

The Languages of East Timor

Some Basic Facts

(Revised 9.2.2002)

Dr Geoffrey Hull

*Instituto Nacional de Linguística
Universidade Nacional de Timor Lorosa'e*

A relatively small area of the globe stretching from the Indonesian islands of Flores and Celebes through New Guinea to the Solomon Islands, Vanuatu and New Caledonia is remarkable for the enormous number of languages existing there. New Guinea has over 400 distinct languages and Timor, many times smaller, has at least nineteen. This linguistic diversity, which amazed the first Europeans who sailed to the Spice Islands, has necessitated the use of simplified contact languages or *lingua francas* bridging the frequent intelligibility gaps. In Eastern Indonesia Creole Malay formerly fulfilled this role (today standard Indonesian having taken its place). The *lingua franca* of Papua-New Guinea, the Solomons and Vanuatu is Pidgin English, and in East Timor the common language is Creole Tetum or *Tetum-Praça*.

East Timor has sixteen indigenous languages, belonging to two different language families or phyla. Twelve of these languages are of Austronesian origin (and therefore 'cousins' to Malay-Indonesian, Javanese, Tagalog, Malagasy, Motu, Fijian, Samoan and Maori). Although the Austronesian languages of Timor belong, with Malay, to the Western Malayo-Polynesian (or Hesperonesian) division of Austronesian, they are too different in structure and vocabulary to be mutually intelligible with Malay-Indonesian.

The **Timoric** (Timorese-Austronesian) vernaculars belong to the **Neo-Butonic** or **Santalic** branch of the Celebic languages and fall into two main groups: **Fabronic** and **Ramelaic**. The Fabronic languages (Tetum, Kawaimina, Habun, Wetarese, Galoli, Bekais and Dawan) and the Ramelaic languages (Tokodede, Kemak, Mambai, Idalaka) descend from **Old Butonese**, introduced from the Muna-Buton-Tukang Besi region of South-East Celebes probably about one thousand years ago. It appears that at the time of the Butonese migrations to Timor only non-Austronesian languages, at least some of them New Guinean, were spoken on the island. Not long after the Butonese settlements, Timor was invaded by people from the Central Moluccas (Ambon, Ceram). As a result of contact with Moluccan languages, the Butonic dialects underwent a drastic grammatical simplification known to linguists as *creolization*. The aboriginal languages were eventually creolized as well.

Another Fabronic language of the extreme east, **Lóvaia** or **Maku'a**, is distinct from the others in that it was introduced much later from a nearby island of the South Moluccas.

Of Timor's aboriginal languages only four (Bunak, Makasai, Makalero and Fatuluku) survive today. These four are distantly related to Papuan languages of the **Trans-New Guinea phylum** spoken in the Bomberai Peninsula of West Papua (Fakfak district). There are also traces of other—apparently non-Papuan—languages in the vocabulary of the various Timoric languages that eventually replaced them, especially in Kawaimina, Mambai, Kemak, Tokodede and Dawan. There is so far no evidence that any pre-Austronesian language of Timor that has left traces was related to Aboriginal languages of nearby Bathurst and Melville Islands or of the Northern Territory or Kimberley coast. Given the time gap involved, it is highly unlikely that connections will be found.

East Timor's sixteen languages are classified and described as follows:

AUSTRONESIAN LANGUAGES

Timoric Stock

Fabronic Group

The Fabronic languages are so named because of their close genetic connection with the language of the Tukang Besi Islands (*Fabronum Insulae* in Latin), derived from Old Butonese, the ancestor of the Timoric languages. Unlike the Ramelaic languages, the Fabronic languages have not been greatly affected in their vocabulary by the aboriginal languages they replaced.

