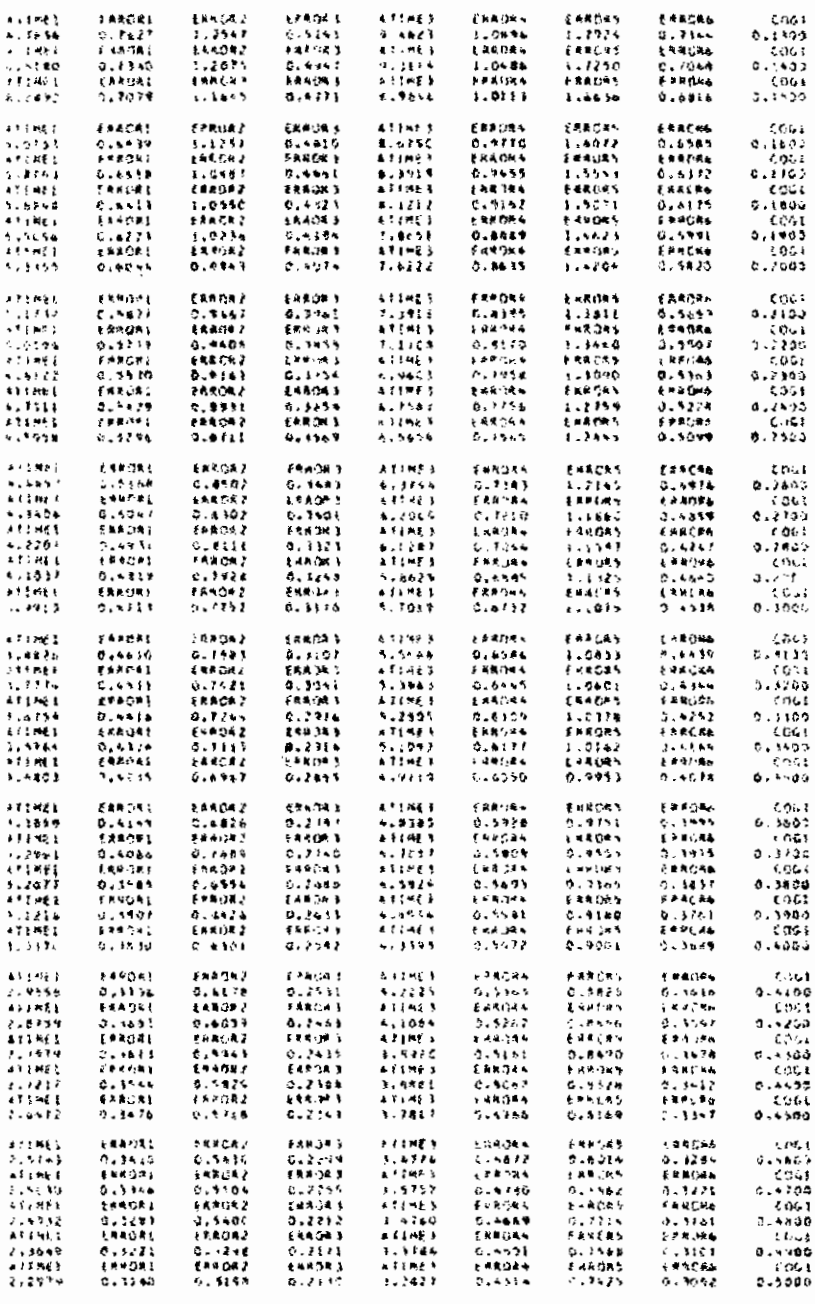
Inhabitants of the east coast also continued to engage in maritime trade with Arabs and Europeans. Trade states, such as that of the Betsimisaraka (Bsk), had formed on this coast by the eighteenth century.

