1

Warramurrungunji's Children

يَا أَيُّهَا النَّاسُ إِنَّا خَلَقْنَاكُمْ مِنْ ذَكَرٍ وَأُنْثَى وَجَعَلْنَاكُمْ شُعُوبًا وَقَبَائِلَ لِتَعَارَفُوا

Oh Mankind, we have created you male and female, and have made you into nations and tribes that ye may know one another

(Koran 49:13, Pickthal translation)

In the oral traditions of northwestern Arnhem Land, the first human to enter the Australian continent was a woman, Warramurrungunji, who came out of the Arafura Sea on Croker Island near the Cobourg Peninsula, having traveled from Macassar in Indonesia. (Her rather formidable name is pronounced, roughly: *worra-moorrooo-ngoon-gee* [wóramùruŋùŋɟi].) Her first job was to sort out the right rituals so that the many children she gave birth to along the way could survive. The hot mounds of sand, over which she and all women thereafter would have to purify themselves after childbirth, remain in the landscape as the giant sandhills along Croker Island's northern coasts. Then she headed inland, and as she went she put different children into particular areas, decreeing which languages should be spoken where. *Ruka kundangani riki angbaldaharrama! Ruka nuyi nuwung inyman!* "I am putting you here, this is the language you should talk! This is your language!" she would say, in the Iwaidja version of the story, naming a different language for each group and moving on.

Figure 1.1 Tim Mamitba telling the Warramurrungunji story (photo: Nick Evans)

Language Diversity and Human Destiny

ŋari-waidbaidjun junbalal-ŋuban wuldjamin daa-walwaljun lilia-woŋa	Speech of different clans, mingling together . . .
duandja mada-gulgdun-maraŋala dualgindiu	Dua moiety clans, with their special distinct tongues.
wulgandarawiŋoi murununŋdu jujululwiŋoi garaŋariwiŋoi garidjalulu mada-gulgdun-maraŋala	People from Blue Mud Bay, clans of different tongues talking together . . .
buduruna ŋari-waidbaidjun woŋa ŋari-ŋariun	Words flying over the country, like the voices of birds . . .

(Song 2, Rose River Cycle, Berndt 1976:86–7, 197–8)

The Judeo-Christian tradition sees the profusion of tongues after the Tower of Babel as a negative outcome punishing humans for their presumption, and standing in the way of cooperation and progress. But the Warramurrungunji myth reflects a point of view much more common in small speech communities: that having many languages is a good thing because it shows where each person belongs. Laycock quotes a man from the Sepik region of Papua New Guinea saying "it wouldn't be any good if we all talked the same; we like to know where people come from."[1] The Tzotzil oral traditions of the Mexican Chiapas give another twist to this tune: "while the sun was still walking on the earth, people finally learned to speak (Spanish), and all people everywhere understood each other. Later the nations and municipios were divided because they had begun to quarrel. Language was changed so that people would learn to live together peacefully in smaller groups."[2]

I recently drove down the dusty road from Wilyi on the coast near Croker Island, to the inland town of Jabiru (figure 1.2), while working with speakers of Iwaidja, the language in which Tim Mamitba (figure 1.1) had told me the Warramurrungunji story. The 200-kilometer transect follows Warramurrungunji's path, traveling inland and southwards from beach through eucalyptus savannah, stretches of tropical wetlands and lily ponds, and occasional sandstone outcrops whose caves hold vast galleries of rock paintings. It is a timeless landscape rich in wild food – magpie geese, fish, bush fruits, and yams. Its Aboriginal inhabitants can live easily through the year, finding all they need on their own clan countries. The few river crossings do not present geographical barriers. But Warramurrungunji's legacy of linguistic diversity is clearly here. In a few hours on the road we passed through the territories of nine clans and seven languages from four language families, at least as different from each other as Germanic, Slavic, Indo-Aryan and Romance (see table 1.1).

To give a rough idea of how different the languages are at the two ends of this transect, consider the useful sentence "you eat fish." Taking Iwaidja from one end, and Gun-djeihmi from the other, we compare *kunyarrun yab* and *yihngun djenj* – of which only the final *-n* in the two languages, which marks non-past tense in both, is historically relatable. Imagine I had driven from London to Moscow – 15 times as far. The Russian equivalent *ty esh rybku*, although incomprehensible to English ears, contains three cognate elements, at least if we cheat a bit by taking the earlier English version *thou eatest fish*: *ty* (with English *thou*), *e* (with English *eat*) and *-sh* (with the older English suffix *-est* in *eatest*). And if I satisfy

Figure 1.2 Clans and languages in northwestern Arnhem Land

myself with a shorter trip to Berlin – still more than five times the Wilyi–Jabiru trip – we get the almost comprehensible *du ißt (isst) Fisch*, in which every element is cognate.