1. TETUM

Called by its speakers *lia-tetun* ('language of the plain'), Tetum is the most widely used vernacular in East Timor and the country's national language. It is spoken in three separate regions of the island: 1) a central strip from the Ombai Strait to the Timor Sea split by the border of East Timor and West Timor (the **Tetun-Belu** or **Belunese** dialect) and including the towns of Atapupu and Atambua (West Timor), and Balibó, Fatomean, Fohorém and Suai (East Timor); 2) the south coastal region around Alas, Luca and Viqueque and including the old kingdoms of Samoro and Soibada (**Tetun-Terik**); 3) the city of Dili and its suburbs (**Tetun-Dili** or **Tetun-Prasa/Tétum-Praça**). Everywhere else in East Timor, Tetum is a second language. Since the final settling of the Dutch-Portuguese colonial border in 1914, the Tetum-speakers of the north-west (Balibó) have been separated from those of the south-west.

Tetun-Belu and Tetun-Terik are not spoken or well understood outside their home territories. Tetun-Praça or *Tetun-Prasa* (*Prasa* 'big town' being a synonym for Dili) is the form of Tetum that is spoken throughout East Timor. Only in the extreme east

(Lospalos region) and in the western Ambeno enclave is Tetum not fully current: here Indonesian is the usual second language, Portuguese having fulfilled this role before 1975. However, the use of Tetum is now rapidly increasing, especially in the Lospalos district.

In the fifteenth century, before the arrival of the Portuguese, Tetum spread through central and eastern Timor as a contact language under the aegis of the Belunese-speaking Kingdom of Wehali, at that time the most powerful kingdom on the island. The Portuguese (present in Timor from c. 1556) made most of their settlements in the west, where Dawan was spoken, and it was not until 1769, when the capital was moved from Lifau (Oecussi) to Dili that the Portuguese began to promote Tetum as an inter-regional language in their colony. (Timor was one of the few colonies of Lisbon where a local language, and not a form of Portuguese, became the *lingua franca*: this is because Portuguese rule was indirect rather than direct, the Europeans governing through local kings who embraced Catholicism and became vassals of the King of Portugal).

In 1769 Dili was a Mambai-speaking district. The variety of Tetum that came to be spoken in the new town was strongly influenced by the local dialect of Mambai. The loss of the consonants **w** (becoming **b**: *wee* 'water' > *bee*, *lawarik* 'child' > *labarik*) and of the glottal stop (*to'o* 'arrive' > *too*, *sa'e* 'rise' > *sae*) and the simplification of the noun and verb systems (loss of suffixes) were the long-term result of Mambai influence. It was this simplified (creolized) form of Tetum that became the *lingua franca* of Portuguese Timor and also the language in which the Gospel was preached by Catholic missionaries. Nevertheless, Tetum (like all other vernaculars) was completely excluded from the education system in the days of Portuguese rule, and it was only the free East Timorese government of 1974-5 that took steps to give it public status and to use it as a medium of education. In 1980 the Church adopted Tetum as its liturgical language, but only after the Indonesian government forbade the use of Portuguese in Catholic public worship: Portuguese had been used alongside Latin in the Mass since the late 1960s.

The varieties of Tetum spoken in East Timor are heavily influenced by Portuguese; the varieties of West Timor have a strong Malay-Indonesian influence. This difference of superstratum hinders mutual intelligibility between the two varieties. Malay, the *lingua franca* of the Malay Archipelago, was introduced from Ambon and became current along the Timorese coasts from the fifteenth century, contributing many new words to Tetum and all other languages of the island. A variety of Creole Malay actually became the vernacular of Kupang, the capital of Dutch Timor with a mixed population of Rotinese, Ndaonese, Atoni, Helong (the original inhabitants) and other groups.

Malay influence ceased in East Timor in the mid nineteenth century, when the colonial government made efforts to increase the use of Portuguese. Malay returned to East Timor in the form of Bahasa Indonesia in 1975, but in spite of 24 years of official imposition, its impact on local languages remains superficial and is likely to wane in the independent nation. Words of Malay origin in Tetum therefore belong to two different historical strata: those absorbed between the 15th and 19th centuries (from Ambonese Malay, e.g. *besi* 'steel', *toko* 'shop', *ukun* 'to rule' (< M. *hukum*), *sarani* 'Christian', *barani* 'brave'), and those borrowed after 1975 (from Indonesian).

