47 · HOW MANY LANGUAGES?

There is no agreed total for the number of languages spoken in the world today. Most reference books give a figure of 5,000 to 6,000, but estimates have varied from 3,000 to 10,000. To see why there is such uncertainty, we need to consider the many problems facing those who wish to obtain accurate information, and also the reasons (linguistic, historical, and cultural) which preclude a simple answer to the question 'What counts as a language?'

DISCOVERIES

An obvious reason for the uncertainty over numbers is that even today new peoples, and therefore languages, continue to be discovered in the unexplored regions of the world – especially in the Amazon basin (as the Transamazonica road system is extended), Central Africa, and New Guinea. However, only a few languages are likely to be encountered in this way; and it is much more usual to find parts of the world where the people are known, but the languages spoken in their area are not. There are in fact many countries where linguistic surveys are incomplete or have not even begun. It is often assumed that the people speak one of the known languages in their area; or that they speak a dialect of one of these languages; but upon investigation their speech is found to be so different that it has to be recognized as a separate language.

ALIVE OR DEAD?

Against this steady increase in the world language total, there is a major factor which decreases it. For a language to count as 'living', there obviously have to be native speakers alive who use it. But in many parts of the world, it is by no means an easy matter to determine whether native speakers are still living – or, if they are, whether they still use their mother tongue regularly.

The speed with which a language can die in the smaller communities of the world is truly remarkable. The Amazonian explorations led to the discovery of many new languages, but they also led to their rapid death, as the Indians became swallowed up by the dominant western culture. Within a generation, all traces of a language can disappear. Political decisions force tribes to move or be split up. Economic prospects attract younger members away from the villages. New diseases take their toll. In 1962, Trumai, spoken in a single village on the lower Culuene River in Brazil, was reduced by an influenza epidemic to a population of fewer than 10 speakers. In the 19th century, there were thought to be over 1,000 Indian languages in Brazil; today, there are only 200. A quarter of the world's languages have fewer than 1,000 speakers; half have fewer than 10,000. It is likely that most of these languages will die out in the next 50 years.

LANGUAGE – OR DIALECT?

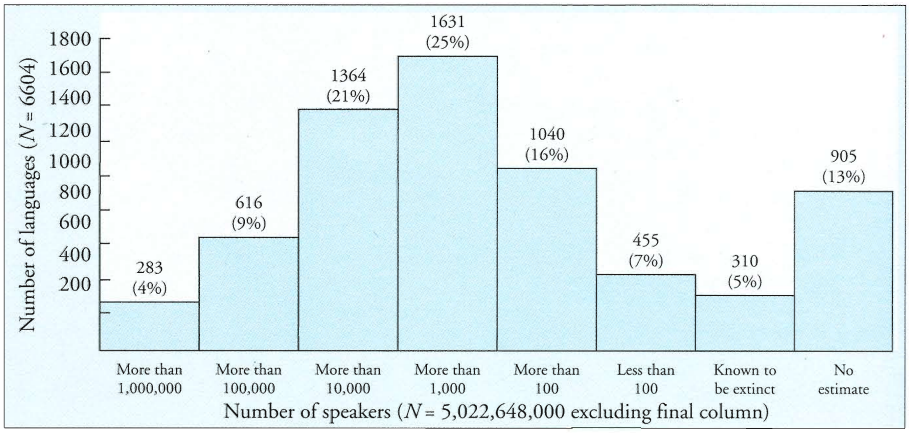
For most languages, the distinction between language and dialect is fairly clear-cut (p. 25). In the case of English, for example, even though regional vocabulary and local differences of pronunciation can make communication difficult at times, no-one disputes the existence of an underlying linguistic unity that all speakers identify as English, and which is confirmed by the use of a standard written language and a common literary heritage. But in hundreds of cases, considerations of this kind are in conflict with each other, or do not clearly apply.

The best-known conflicts occur when the criteria of national identity and mutual intelligibility do not coincide. The most common situation is one where two spoken varieties are mutually intelligible, but for political and historical reasons, they are referred to as different languages. For example, using just the intelligibility criterion, there are really only two Scandinavian languages: Continental (Swedish, Danish, and two standard varieties of Norwegian) and Insular (Icelandic, Faeroese). Swedes, Danes, and Norwegians can understand each other's speech, to a greater or lesser extent. But as soon as non-linguistic criteria are taken into account, we have to recognize at least five languages. To be Norwegian is to speak Norwegian; to be Danish is to speak Danish; and so on. In such cases, political and linguistic identity merge. And there are many other similar cases where political, ethnic, religious, literary, or other identities force a division where linguistically there is relatively little difference – Hindi vs Urdu, Bengali vs Assamese, Serbian vs Croatian, Twi vs Fante, Xhosa vs Zulu.

A new road cuts a swathe through the Brazilian rainforest.

NUMBER OF SPEAKERS

Number of speakers of the world's languages, based on the data provided in *the International Encyclopedia of Linguistics* (Bright, 1992). The total number of languages (including extinct ones) is 6,604. Most of the estimates were made during the 1980s, with some from the late 1970s. The world population total passed 5,000 million in July 1986, and had reached 5,111 million by mid-1988, which gives an indication of the order of magnitude unaccounted for in the final column.



The opposite situation is also quite common. Here we find cases where spoken varieties are mutually *unintelligible*, but for political, historical, or cultural reasons they are nonetheless called varieties of the same language. The three main 'dialects' of Lapp fall into this category, for example. Chinese is a case where linguistic criteria alone are in conflict with each other. From the viewpoint of the spoken language, the many hundreds of dialects in China can be grouped into eight main types (p. 314), which are mutually unintelligible to various degrees. But speakers of all these dialects share the same written language tradition, and those who have learned the system of Chinese characters are able to communicate with each other. Despite the linguistic differences, therefore, Chinese is considered by its speakers to be a single language.