Some of these languages are now down to just a couple of speakers (Amurdak) or have recently ceased to be spoken (Manangkardi), but others are still being learned by children. Bininj Gun-wok, the largest, now has about 1,600 first-language speakers as members of other groups shift to it. But the average population per language in this region is much smaller, probably less than 500 speakers. And many are even smaller: a recent study by

Table 1.1 Clans and languages along the 200-kilometer track from Wilyi to Jabiru[3]

Clan	*Language*	*Language family*
Murran	*Iwaidja*	Iwaidjan; Iwaidjic
Manangkali	*Amurdak*	Iwaidjan; Southern
Minaka	*Manangkardi*	Iwaidjan; Iwaidjic
Born/Kardbam (Alarrju)	*Bininj Gun-wok (Kunwinjku dialect)*	Gunwinyguan (Central)
Mandjurlngun	*Bininj Gun-wok (Kunwinjku dialect)*	Gunwinyguan (Central)
Bunidj	*Gaagudju*	Gaagudjuan (Isolate)
Mandjurlngun Mengerr	*Mengerrdji*	Giimbiyu
Manilakarr	*Urningangk*	Giimbiyu
Bunidj Gun-djeihmi, Mirarr Gun-djeihmi	*Bininj Gun-wok (Gun-djeihmi dialect)*	Gunwinyguan (Central)

Rebecca Green[4] on Gurr-goni, a few hundred kilometers to the east of the Warramurrungunji track, suggests it has been quite stable for as long as anyone remembers, never with more than around 70 speakers.

Each person from this region has one "father language," which they have special rights in, by virtue of the clan membership they get from their father. This vests them with authority and spiritual security as they travel through their ancestral lands. In traveling to places that have not been visited for some time, clan members should call out to the spirits in the local language, to show they belong to the country. Doing this with visitors is the duty and right of a host. It is said that many resources, such as springs, can only be accessed if you address them in the local idiom. For these reasons there are intimate emotional and spiritual links between language and country. Travelers sing songs listing the names of sites as they move through the land, and switch languages as they cross creeks and other clan boundaries. In epics of ancestral travels it is common to flag where the characters have got to simply by switching the language the story is told in – as if the *Odyssey* were told not just in Greek, but in the half a dozen ancient Mediterranean languages Ulysses would have encountered in his travels.

Throughout Aboriginal Australia, speaking the appropriate local language is a kind of passport, marking you – both to local people and to the spirits of the land – as someone known and familiar, with the right to be there. I once went out in a boat with Pat Gabori to map a Kayardild site a few kilometers off shore, in the company of several Kayardild-talkative senior women and a few children who did not know their ancestral language. Pat and the women called out in Kayardild to the spirits and ancestors of the place, identifying themselves and introducing the silent children, and explaining gently that the children's inability to speak Kayardild did not make them strangers – they just hadn't learned the language yet.

A more extreme illustration of this principle comes from a story Pluto Bentinck, another old Kayardild man, related during a Native Title claim. When asked if traditional

law included sanctions to be taken against trespassers, he cited an incident during World War II, when a hapless white airman swam ashore on Bentinck Island after his plane crashed in the sea. Pluto told me the man had said *danda ngijinda dulk, ngada warngiida kangka kamburij* ("this is my country, I just speak this one language"), as he struggled ashore without his Berlitz Kayardild phrasebook. When I asked him how he knew what the man had said, when he himself knew no English, Pluto replied: *Marralwarri dangkaa, ngumbanji kangki kamburij!* ("He was an ear-less (crazy) man, he spoke your language!"). Speaking English on Bentinck Island, in Pluto's view, was tantamount to claiming it for English speakers. *Nyingka kabatha birdiya kangki! Ngada yulkaanda mirraya kangki kabath!* he had replied to the man ("You found the wrong words! I've found the right words, since forever"). *Ngada bunjiya balath, karwanguni*, Pluto continued: "And I clubbed him in the back of the neck".[5]

Normal members of Arnhem Land society are highly multilingual, often speaking half a dozen languages by the time they are adults. This is helped by the fact that you have to marry outside your clan, which likely means your wife or husband speaks a different language from you. It also means that your parents each speak a different language, and your grandparents three or four languages between them. The late Charlie Wardaga, my Ilgar teacher, was typical. Knowledge of Ilgar, Manangkardi, Marrku, Iwaidja, and Kunwinjku came to him from his grandparents and parents. Although he lived mostly on lands where Ilgar, Marrku, and Iwaidja were the locally appropriate languages, he married a Kunwinjku-speaking woman from a mainland clan and would regularly speak Kunwinjku with her and her relatives, or when traveling to distant communities as a songman. In this system your clan language is your title deed, establishing your claims to your own country, your spiritual safety and luck in the hunt there. Meanwhile the knowledge of other languages gives you the far-flung network of relatives, spouses actual and potential, ceremonial age-mates and allies, which makes you someone who counts in the greater world. This combination of highly developed multilingualism with strong attachments to small local languages is by no means an Arnhem Land oddity – around the world, it is common in zones of high linguistic diversity, like Nagaland in northeastern India, or the Mandara Mountains of Cameroon (see box 1.1).