Numerals 1-10:

Tetum: **ida, rua, tolu, haat, lima, neen, hitu, ualu, sia, sanulu**

Tetun-Belu: **ida, rua, tolu, haat, lima, neen, hitu, ualu, siwi, sanulu**

2. HABUN

Habun is a numerically small language spoken in the Cribas district south of Manatuto. It may be considered an archaic variety of Tetum and in the past was close to its eastern neighbour Kawaimina. Habun also shows some similarities to the Idaté dialect of Idalaka spoken to the west.

Numerals 1-10:

isa, rua, tolu, haa, lima, neen, hitu, ualu, sia, sanulu

3. KAWAIMINA

This acronym (used by linguists rather than by speakers of the language) refers jointly to the similar **Kairui**, **Waimaha**, **Midiki** and **Naueti** dialects, the easternmost Timoric vernaculars. Kairui is spoken in and around the village of that name; Waimaha (*Waima'a* to its speakers, *Waimoa* to other Timorese) is spoken along the north coast from Vemasse to Bucóli and the outskirts of Baucau; Midiki is heard in the districts of Lacluta, Liaruca and Venilale; and Naueti (separated geographically from the other three dialects) is used on the south coast around Uatolári and Uato Carabau, surrounded by Makasai-speaking territory.

The Kawaimina dialects have both remarkable archaisms and strange innovations (e.g. vowel harmony, aspirated and postglottalized consonants) in their sound-systems. Their grammar, on the other hand, is (with the partial exception of Naueti) very simple in structure.

Numerals 1-10:

Kairui: **se, kirua, kitele, kihoo, kiliim, kinee, kihiti, kikoho, kisia, bosé**

Waimaha: **se, kairuo, kaitelu, kaihaa, kailime, kainena, kaihitu, kaikaha, kaisiwe, basé**

Midiki: **se, kairuo, kaitelu, kaihaa, kailime, kainee, kaihitu, kaikaha, kaisiwe, basé**

Naueti: **se, kairua, kaitelu, kaihaa, kailima, kailima-resin, kailima-resi-kairua, kailima-resi-kaitelu, kailima-resi-kaihaa, welisé**
(NB: quinary counting system for 6-9)

4. IDALAKA

The scientific acronym 'Idalaka' has been coined to refer jointly to three very similar dialects of central East Timor: **Lakalei**, spoken in the Fahinehan district, and **Idaté**, the vernacular of Laclúbar, and **Isní**, proper to the intervening district east of Turiscai. These dialects are closely related to Tetum and Habun and have numerous features in

common with Galoli. An early variety of Idalaka was the immediate ancestor of the Ramelaic languages (Mambai, Tokodede and Kemak). The **Lolein** subdialect of the Becora Leten district south-west of Dili is a variety of Isní introduced by nineteenth-century immigrants from the Turiscai region.

Lakalei: **isa, rua, telu, aat, lima, neen, hitu, ualu, sia, sakulu**
Isní: **is, rua, tel, aat, lim, neen, hitu, ualu, sia, sakúl**
Lolein: **isa, rua, telu, aat, lima, neen, hitu, ualu, sia, sakulu**
Idaté: **isa, rua, telu, aat, lima, neen, hitu, ualu, sanulu**

5. GALOLI

Galoli (*Galolen, Glolen*) is the vernacular of the Manatuto and Laleia districts of the north coast between Dili and Baucau. Because of the very arid terrain of their region, the *Lo'ok* or Galoli speakers have traditionally turned to the sea for their livelihood. There is an old Galoli colony on the south coast of nearby Wetar (speakers of the Talo dialect, still mutually intelligible with Timorese Galoli).

Manatuto has for centuries been a main landing-place in East Timor, which explains the large number of foreign (Central and Southern Moluccan and Malay) loanwords in its vocabulary. Manatuto was the spearhead of new Catholic missionary endeavour in the late nineteenth century, and among the earliest specimens of East Timorese literature is a catechism in the Galoli language, a Galoli grammar and a Portuguese-Galoli dictionary.