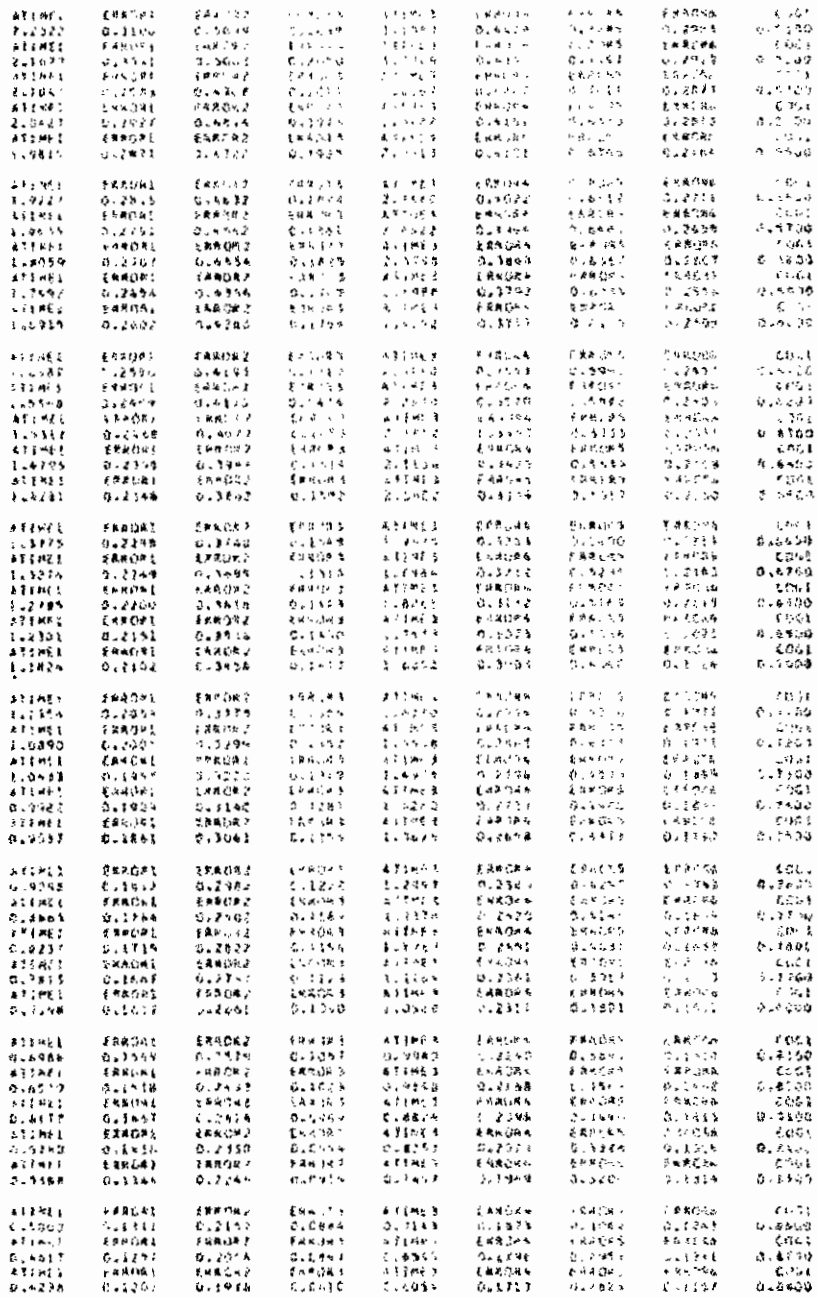
Glottochronology has contributed to our understanding of Malagasy culture history, confirming the genetic relationship and unity of all Malagasy speech communities and showing the relative time depths and the nature of divergences which have given rise to the different kinds of cultural adaptations encountered in Madagascar today.