Most non-Aboriginal people are astonished when they learn how many demographically tiny languages etch their distinctive local domains across the Australian map. Modern citizens of industrialized countries like Britain or Japan take it for granted that they can use their languages with hundreds of millions of people and that a single language occupies the entire territory of their nation, bar dialect variation, immigrants, and one or two beleaguered minorities like Welsh or Ainu. For speakers of big languages, the question is: why are there so many languages in the world/in Papua New Guinea/in Australia/etc.? The naïve explanation sometimes offered, that they result from mutual isolation in distant valleys and gorges, just does not bear up. In Arnhem Land there are no significant geographical barriers at all. And marriage patterns, in Arnhem Land or the Vaupés region of Amazonia, mean that several languages are spoken on a daily basis inside the one household – hardly a case of mutual isolation.

But maybe we are approaching the problem from the wrong end. Doesn't it make more sense to turn the question round and ask, not why Melanesia, the Amazon, Arnhem Land,

the Cameroon, or the Caucasus have so many languages, but rather why Europe or parts of Asia have so few?

Indeed, there are good reasons to believe that our little transect through Arnhem Land is a good representation of how humans have been for most of our past – not just for the 99 percent or so of our history up to 10,000 years ago when we were all hunter-gatherers, but also for much of the time that followed. This is because the dawn of agriculture, although it led to an explosion in human populations, did not automatically lead to the development of much larger societies. Speech communities got a bit bigger but it was rare for them to exceed the few score thousand that could be held together as a homogeneous unit without the panoply of state control that only began with our incorporation into large centralized political entities like the Roman Empire or a modern nation state. Mapping the million or more years of human history onto a 24-hour clock, incorporation into large centralized states did not start for any human society before ten minutes to midnight

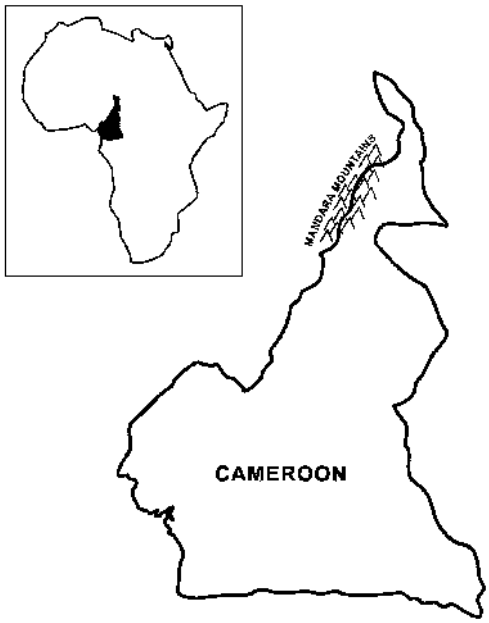
Box 1.1 The many paths to multilingualism in up-country Cameroon

Jonas courting Gogo in Jilve village in the presence of other villagers (photo: Leslie Moore)

Jonas, the boy in this photo, comes from the village of Jilve in the Mandara region of Cameroon, another region of daunting linguistic diversity, where people speak small "montagnard" languages of the Chadic family, very distantly related to the Semitic languages. Here he is shown courting Gogo, the girl he wants to marry, in her mother's compound, with her girlfriends in attendance. They are speaking primarily in Mada, Gogo's paternal language. Mada is one of eight languages that

(five millennia ago in the Fertile Crescent). For many groups it has only begun to happen in the last seconds.

The island of New Guinea and its Melanesian surrounds, a few hundred kilometers to the north of Warramurrungunji's territory, is a good illustration of a region almost completely made up of Neolithic agricultural societies, with no centralized states until recent colonization by Europeans and Indonesians. Its population of around 10 million people speaks some 1,150 languages – under 10,000 people per language. In the Central Highlands, where the population density is highest thanks to intensive agriculture and pig breeding, elaborate networks of production and ceremonial exchange have gradually bound people together into larger speech communities. The more intensified the system, the more speakers per language. But, even in the most elaborate and intensified highland Papuan communities, the average number of speakers per language rarely exceeds 40,000. And in many other parts of Melanesia languages of that size seem unimaginably large: the nation



The Mandara mountains, Cameroon

Jonas speaks. Although he began learning Mada in order to court Gogo, the two of them already had two languages in common: Wandala (the local lingua franca) and Wuzlam, the first language of Jonas' father and of Gogo's mother. Prior to this visit, Jonas had prepared a list of topics of conversation and relevant Mada vocabulary, which he had noted on a piece of paper he brought with him but did not consult during the visit.

We know very little about how the impressive levels of multilingualism are acquired in small-scale preliterate societies, but Leslie Moore's pioneering ethnographic work (from which the above vignette is taken) has taught us something about how multilingualism works in Cameroon communities like Jilve. Besides the "normal" acquisition of their mother tongue, children learn French and later English in school, the regional lingua franca Fulfulde from night-time storytelling by their elders, and the languages of neighboring villages through self-instruction of the type we see here. From an early age parents ask their children to memorize messages, in languages they do not yet know, and to go and deliver them orally to people from neighboring villages. Even young children develop a strong metalinguistic awareness, for example using knowledge of cognates in related languages to help them remember new vocabulary.