Numerals 1-10:

Manatuto dialect: **nehe, irua, itelu, ihaat, ilima, ineen, ihitu, ihaa, isia, sanulu**
Talo dialect of Wetar: **nehe, erua, etelu, ehaat, elima, eneem, ehitu, ehaa, esia, sanulo**

6. WETARESE (Atauran dialect)

The small East Timorese island of Ataúro, known as *Pulau Kambing* ('Goat Island') in Malay for the large numbers of goats traditionally raised there, lies between the larger Indonesian islands of Alor (Ombai) and Wetar and faces the city of Dili. Ataúro's three dialects are **Rahesuk**, spoken in the northern villages of Biquéli and Beloi, **Resuk** in the south-eastern villages of Maumeta and Makili, and **Raklungu** in the south-western villages of Makdadi and Manroni. Ataúro is culturally unique in East Timor in that many of its northern inhabitants (Rahesuk speakers) are not Catholics but Protestants, having been evangelized from a Dutch Calvinist mission on Alor earlier this century. There are also numbers of Protestants in the predominantly Catholic south.

Ataúro's three dialects, mutually very similar, do not form an 'Atauran' language, but are variants (subdialects) of the Wetarese language proper to Wetar and the tiny island of Lirar situated between Ataúro and Wetar. Wetarese is closely related to Galoli, but

it is more conservative in structure and its vocabulary has been more influenced by Malay and languages of the Moluccas.

Numerals 1-10:

Rahesuk: **iha, barua, batelu, waa, balima, baneen, baítu, pawau, pasia, sanulu**

Resuk: **hia, harua, hatelu, haát, halima, haneen, haítu, ha'au, hasé, sangulu**

Raklungu: **hea, herua, hetelu, heát, helima, heneen, heítu, he'au, hesé, sengulu**

7. BEKAIS

Bekais (named *Welaun* by its speakers), is used in the small Sanirin district north of Balibó, bordering on the Tetun-Belu and the Kemak-speaking zones. The most recent research has shown it to be a distinct Fabronic language, though largely assimilated to Tetum today. In the past Bekais was probably spoken along a considerable tract of coastal territory between the Tokodede and Dawan speech-zones. After the sixteenth century it was gradually replaced in most of this area by the Belunese Tetum dialect of the Kingdom of Wehali, advancing from the south.

Numerals 1-10:

Bekais: **isa, rua, tolu, hoat, lima, inan, hitu, ualu, siwi, sakulu**

8. DAWAN (Baikenu dialect)

Dawan is the mother tongue of the Atoni Pah Meto people who inhabit most of the western half of Timor, and is the Timorese language with the largest number of native speakers (over 600,000: Tetum is used by a larger number of people, but only a third of these have it as their first language). This language is counted among those of East Timor because one of its dialects, **Baikenu** (Vaiqueno), is the vernacular of the Ambeno (Oecussi) enclave on the north-west coast.

Lifau, the old capital of Ambewno (the new capital is nearby Pante Macassar, known locally as *Oekusi*), was until 1769 the seat of Portuguese colonial rule in Timor. When, later, most of the adjacent parts of western Timor were occupied by the Dutch, Ambeno continued to fly the Portuguese flag, though its rulers, the Topasses or 'Black Portuguese' (a Catholic dynasty of mixed indigenous and European blood) paid only nominal allegiance to the 'White Portuguese' governor in Dili. Direct Portuguese rule over Ambeno did not come until the late nineteenth century.

Although it has the same immediate origin as Tetum, Dawan is very difficult for other East Timorese to understand because its sound-system has been drastically altered by aboriginal and, later, Central Moluccan influences. Dawan—even the Baikenu dialect—has borrowed more extensively from Malay than the languages of the east. Portuguese influence has at the same time been strong on Baikenu.

NOTE: The correct name of the enclave is *Ambeno*. *Oecussi* (< *Oe-Kusi* 'the water pot') is the name of the capital, known officially by its Malay name of Pante Macassar

('the beach of the Macassarese'), a reference to the Celebean traders who used to land and congregate there.