Appendix

Using the following simple computer program (written in FORTRAN IV for the IBM 7094 computer) a table can be generated giving lexicostatistical information that is desirable in analyzing matrices of shared cognate percentages. The table gives time depths, both for Gudschinsky's and Hattori's formulas, and error estimates for both methods at the 68 percent, 90 percent and 50 percent levels of confidence.

```
FINITIONS FOR VARIABLES AND CONSTANTS

  ATIME1   ESTIMATED TIME DEPTH IN MILLENIA
  ATIME2   ATIME1 PLUS ONE STANDARD ERROR IN MILLENIA
  ATIME3   ESTIMATED TIME DEPTH IN MILLENIA WITH ARITTORI MODIFICATION
  ATIME4   ATIME3 PLUS ONE STANDARD ERROR IN MILLENIA
  COG1     COGNATES REPORTED
  COG2     COGNATES PLUS ONE STANDARD ERROR
  CONST    RR CONSTANT
  SIGMA    STANDARD ERROR OF COGNATE ESTIMATE
  SAMPLE   WORD LIST LENGTH
  ERROR1   68 PERCENT CONFIDENCE VALUE FOR ATIME1
  ERROR2   90 PERCENT CONFIDENCE VALUE FOR ATIME1
  ERROR3   50 PERCENT CONFIDENCE VALUE FOR ATIME1
  ERROR4   68 PERCENT CONFIDENCE VALUE FOR ATIME3
  ERROR5   90 PERCENT CONFIDENCE VALUE FOR ATIME3
  ERROR6   50 PERCENT CONFIDENCE VALUE FOR ATIME3

  WRITE (6, 50)
0 FORMAT (31H LEXICOSTATISTICAL CALCULATIONS//)
  READ (5,10) CONST, SAMPLE
0 FORMAT (2F10.2)
  COG1 = 0.01
0 ATIME1 = ALOG10 (COG1) / (2. * ALOG10 (CONST))
  SIGMA = SQRT ((COG1 * (1. - COG1)) / SAMPLE)
  ATIME2 = ALOG10 (COG1 + SIGMA) / (2. * ALOG10 (CONST))
  ERROR1 = ATIME1 - ATIME2
  ERROR2 = ERROR1 * 1.645
  ERROR3 = ERROR1 * .674
  ATIME3 = ALOG10 (COG1) / (1.4 * ALOG10 (CONST))
  ATIME4 = ALOG10 (COG1 + SIGMA) / (1.4 * ALOG10 (CONST))
  ERROR4 = ATIME3 - ATIME4
  ERROR5 = ERROR4 * 1.645
  ERROR6 = ERROR4 * .674
  WRITE (6, 60)
50 FORMAT (10H    ATIME1    ERROR1    ERROR2    ERROR3    ATIME3    E
 2RROR4    ERROR5    ERROR6    COG1//)
  WRITE (6,70) ATIME1, ERROR1, ERROR2, ERROR3, ATIME3, ERROR4, ERROR
 25, ERROR6, COG1
70 FORMAT (9F10.4)
  COG1 = COG1 + .01
  IF (COG1 .LE. 1.0) GO TO 20
  STOP
  END

$DATA
LEXICOSTATISTICAL CALCULATIONS

 ATIME1    ERROR1    ERROR2    ERROR3    ATIME3    ERROR4    ERROR5    ERROR6     COG
15.2665    2.2696    3.7[illegible]    1.[illegible]    21.[illegible]    3.7208    5.3605    3.2055    0.010
 ATIME1    ERROR1    ERROR2    ERROR3    ATIME3    ERROR4    ERROR5    ERROR6     COG
12.[illegible]    1.7591    2.[illegible]    1.1936    [illegible]    2.[illegible]    4.[illegible]    1.6938    0.020
 ATIME1    ERROR1    ERROR2    ERROR3    ATIME3    ERROR4    ERROR5    ERROR6     COG
11.6296    1.[illegible]    2.[illegible]    [illegible]    16.[illegible]    2.[illegible]    3.5673    1.4320    0.030
 ATIME1    ERROR1    ERROR2    ERROR3    ATIME3    ERROR4    ERROR5    ERROR6     COG
10.6710    1.5258    2.5149    [illegible]    15.2466    1.0447    1.[illegible]    1.[illegible]    0.040
 ATIME1    ERROR1    ERROR2    ERROR3    ATIME3    ERROR4    ERROR5    ERROR6     COG
 9.9313    1.3996    1.9736    0.8024    15.[illegible]    1.7[illegible]    2.8385    1.1568    0.050

 ATIME1    ERROR1    ERROR2    ERROR3    ATIME3    ERROR4    ERROR5    ERROR6     COG
 9.3249    1.1035    1.8136    0.[illegible]    13.[illegible]    1.[illegible]    2.5445    1.0445    0.060
 ATIME1    ERROR1    ERROR2    ERROR3    ATIME3    ERROR4    ERROR5    ERROR6     COG
 8.8158    1.0303    1.6[illegible]    [illegible]    12.[illegible]    1.4[illegible]    2.6[illegible]    0.9920    0.07
 ATIME1    ERROR1    ERROR2    ERROR3    ATIME3    ERROR4    ERROR5    ERROR6     COG
 8.3732    0.9[illegible]    1.55[illegible]    0.6[illegible]    11.[illegible]    1.5629    2.2749    0.9321    0.08
 ATIME1    ERROR1    ERROR2    ERROR3    ATIME3    ERROR4    ERROR5    ERROR6     COG
 7.9627    0.9[illegible]    1.5[illegible]    0.6269    11.4036    1.3[illegible]    2.1510    0.8813    0.09
 ATIME1    ERROR1    ERROR2    ERROR3    ATIME3    ERROR4    ERROR5    ERROR6     COG
 [illegible]    0.8698    1.4[illegible]    0.5[illegible]    10.4049    1.2[illegible]    2.0440    0.8[illegible]    0.10

 ATIME1    ERROR1    ERROR2    ERROR3    ATIME3    ERROR4    ERROR5    ERROR6     COG
 7.3374    0.8299    1.3651    0.5[illegible]    10.4535    1.1855    1.9502    0.7990    0.11
 ATIME1    ERROR1    ERROR2    ERROR3    ATIME3    ERROR4    ERROR5    ERROR6     COG
 7.0290    0.7949    1.3064    0.5[illegible]    10.0[illegible]    1.1150    1.8670    0.7650    0.12
```