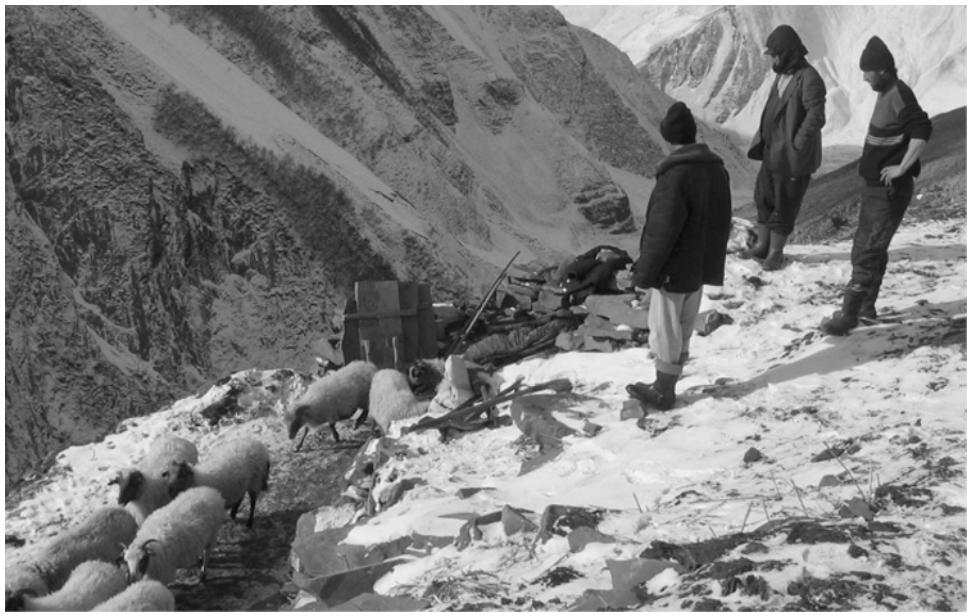


Figure 1.3 Archi men herding sheep into a *maɬɬi* or underground sheep fold (photo: Marina Tchoumakina)

of Vanuatu (total population 195,000), most of whose population are village agriculturalists, counts 105 languages – an average of less than 2,000 speakers per language for the whole country! Apart from the recently developed national lingua franca, Bislama, its biggest language (Lenakel) has just 11,500 speakers and only 13 languages have 5,000 speakers or more.[6]

We see small languages wherever in the world societies have lain beyond the homogenizing reach of great empires. But the situation is most extreme where groups can maintain themselves self-sufficiently without needing to call on the hospitality of others. The village cluster of Archib (population 1,237) in the Caucasus is the only place in the world where the Archi language is spoken – a language whose morphology is so complex that it has been calculated that a single verb possesses more than 1.5 million inflected forms. Most of its inhabitants will be born, married, and laid to rest in this one village, basing their economy on specially adapted mountain sheep, which they tuck up on freezing winter nights into special underground sheep folds called *maɬɬi* (see figure 1.3). Or, in northwestern California, the entire territory of the Chimariko people and their language consisted of a 20-mile stretch along a narrow canyon of the Trinity River.[7] Until the Gold Rush, their economic self-sufficiency on this small patch was assured by the rich salmon stocks in the river.

There is evidence from many parts of the world that small groups in favored areas did not simply rely on the drift of time to carry their languages apart from those of their neighbors. In northern Australia the reigning ideology is that each clan should have its own

distinct language variety. This then sanctions the investiture by tribal elders of variant forms as proper to their local languages, driving along a relentless diversification.

Peter Sutton, working on clan identities in the Cape York Peninsula of Australia, reports cases where the fission of clans is rapidly followed by the emergence of new language varieties. In settings where fewer than a hundred people may speak a "clan lect," one or two powerful individuals can readily impose what may have started out as individual idiosyncrasies, and seed the emergence of a new system. In Iwaidja, many forms of nouns and verbs mutate their initial consonants – "his or her arm" is *bawurr*, from the root *mawurr*, for example, whereas the corresponding words in related languages like Charlie Wardaga's language Ilgar keep the original *m*. The mutated forms are based on an obscure "miscellaneous" gender so rare in all the languages of the family that it would never have won out as the standard form by processes of normal change. More likely, at some point in the past, Iwaidja speakers deliberately extended the use of the miscellaneous gender to set their language off from their neighbors, on the "you say tomahto, I say tomato" principle.