Numerals 1-10:

mese, nua, teun, haa, niim, nee, hiut, faun, sio, boés

A NOTE ON HELONG, ROTINESE AND NDAONESE

There exist two more Fabronic languages, both of which belong exclusively to West Timor. **Helong**, the original vernacular of Kupang (which today speaks Malay), is now restricted to a few villages south of the city and along the east coast of the adjacent island of Semaui. **Rotinese**, the vernacular of Roti, is broken up into many dialects. Rotinese speakers are found today in many districts of West Timor, the result of transmigration programmes initiated by the Dutch in the nineteenth century. Helong and Rotinese are both closely related to Dawan. Although their structure is less 'eroded', they show many aboriginal and Central Moluccan influences in their vocabularies.

The **Ndaonese** language of the small island of Ndao, the westernmost in the Timor archipelago, is Austronesian, but an offshoot of Savunese (from Savu Island) and a member of the Florinic stock rather than an Fabronic language. Colonies of Ndaonese speakers are also found in Kupang and other districts of West Timor.

Numerals 1-10:

Helong: **mesa, dua, tilu, aat, lima, eneng, itu, palu, sipa, sngulu**

Rotinese: **esa, dua, telu, haa, lima, ne, hitu, falu, sio, sanahulu**

Savunese: **ahi, du'e, telu, apa, lami, ana, pidu, aru, he'o, henguru**

Ramelaic Languages

The Ramelaic languages are so named because they are spoken in a broad area centred around the Ramelau range. They are younger languages than the Fabronic ones because all of them evolved from an earlier form of **Idalaka**: the Ramelaic zone was therefore the last one in the island to be austronesianized. These languages belong in a category of their own also because although derived from a Fabronic language, they have all been considerably transformed by the aboriginal languages spoken in each speech-zone. This pre-Austronesian influence is most apparent in the vocabularies of Mambai, Kemak and Tokodede.

9. MAMBAI

In numerical terms, Mambai (named in the vernacular *Manbae*) is the third most important language in East Timor, with over 90,000 speakers. As the language of the

most mountainous area of Timor, Ramelau and the surrounding ranges, Mambai is broken up into many local dialects. Its main centres are Ermera, Aileu, Remexio, Turiscaí, Maubisse, Ainaro and Same. Mambai dialects are divided into a southern group which keep the Austronesian consonant /p/ intact (e.g. *pat* 'four') and a northern group which change it to [f] (cf. *fat*).

Mambai has a simplified morphology (the model for Tetum-Praça) and a partly aboriginal vocabulary. The southern dialects have many words shortened by syncope and apocope, e.g. Ainaro *mlua* 'wide', *lel* 'sun', *ton* 'year' compared to Remexio *mulua*, *lelon*, *tonan*. This feature (also occurring in the north, but less marked there) is shared by the neighbouring Papuan Bunak language and by Dawan, and may be due to the influence of an extinct aboriginal language.

Numerals 1-10:

Ainaro dialect: **id, rua, tel, pat, lim, lim-nain-ide, lim-nai-rua, lim-nai-tel, lim-nai-pata, sagúl** (N.B.: quinary numerals 6-9)

Aileu dialect: **id, ru, teul, fat, lim, nen, hitu, ualu, sia, sakúl**

10. KEMAK

Kemak (*Ema*) is spoken in the north-west of East Timor, in and around Atabae, Cailaco, Atsabe and Maliana. This language is a close relative of Tokodede, but unlike Tokodede and Mambai it retains a certain morphological complexity inherited from Idalaka. As well as aboriginal words in its vocabulary, Kemak has an element which points to influence in the past from the South Moluccas. Malay has also left a stronger imprint on this language than on its neighbours to the east.