ATIME1	ERROR1	ERROR2	ERROR3	ATIME3	ERROR4	ERROR5	ERROR6	COG1
0.3367	0.1145	0.1886	0.0772	0.5519	0.1636	0.2692	0.1101	0.8900
ATIME1	ERROR1	ERROR2	ERROR3	ATIME3	ERROR4	ERROR5	ERROR6	COG1
0.3545	0.1087	0.1185	0.0733	0.4450	0.1453	0.2553	0.1047	0.9000
ATIME1	ERROR1	ERROR2	ERROR3	ATIME3	ERROR4	ERROR5	ERROR6	COG1
0.3827	0.1077	0.1685	0.0592	0.4468	0.1466	0.2412	0.0948	0.9100
ATIME1	ERROR1	ERROR2	ERROR3	ATIME3	ERROR4	ERROR5	ERROR6	COG1
0.2764	0.0963	0.1385	0.0649	0.3969	0.1376	0.2264	0.0920	0.9200
ATIME1	ERROR1	ERROR2	ERROR3	ATIME3	ERROR4	ERROR5	ERROR6	COG1
0.2504	0.0797	0.1436	0.0405	0.3437	0.1207	0.2109	0.0866	0.9300
ATIME1	ERROR1	ERROR2	ERROR3	ATIME3	ERROR4	ERROR5	ERROR6	COG1
0.2051	0.0827	0.1341	0.0358	0.2930	0.1187	0.1944	0.0796	0.9400
ATIME1	ERROR1	ERROR2	ERROR3	ATIME3	ERROR4	ERROR5	ERROR6	COG1
0.1700	0.0752	0.1237	0.0507	0.2424	0.1074	0.1767	0.0724	0.9500
ATIME1	ERROR1	ERROR2	ERROR3	ATIME3	ERROR4	ERROR5	ERROR6	COG1
0.1353	0.0610	0.1102	0.0452	0.1933	0.0797	0.1574	0.0645	0.9600
ATIME1	ERROR1	ERROR2	ERROR3	ATIME3	ERROR4	ERROR5	ERROR6	COG1
0.1010	0.0578	0.0951	0.0390	0.1445	0.0825	0.1358	0.0556	0.9700
ATIME1	ERROR1	ERROR2	ERROR3	ATIME3	ERROR4	ERROR5	ERROR6	COG1
0.0670	0.0478	0.0774	0.0317	0.0957	0.0672	0.1105	0.0453	0.9800
ATIME1	ERROR1	ERROR2	ERROR3	ATIME3	ERROR4	ERROR5	ERROR6	COG1
0.0333	0.0332	0.0565	0.0223	0.0476	0.0474	0.0779	0.0319	0.9900
ATIME1	ERROR1	ERROR2	ERROR3	ATIME3	ERROR4	ERROR5	ERROR6	COG1
0.0000	0.0001	0.0002	0.0001	0.0000	0.0002	0.0003	0.0001	1.0000