In the above cases, the languages in question have been well studied, and many speakers are involved. When languages have been little studied, or have very few speakers, it is much more difficult for linguists to interpret all the factors correctly. For example, when two languages are in close proximity, they often borrow words from each other – sometimes even sounds and grammar. On first acquaintance, therefore, the languages may seem more alike than they really are, and analysts may believe them to be dialects of the same language. This has proved to be a real problem in such parts of the world as South America, Africa, and South-east Asia, where whole groups of languages may be affected in this way. Similarly, decisions about how to analyse all cases of dialect continua (p. 25) will affect our final total of languages.

LANGUAGE NAMES

A big problem, in working on lesser-known language areas, is deciding what credence to give to a language name. This issue does not arise when discussing the main languages of the world, which are usually known by a single name that translates neatly into other languages – as in the case of *Deutsch, German, Tedesco, Nemetskiy,* and *Allemand*, for instance. But in many cases the situation is not so straightforward.

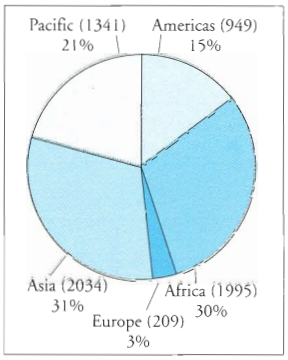
At one extreme, many communities have no specific name for their language. The name they use is the same as a common word or phrase in the language, such as the word for 'our language' or 'our people'. This is often so in Africa (where the name *Bantu*, which is given to a whole family of languages, means simply 'people'), and also in Meso- and South America. In the latter areas, we find such examples as *Carib* = 'people', *Tapuya* = 'enemy', and *Macu* = 'forest tribes'. Some tribes were called *chichimecatl* (= 'lineage of dogs'), *chontalli* (= 'foreigners') or *popoloca* (= 'barbarians'), and these labels led to the modern language names Chichimeca, Chontal, and Popoloca. Frequently, the name is the same as a river on which a tribe has been observed to live, as with the many groups of Land Dayak, in the West Indonesian family. In several Australian aboriginal languages, the name for the language is the word for 'this': for example, the nine languages within the Yuulngu family are known as *Dhuwala, Dhuwal, Dhiyakuy, Dhangu, Dhay'yi, Djangu, Djinang, Djining,* and *Nhangu*. Asking native speakers what language they speak is of little practical help, in such circumstances, if they only answer 'this'!

At the other extreme, it is quite common to find a community whose language has too many names. A South American Indian tribe, for instance, may have several names. A tribe, first of all, will have a name for itself (see above). But adjacent tribes may give the people a different name (e.g. *Puelche* means 'people from the east' in Araucanian). The Spanish or Portuguese explorers may have given them a third name – perhaps a characteristic of their appearance (e.g. *Coroado* means 'crowned' in Portuguese). More recently, anthropologists and other investigators may have used another name, often based on the geographical location of the tribe (e.g. 'up-river' vs 'down-river'). And lastly, the same language may be spelled differently in Spanish, Portuguese, English, or in its own writing system (if one has been devised). For example, Machacali, spoken in Minas Gerais, Brazil, is sometimes spelled Maxakali, sometimes Maxakari. When the initial letters vary (as when the Peruvian language Candoshi is spelled Kandoshi), indexing is especially awkward.

There are further complications. Sometimes, the same name is applied to two different languages, as when *mexicano* is used in Mexico to refer to Spanish (otherwise known as *español* or *castellano*) and to the main Indian language (*nahuatl*). Sometimes, speakers from different backgrounds may disagree about whether their ways of speaking should be related at all. Speakers of Luri, spoken in south-west Iran, say that their speech is a dialect of Persian; speakers of Persian disagree. Asking the native speakers is evidently no solution, for their perceptions will be governed by non-linguistic considerations, especially of a religious, nationalistic, or socioeconomic kind.

TO CONCLUDE

When all these factors are taken into account, it is plain that there will be no single answer to the question 'How many languages?' In some parts of the world, there has been a tendency to over-estimate, by taking names too literally and not grouping dialects together sufficiently – the Malayo-Polynesian languages are often cited in this connection. In other places, the totals are likely to have been underestimated – Indonesian languages, for example. There are over 37,000 language names listed in the 12th edition of *Ethnologue* (Grimes, 1992), and these have been grouped into 6,528 living languages. The number listed in the Index to the *Atlas of the World's Languages* (Moseley & Asher, 1994) is 6,796. The *International Encyclopedia of Linguistics* (Bright, 1992) lists 6,604, but this includes some 300 extinct languages. These surveys generally use data from the 1970s and 1980s. A total of 6,000 would seem to be a safe estimate for the 1990s.



WHERE ARE MODERN LANGUAGES SPOKEN?

The geographical distiribution of living languages, according to *Ethnologue* (Grimes, 1992), based on a total of 6,528 languages.

HOW MANY LANGUAGES HAVE THERE BEEN?

Based on what is known about the rate of language change at which new languages develop from a common origin (p. 331), it is possible to speculate about the number of languages which may have existed since the emergence of a human language faculty. Cautious estimates suggest 30,000; radical ones, over 500,000. A plausible 'middle of the road' figure is 150,000.