Numerals 1-10:

sia, rua, telu, pata, lima, neme, icu, balu, sibe, sapulu

11. TOKODEDE

Tokodede (*Tukodede*) is the vernacular of the north-western coastal strip of East Timor around Vatoboro, Maubara, Liquiçá and Bazartete. Like Kemak, Tokodede has aboriginal and Moluccan influences in its vocabulary, some of the latter shared with Galoli and the dialects of Ataúro. Unlike Kemak, however, Tokodede has a simple morphology similar to that of Mambai.

Numerals 1-10:

iso, ru, telu, paat, liim, hohoniso, hohorú, hohotelu, hohopaat, sagulu
(N.B.: quinary numerals 6-9)

Arafuric Stock

12. LOVAIA (MAKU'A)

Lovaia, formerly spoken in two small pockets near Tutuala in the extreme east, was probably the last regional language to be introduced to East Timor, its speakers' ancestors having migrated from a neighbouring South-East Moluccan island, probably Kisar. Lovaia is similar to the Meher dialect of Kisar and to the Luangese of the Leti-Babar Islands. This language is on the verge of extinction today, with only a handful of elderly speakers left: the younger generations have undergone a language shift to Fataluku. Another Lóvaia colony living further west in the Makasai-speaking district of Luro appears to have already lost its language.

Numerals 1-10:

itetlá, urua, okelo, o'aka, olima, oneme, oíko, oava, osia, ideli

TRANS-NEW GUINEAN LANGUAGES

Of East Timor's four identifiably Papuan languages Makasai, Makalero and Fataluku are all closely interrelated and akin to the languages of the Indonesian islands of Alor, Pura and Pantar off the north-west coast of Timor. Their relationship with Bunak is more remote. These four languages appear to have lost most of their original morphosyntactic features, becoming largely assimilated in structure to the surrounding creolized Austronesian languages. The most resistant Papuan grammatical characteristic is the typical *subject + object + verb* word order, contrasting with the *subject + verb + object* word order of the Austronesian languages. The traditional vocabularies of Fataluku, Makalero, Makasai and (especially) Bunak — which show the effects of three distinct pre-Papuan substrata — have been replaced in part by new Austronesian words.

13. BUNAK

Bunak (*Gai'*), spoken in the districts of Bobonaro, Lalotoc, Tilômar, Zumalai and Cassa in south-western East Timor, is unintelligible to speakers of the neighbouring Austronesian languages. Along the south coast many Bunak and Belunese (Tetun-Belu) speaking communities co-exist. Modern Bunak is only remotely related to Makasai and Fataluku.

Bunak words have a highly eroded structure, many of them reduced to a single syllable; for example the Bomberaian word for 'dog', rendered as *iparu* in Fataluku and *defa* in Makasai, is reduced to *zaf* in Bunak. The vocabulary of Bunak and its western dialect Marae is very mixed and contains words which are apparently

aboriginal (i.e. pre-Papuan) and Timoric words, as well as the usual borrowings from Malay and Portuguese.

Numerals 1-10:

Bobonaro dialect: **uen, hiro-on, goni-on, goni-il, goni-ciet, thomor, hicu, walu, siwe, sogo**

Zumalai dialect: **wen, hili-on, goni-on, goni-il, goinseet, temol, hitu, alu, sie, sego**

Marae dialect of Indonesian Timor: **uwen, hile-on, koni-on, koni-il, koni-tiet, tomol, hitu, walu, siwe, soko**

14. MAKASAI

Makasai (*Makasae*) is numerically the second most important language of East Timor, with over 100,000 speakers today. It is the vernacular of the districts of Baucau (where Waimaha is also spoken), Quelicai, Ossu, Baguia, Laga, Laivai and Luro. The Ossu dialect stands apart in its preservation of the consonant *p*, which became *f* elsewhere (cf. Ossu *pi*, Baucau *fi* 'we'; Ossu *apa*, Baucau *afa* 'stone'). The nickname given to the inhabitants of the eastern half of East Timor, *Firaku*, is a Makasai word (*fi raku* 'we comrades').

Numerals 1-10:

u, lola'e, lolitu, loloha, lima, daho, fitu, afo, siwa, ruru-u

15. MAKALERO

Makalero (*Maklere*) is spoken on the south-east coast of Timor in the district of Iliomar. It is transitional between Fataluku and Makasai though generally more similar to the latter.