UNIVERSITÉ DE MADAGASCAR, UNIVERSITY OF MICHIGAN,
and
COLUMBIA UNIVERSITY

Notes

[1] Conrad P. Kottak wishes to thank the Foreign Area Fellowship Program of New York for a research grant and postdoctoral fellowship which supported a part of the research on which this paper is based.

[2] Hudson's recent (1967) publication of word lists for several Barito speech communities of Borneo does much to confirm the close relationship between Malagasy, Maanyan, and other related speech communities of Southeast Borneo. Preliminary comparison of Malagasy dialects with those in the Bornean East Barito subgroup suggests that the percentage of shared cognates between Malagasy and East Barito (including Maanyan) dialects is certainly as high for the 100-word list as Dyen had estimated for the 200-word list and may, in fact, be significantly higher.

[3] As will be apparent, we will be treading on the historical boundary of "language" and "dialect." For this reason, insofar as possible, we shall try to stick to the label "speech community" throughout the paper.

[4] But the actual number of *ethnies* mentioned in governmental statistics varies with different administrators. Some subprefects and canton chiefs add new categories to the officially printed lists to take account of novel cultural divisions recognized by the inhabitants of their administrative areas.

[5] The criteria for ethnic, cultural, or tribal differentiation have been mainly geographical or political. Thus, the term Tanala means 'people of the forest', and includes clans sharing a common habitat, many of them refugees from other areas, but never unified on a kinship or political basis. On the other hand, the term Sakalava refers to an agglomeration of clans or tribes such as Manadabo, Kajemby, and others, who were eventually united into a political confederation, a west coast kingdom based on intensive trade with Arabs and Europeans.

[6] For their assistance in collecting Malagasy word lists we are very grateful to Pastor Randrianaivo, for his help in compiling the Sihanaka and Betsimisaraka word lists, to Father Blot for putting us in contact with missionaries stationed in southeast Madagascar, to Pastor Lanvin for contacting a Tsimihety informant, and to Mr. Themistocle for obtaining a basic vocabulary list for the Sakalava of Morondava.

[7] Eastern and central speech communities are well represented. We regret that it was not possible to obtain additional lists from the Betsimisaraka, Sakalava, and Bara areas. A basic vocabulary list is also lacking for Tanosy. This southern dialect has previously been regarded as intermediate between hypothesized eastern and western subgroups of Malagasy (Dez 1963).

[8] We were somewhat reluctant to take this step at first, since we felt that we were discarding valuable information about subgrouping. The borrowing

from Swahili of the term for *lava*, common to all Malagasy dialects, could be considered a shared innovation which affirms a common, unique subgrouping for the Malagasy as a whole. Nonlinguistic evidence suggests that this and other lexical borrowings probably took place when the speakers of the group of dialects in question were a single ancestral population. Consequently, borrowing may have been one source of shared innovation, and therefore evidence of genetic relationship. However, since we could not conclusively demonstrate that this had been the case for all borrowings from Swahili, we followed Gudschinsky's dictum.

Dr. George Grace (personal communication) has suggested an alternative interpretation of the data which the authors find intriguing. He has suggested that the original settlement of Madagascar may have been small and may have occurred in the far north of the island, in the areas currently inhabited by the Tsimihety and the Tankarana. Gradually the population increased and settlements spread to the south along the east coast. Subsequently, after considerable dialect differentiation had taken place, a group from the southeast may have pushed into the central highlands where it has become our central subgroup. Dr. Grace further suggests that this central group may then have had extensive contact and dialect borrowing with Betsimisaraka (Bsk). We know, in fact, that there has been extensive contact between Me and Bsk for at least the past two hundred years. The Sihanaka (Si) are the nearest neighbors of the northern Betsimisaraka, whose dialect is the basis for our Bsk list. Extensive borrowing could have distorted the shared cognate percentages for Me and Si with Bsk. If Me and Si are excluded from consideration because of the possibility of extensive borrowing, Bsk emerges as distinct from other Malagasy speech communities to about the same extent as Tbk. This would accord with its intermediate geographical position on the east coast between Ts and Tbk.