In New Guinea language differentiation is sometimes fostered even more deliberately. When we compare the Uisai dialect of Buin (1,500 speakers), on Bougainville Island, with the other dialects of Buin (about 17,000 speakers all up), we see that it has completely flipped over all its gender agreements:[8] all the masculines have become feminine, and all the feminines have become masculine. Because no known mechanism of normal linguistic change could produce this effect, Don Laycock has suggested that "an influential Uisai speaker innovated a linguistic change to differentiate his community from the rest of the Buins." Again we see how much influence a single individual can have in a small speech community. For another Papuan language, Selepet, we actually have a reported instance, witnessed by linguist Ken McElhanon, where one community decided at a meeting to replace the standard Selepet word *bia*, for "no," with the word *bunge*, to differentiate themselves from other Selepet villages.[9]

Although for illustrative purposes we have concentrated on decisions affecting a single word or grammatical feature, this is just the thin end of the wedge. William Thurston studied "esoterogeny" – the engendering of difference and linguistic obscurity – with Anem speakers on the island of New Britain, off the New Guinea mainland. He found that "esoterogenic" languages tend to streamline pronunciation in ways that make the overall structure harder to see, comparable to saying *dja* for *didja* from *did you* in English. They replace clear regular relationships with "suppletive" (totally irregular) ones, reveling in alternations like *good:better* at the expense of the more transparent *big:bigger* style. They have huge numbers of opaque idioms, of the *kick the bucket* type, and entrench prescriptive traditions that limit flexibility of language: "you must speak this way to be a member of our community!" They also elaborate terminology to make subtle distinctions, and speakers take pride in the greater richness of their language than the neighboring language of Lusi in this regard.

During Thurston's research on Anem he found that "some of the boys had devised a competitive word game aimed at exposing one another's ignorance of the name for an obscure vine or bush; in order to keep ahead, boys were asking older people, secretively, for words they could use to try tricking other boys." All these forces conspire to maximize difference between one language and its neighbors – although I should stress that, up till now,

we have no more than the sorts of anecdotal evidence reported on here, and systematic studies of the causes and processes of change in small languages are badly needed.

Small-scale societies in such parts of the world are economically self-sufficient, and proudly form the center of their own social universe without needing to defer unduly to more powerful outside groups. Their constructive fostering of variegation – which holds social groupings to a small and manageable size, and keeps outsiders at a suitable distance – is not offset by the need to align their language with large numbers of other people in the world. The great Swiss linguist Ferdinand de Saussure saw language as being pulled in opposite directions by the "spirit of the steeple" – the parochialism of showing which little community you belong to – and the "spirit of wider communication." But, for those small-scale societies able to subsist mostly on their own resources, the force of the steeple is dominant.

Language Diversity through Time and Space

The classic estimate of the world's population on the eve of the Neolithic, ten millennia ago, is 10 million.[10] Combining these figures with a very generous 2,000-speaker maximum for hunter-gatherer languages suggests that, on the eve of agriculture and fixed settlements, there were already from 3,000 to 5,000 languages in the world – roughly the same number as now, even though the population was less than half a percent of its current level. If we assume 1,000 speakers per language, a more realistic figure in my view, the number doubles to between 6,000 and 10,000.[11] Levels of language diversity whose full magnitude we can barely grasp have been with us for a very long time.

Going much further back in time, to the population bottleneck about 150,000 years ago that preceded the long trek out of Africa, Rob Foley estimates from mitochondrial DNA that there were probably between 10,000 and 20,000 women of reproductive age – say 50,000 humans all up, who on the language-size estimates given above would already have been speaking 10–20 distinct languages (and possibly a hundred already if we go with the plausible language population of 500). Already then, at a point when humans went through a population bottleneck probably caused by environmental crises, just before their fortunes turned and they ventured forth across the planet, there were scores of languages. Since our forebears probably began to speak and develop languages long before that,[12] the figure 150,000 years ago is likely to have included tongues that were already quite different from one another.

Let us come back to the emergence of agriculture at the beginning of the Neolithic. As agriculture spread around much of the world from then on, it is likely that the increases in language populations of cultivating groups – perhaps to New Guinea-like levels of around 10,000 – would have been more than offset by the explosion in overall world population, so that the number of languages in the world may have risen to 10,000 or 15,000. However, scholars like Colin Renfrew and Peter Bellwood[13] have argued that the bearers of agricultural expansion would have been just a small number of groups who had made a radical cultural transition. From the hunter-gatherer perspective, agricultural life looks pretty unappealing. In return for the security of regular crop supplies you have to put up with a poorer and less varied diet, monotonous year-round residence in a circumscribed

area, and the diseases that come from having so many people living in relatively close quarters, on top of each other's excrement. Most hunter-gatherers would have walked away from the deal. Inexorably, though, the "demic expansion" that agriculturalists could feed from the increased food yields they drew from the land would gradually have squeezed out or assimilated the original populations of hunter-gatherers.

The expansion of these few agriculture-based lineages would have produced what I will call the Renfrew–Bellwood effect: a decrease in deep-level diversity, i.e. in the number of unrelated stocks or deep lineages, as clusters of closely related languages spread outward from the dozen or so foci where agricultural complexes were developed, obliterating the deep-time variability that was there beforehand. If agricultural expansion is old enough, however, there is time for significant new multiplicity to develop, as has happened with the Afro-Asiatic languages stretching from Hausa in Nigeria to Hebrew and Arabic in the Middle East, or the Austric family in Southeast Asia. Each of these families probably goes back to the very dawning of agriculture in their respective regions, and contains languages so different that it has taken many years of work to demonstrate their relatedness – a topic we return to in chapters 5 and 6.