Numerals 1-10: **u, loloi, lolitu, lolo'e, lima, douhisi, fitu, afo, siwa, ruru-u**

16. FATALUKU

Fataluku (*Fatalukunu*) is the vernacular of the culturally diverse people of the far eastern end (*Ponta Leste*) of the island who inhabit the districts of Lautém, Lospalos, Loré and Tutuala. There are five dialects. Although Fataluku is similar in structure and vocabulary to Makasai and Makalero, these languages are no longer intelligible to Fataluku speakers. The north-western dialect has been somewhat influenced by Makasai. Fataluku has some unique phonetic characteristics, for example an absence of the voiced consonants *b*, *g*, *d* (except in the north-western dialect), and most words ending in vowels.

A dialect of Fataluku, Oiratan, is spoken in one village in the south of the small Indonesian island of Kisar, lying off the north coast.

Numerals 1-10:

Lospalos dialect: **ukani, ece, utue, fate, neme, fetu, ikafa, siwa, ta'ane**

Lautém dialect: **ukani, ece, itue, fate, lime, neme, fitu, kafa, sife, taane**

Oirata dialect of Kisar: **uani, ei, utu, pata, limi, neme, pitu, kapa, siwa, taanauni**

To these sixteen languages may be added four more which are not indigenous to East Timor: **Portuguese**, the official and historic second language and the principal influence on all the vernaculars; **Indonesian**, current in the territory only since 1975; and the **Hakka** dialect of the Sino-Timorese community, most of whom trace their origin to the city of Meixian in Guangdong Province. The small number of Macanese Chinese resident in East Timor have traditionally spoken **Cantonese**. East Timor's variety of **Creole Portuguese** (*português de Bidau*), spoken in Dili, became extinct in the 1960s. This dialect was never widespread in the colony, having been introduced and mainly used by Larantucan (Florinese Mestiço) residents of the eastern Bidau suburb of Dili.

Numerals 1-10:

Portuguese: **um, dois, três, quatro, cinco, seis, sete, oito, nove, dez**

Indonesian: **satu, dua, tiga, empat, lima, enam, tujuh, delapan, sembilan, sepuluh**

Hakka: **it, nyi, sam, syi, ng, liuk, chit, bat, giu, sip**

Cantonese: **yat, i, saam, sei, ng, luk, chat, baat, gau, sap**

Mandarin: **yì, er, san, sì, wu, liú, qí, ba, jiú, shí**

© Geoffrey Hull 1999, University of Western Sydney, and 2002, Instituto Nacional de Linguística, Universidade Nacional de Timor Lorosa'e

For a comprehensive bibliography on the languages of East Timor to 1999, see the following article:

Geoffrey Hull. "The Languages of East Timor: 1772-1997: A Literature Review", *Studies in Languages and Cultures of East Timor*. University of Western Sydney Macarthur, 1999, pp. 1-38.

See also various articles in the above journal (Volumes 1 - IV) for more information on the vocabulary, sound system and grammar of the Austronesian languages of East Timor.

Prospective Language Researchers

In the new independent state of East Timor research on the languages of the country will be officially co-ordinated by the Instituto Nacional de Linguística at the National University of East Timor. A great deal of data has already been collected from the regional languages and their dialects but there is a need for more trained linguists to assist in the analysis of the data and the creation of language profiles (four of which have already been produced by INL, for Mambai-Ainaro, Galoli, Waimaha and Baikenu), grammars and dictionaries. A prerequisite for participation in the work of the Institute is a good command of the Portuguese and Tetum languages and at least a passive reading knowledge of Malay (Indonesian). Individuals possessing these skills or willing to acquire them, and who have at least an undergraduate Arts degree including a course or courses in general or descriptive linguistics may contact Dr Geoffrey Hull (c/- SAdS Project auriga6@optusnet.com.au) for more information concerning training and research opportunities in East Timor.