References

Cross, Ephraim. 1962. Lexicostatistics has not yet attained the status of a science. In: Proceedings of the ninth international congress of linguistics, ed. by Horace G. Lunt, pp. 480-489. The Hague, Mouton & Co.

Dahl, Otto C. 1951. Malgache et Maanjan: une comparaison linguistique. Oslo, Egede Instituett.

Dempwolff, O. 1934-38. Vergleichende Lautlehre des austronesischen Wortschatzes. 3 vols. Berlin, Dietrich Reimer.

Deschamps, H. 1965. Histoire de Madagascar. 3d ed. Paris, Berger-Levrault.

Dez, J. 1963. Aperçus pour une dialectologie de langue malgache. Bulletin de Madagascar, No. 204, 205, 206, 210.

Dyen, Isidore. 1953. Review of Otto Dahl, Malgache et Maanjan. Language 29(4):577–590.

———. 1965. A lexicostatistical classification of the Austronesian languages. International Journal of American Linguistics 17, Memoir 19 (Vol. 31, no. 1).

———. 1966. Comment. Oceanic Linguistics 5(1):32–49.

Elmendorf, William W. 1951. Word taboos and lexical change in Coast Salish. International Journal of American Linguistics 17:205–208.

Ferrand, G. 1903. Essai de grammaire malgache. Paris, Leroux.

Flacourt, Etienne de. 1661. Histoire de la grande île Madagascar. Paris.

Grandidier, Alfred, and Guillaume Grandidier. 1903–13. Collection des ouvrages anciens concernant Madagascar. 8 vols. Paris, Comité de Madagascar.

Grace, George W. 1966. Austronesian lexicostatistical classification: a review article. Oceanic Linguistics 5(1):13–31.

Gurschinsky, Sarah. 1956. The ABC's of lexicostatistics. Word 12:175–210.

Hattori, Shirô. 1953. On the method of glottochronology and the time depth of Proto-Japanese. Gengo Kenkyu [Journal of the Linguistic Society of Japan] nos. 22, 23, pp. 29–77 (English summary pp. 74–77).

Hebert, J. C. 1965. La cosmographie ancienne malgache suivie de l'énumeration des points cardinaux et l'importance du Nord-Est. Annales de la Faculté des Lettres et Sciences Humaines [Tananarive] 1:83–195.

Houtman, Frederick de. 1603. Spraeckende woord-boeck inde Maleysche ende Madagascarsche talen met vele Arabische ende Turesche woorden. Amsterdam, Jan Evertsz.

Hudson, Alfred B. 1967. The Barito Isolects of Borneo. Cornell University, Department of Asian Studies, Data Paper no. 68. Ithaca, New York.

Hymes, Dell H. 1960. Lexicostatistics so far. Current Anthropology 1(1):3–44.

———. 1966. Comment. Oceanic Linguistics 5(1):50–57.

Matsac, V. 1888. Dictionaire Malgache-Français. Tananarive, Imprimerie Catholique.

Mariano, L. 1613–14. Relation du voyage de decouverte fait à l'île Saint-Laurent dans les années 1613–1614. In: Collection des ouvrages anciens concernant Madagascar, ed. by A. and G. Grandidier, 2:1–64. Paris, Comité de Madagascar.

Molet, Louis. 1953. État actuel de la cartographie ethnographique à Madagascar. Bulletin de l'Académie Malgache [Tananarive] n.s. 31:41–93.

Swadesh, Morris. 1954. Perspectives and problems of Amerindian comparative linguistics. Word 10:306–332.

———. 1955. Towards greater accuracy in lexicostatistic dating. International Journal of American Linguistics 21:121–137.

———. 1956. Some limitations of diffusional change in vocabulary. American Anthropologist 58:301–360.

Teuk, H. N. van der. 1864. Outlines of grammar of Malagasy language. Journal of the Royal Asiatic Society 8(2–.