We can see the Renfrew–Bellwood effect clearly in Indonesia where, over most of the archipelago, no linguistic trace remains of the hunter-gatherers who must have occupied its fertile lands until the coming of the Austronesian agriculturalists a few thousand years ago. It is also clear in New Guinea, where the whole fertile highland cordillera along which root-and-pig farmers expanded over the last few millennia is occupied by a single Trans New Guinea family, albeit one with around 400 member languages. New Guinea's most mind-boggling lineage diversity is found in lowland areas like the Sepik and the Trans-Fly region, north and south of the cordillera, where people either practice a much less intensive form of agriculture or mix agriculture, fishing, hunting, and gathering.

We have seen already, though, that even in the densely populated New Guinea Highlands, and even after six to seven thousand years of intense agriculture, there were no really large languages. This is because it was only with the advent of centralized and then industrialized state societies that a few languages began to spread to the point where they counted hundreds of thousands and then millions of speakers. Unfortunately, most of these expansionist new societies had no interest in recording anything about the languages of the peoples they subjugated, as we will see in the next chapter. But we can get some idea of what the world was like as the first great empires emerged by looking at the Italian peninsula in the fifth and sixth centuries BC. There, under Greek influence, a number of different civilizations developed their own writing systems in time to leave some record of their languages before they were all sucked into the Latin-speaking vortex of the Romans.

Inscriptions in pre-Roman Italy attest between 12 and 15 distinct languages, quite different from one another, and belonging to between 5 and 10 branches of at least 4 distinct families – 3 branches of Indo-European (Celtic, Italic, and Greek) plus Etruscan, which was non-European. The Romans did not actively try to stamp out other languages – indeed, the retention of other languages by non-Romans favored the policy of *diuide et impera* ("divide and rule"). Umbrians, for example, continued to make inscriptions in their language for centuries after Roman annexation. But eventually the power and status of Latin prevailed, particularly after all residents of Italy became Roman citizens in the middle of

the last century BC. At first other groups would just have used Latin for "outside" purposes, but gradually the centralizing power of Rome "relegated the local speech, just as it did political initiative and concerns, to a secondary, subordinate, and ever retreating position."[14] We do not know exactly when the last speakers of Oscan, Umbrian, Etruscan, and other languages of the peninsula finally passed away, but the elimination of all non-Latin languages from the Italian peninsula is likely to have been almost complete by the time Herod washed his hands of another death in another part of the Roman Empire.[15]

Emerging kingdoms in many parts of the world – in Egypt, Arabia, Persia, Mali, China, Korea, India, Mexico, the Andes – had similar impacts on the smaller peoples in their domains. It was probably in this period – beginning around 2,000 BC – that the first languages with more than a million speakers emerged. Expansionist agricultural–military complexes like the Bantus in the southern half of Africa obliterated vast mosaics of diversity. And then, from 1492, European colonial expansion began to take its toll. Little more than half a century after the Spaniards reached Cuba and Puerto Rico, the Arawakan language Taino would cease to be spoken, although some of its words have survived as loanwords into Spanish (*cacique* "chief") and others passed further into English ("barbecue" < *barbacoa*; "canoe" < *canoa*; "tobacco" < *tabaco*). Thousands of other languages around the world would suffer a similar fate, leading to the accelerating loss of linguistic diversity we see today, and the concomitant dominance of the dozen or so languages with more than a hundred million speakers.

Where the Hotbeds Are

The upshot of what we have been discussing is that language diversity is now distributed very unevenly around the world. On one estimate,[16] 17 countries hold 60 percent of all languages, although these countries make up only 27 percent of the world's population and 9 percent of its land area.[17] Table 1.2 shows two slightly different rankings of the top 25 language-diverse countries – a calculation of sheer number of endemic languages,[18] and an alternative measure showing the number of linguistic lineages, which is a better measure of deep-level language diversity. It also shows the top 25 countries for biological diversity, for reasons we will return to below.

As the preceding discussion should have made clear, the current distribution of languages reflects many influences. The effects of each region's history have been superimposed on original patterns that are likely to have shown even closer parallels between linguistic and biological diversity:

(1) An original stratum of deep-time language diversity goes back to when all humans were hunter-gatherers. This is visible in regions where people have remained hunter-gatherers until recently. Here there is high language diversity on both measures (i.e. total number of languages and number of independent lineages) except in spread zones such as deserts and other less favored regions that show the effects of repeated recolonization and cultural pressures to extend intercommunicating networks.

Table 1.2 The top 25 megadiverse countries, for endemic species and two measures of language diversity. Left and right columns reprinted with permission from Harmon (1996)

Rank	*Endemic languages*		*Endemic linguistic lineages*		*Endemic higher vertebrate species*	
1	PNG	847	USA	64	Australia	1,346
2	Indonesia	655	PNG	58	Mexico	761
3	Nigeria	376	Indonesia	37	Brazil	725
4	India	309	Brazil	31	Indonesia	673
5	Australia	269	Mexico	24	Madagascar	537
6	Mexico	230	Colombia	24	Philippines	437
7	Cameroon	201	Australia	22	India	373
8	Brazil	185	Peru	21	Peru	332
9	Zaire	158	Russia	17	Colombia	330
10	Philippines	153	Sudan	15	Ecuador	294
11	USA	143	Canada	14	USA	284
12	Vanuatu	105	Bolivia	13	China	256
13	Tanzania	101	Venezuela	11	PNG	203
14	Sudan	97	India	10	Venezuela	186
15	Malaysia	92	China	9	Argentina	168
16	Ethiopia	90	Ethiopia	8	Cuba	152
17	China	77	Chad	8	South Africa	146
18	Peru	75	Argentina	8	Zaire	134
19	Chad	74	Ecuador	8	Sri Lanka	126
20	Russia	71	Nigeria	6	New Zealand	120
21	Solomon Islands	69	Burkina Faso	6	Tanzania	113
22	Nepal	68	Tanzania	5	Japan	112
23	Colombia	55	Cameroon	5	Cameroon	105
24	Ivory Coast	51	Georgia	5	Solomon Islands	101
25	Canada	47	Chile	5	Ethiopia	88
			Laos	5	Somalia	88

(2) A second stratum results from small-scale agricultural expansion since the end of the Neolithic in some regions, although it was much more recent in some areas. This expansion wiped out hunter-gatherer languages but regrew a more recent pattern of diversification – with more or less lineage density depending on the time-depth of agriculture – leaving a pattern of large numbers of languages groupable into deep-level families like Indo-European or Austronesian.

(3) The effects of state formation, between 3,000 BC and AD 1,000 depending on the area, produced a steady fall in linguistic diversity. Thus China has a relatively low score, of just 77 endemic languages – fewer than its tiny southern neighbor Laos. This undoubtedly reflects the gradual assimilation of minority populations into the majority Han population over the millennia of centralized state rule, with a shift to speaking Mandarin and other

Chinese varieties. In most countries of Europe, North Africa, and the Middle East the low scores are due to comparable effects over the last two or three millennia.

(4) Most recently the effects of expansionist colonization by Europeans and the elites of the nation states they created have been to wipe out indigenous linguistic diversity in many of their colonies. In Uruguay, Cuba, Haiti, and all islands of the Caribbean – ironically, the only sea to be named after an indigenous language (Carib) – have the dubious distinction of having completely silenced their indigenous languages.

If we look at countries like Australia, the USA or South Africa, colonized by Europeans in the last few centuries, we get very rapid rates of language death under the impact of English. Comparably rapid rates of language extinction are occurring in much of Brazil under the impact of Portuguese, in Siberia under Russian, in the Sudan under Arabic, throughout Indonesia under Indonesian, and in even quite remote parts of Papua New Guinea under Tok Pisin, the newly developed national lingua franca.

Looking back beyond the recent flattening of multiformity by colonial languages – and the scores in table 1.2 largely bracket this off by giving known numbers of languages at the moment of colonial contact – we can see strong correlations between linguistic and biological diversity. Arizona linguistic anthropologist Doug Harmon first looked at this correlation in an important 1996 study, and since then his findings have been replicated worldwide on a country by country basis,[19] confirmed for Africa at a coarse resolution,[20] and at quite fine resolutions for the Americas[21] using passerine birds as the index of biodiversity and sampling geographical cuts down to squares one degree wide. More recent approaches to this question have used broad ecological areas instead of the rather accidental boundaries given by countries, and again found strong correlations.[22] Ten of the world's top dozen "megadiversity" countries on biological measures also make it into the A-league of the world's top 25 countries for endemic linguistic diversity.[23] Harmon's work also makes it clear where both types of diversity are concentrated: Central and South America, tropical Africa, and South and Southeast Asia across through Indonesia, Melanesia, and Australia to the Western Pacific.

The Wellsprings of Diversity in Language, Culture, and Biology

> [Continued loss of biocultural diversity will] staunch the historical flow of being itself, the evolutionary processes through which the vitality of all life has come down to us through the ages.
>
> (Harmon 2002:xiii)

The arguments for conserving diversity are similar whether we consider the loss of a rare bird or tree species, a body of cultural knowledge that will soon be forgotten, or an endangered language. Since Darwin, we have begun to articulate, at the scientific level, what most cultures have had enshrined in their aesthetics and cosmologies for a long time: that variety is the reservoir of adaptability.

Having a genetically homogenous population of Cavendish bananas is great for maximizing yield and efficiency, but it just takes one new strain of fungus to wipe out the worldwide population. The traditional (agri)cultural practices that were displaced by the technology-driven Green Revolution, with its fertilizers, standardized seeds, and productive new breeds, are increasingly being seen as having strengths that were often overlooked 40 years ago: drought-resilience, disease resistance, lower demands on local water-tables – and, ironically, higher yields if we factor in water consumption rather than just tons of yield per area. The many "land-races" of traditional agriculture around the world are now being genetically archived,[24] so that at least it is possible to draw on their genetic diversity. But we also need all the cultural knowledge that grew up around them. Which variety should be planted in which conditions? What grows best where? Which is resistant to what crop disease? Which are good companion plants? Agriculture, like all technology, only works when artifacts marry know-how, and if we only store seeds without the accompanying knowledge we still have an impoverished picture.

Within the western scholarly and scientific traditions, we can identify two types of attitude to global knowledge. One is universalizing, and sees it as possible to incorporate all knowledge into the world language of the era – Latin, Arabic, French, and English have all had their turn – spoken by a "unified mankind within a single unified realm, subscribing to a universal value system."[25] The other recognizes the strength and richness that comes from distinct traditions that can never be straightforwardly mapped onto a single value system speaking a single world language: "any reduction of language diversity diminishes the adaptational strength of our species because it lowers the pool of knowledge from which we can draw."[26]

Joshua Fishman outlines this alternative view, showing how it has been developed by a succession of thinkers from Vico and Herder to Boas, Sapir, and Whorf:

> the entire world needs a diversity of ethnolinguistic entities for its own salvation, for its greater creativity, for the more certain solution of human problems, for the constant rehumanization of humanity in the face of materialism, for fostering greater esthetic, intellectual, and emotional capacities for humanity as a whole, indeed, for arriving at a higher stage of human functioning . . . the great creative forces that inspire all humanity do not emerge out of universal civilization but out of the individuality of separate ethnic collectivities – most particularly, out of their very own authentic languages. Only if each collectivity contributes its own thread to the tapestry of world history, and only if each is accepted and respected for making its own contribution, can nationalities finally also be ruled by a sense of reciprocity, learning and benefiting from each other's contributions as well. (Fishman 1982:7)

Fishman's wording makes it clear that this is neither a plea for the thousands of particularistic small societies to become mutually isolated museum pieces, nor for a few cute local words to be lifted into a world language like English to form a sort of linguistic theme-park. Rather, it recognizes the deep creative interactions and synthetic insights that come up when we look at one language or culture through the prism of another. A few hundred kilometers east of Warramurrungunji territory, the Yolngu peoples of eastern Arnhem Land, in conceptualizing the value of the bilingual and bicultural schools they are striving to develop on their lands, employ the Yolngu metaphor of *ganma*. This denotes

the special mixing that develops when the outgoing freshwater current of a river mingles with the saltwater of the incoming tide.

I will assemble the evidence for the humanistic and scientific value of Warramurrungunji's bequest as this book unfolds. But to close this chapter, as we have been talking about the relation between linguistic diversity, species diversity, and ecology, let us look at some of the ways in which small languages hold detailed biological and ecological knowledge, which generations of speakers have gradually discovered and recorded in their languages. These examples will also illustrate the point – developed eloquently in Daniel Nettle and Suzanne Romaine's book *Vanishing Voices* (2000) – that the loss of our linguistic heritage is intimately tied up with the loss of cultures and habitats. The sorts of knowledge and vocabulary I am about to turn to are typically the first things to be lost when speakers of a language are shifted from their traditional lands to a reserve in Oklahoma, a rubber plantation in Malaysia, a ghetto in Ibadan, or pushed into a sedentary lifestyle where they no longer practice their traditional ecological knowledge.

Words on the Land

> [I]ndividuals draw on cultural resources to structure and accomplish problems with which they engage in everyday socially organized activities.
>
> (Saxe and Esmonde 2005:173)

Small languages and societies have kept their place in the world by being finely tuned to their local ecologies and amassing a rich fund of knowledge about them. Much of this has been carried forward just in their languages. Many aspects of their traditional knowledge are still unknown to western science, and in fact languages are arguably the most important and distinctive of the "cultural resources" that Saxe and Esmonde are referring to in the above quote.

Consider Seri, spoken by around 500 hunter/gatherer/fisherpeople in Baja California in Mexico. This is most probably an isolate language without known relatives, although some linguists argue it is a southern outpost of the Hokan languages of California. In the course of documenting the Seri lexicon, linguists Edward and Mary Moser were told by Seri speakers about their use of eelgrass (*Zostera marina L.*) as a source of grain, leading to the involvement of ethnobotanist Richard Felger. The resultant research was published in *Science* with the appropriate title "Eelgrass (Zostera marina L.) in the Gulf of California: discovery of its nutritional value by the Seri Indians."[27] The authors concluded that this is the only known case of a grain from the sea being harvested as a human food source, and emphasized its considerable potential as a general food resource for humankind, which can be cultivated without fresh water, pesticides, or artificial fertilizer. Despite its potential importance in a world likely to need new crops, this crucial knowledge had been locked up inside the almost impenetrable Seri language, known only to members of the tiny Seri world. Many other words in Seri contain information about the treatment, products, and harvesting of eelgrass. For example, the month of April is called *xnois iháat*

Figure 1.4 A group of Seri people in what outsiders call the "desert," but which they call *heheán* ("place of the plants") (photo: Arizona State Museum)

iizax ("moon of the eelgrass harvest"), and the onset of harvest time is signaled when the black brant bird known as *xnois cacáaso* ("the foreteller of eelgrass seed") dives into the sea to feed on the plant.

All around the world indigenous people transmit, through the words and expressions of their languages, the fruits of millennia of close observation of nature and experimentation with its products. In Arnhem Land there have been a number of cases where the impulse for western natural scientists to recognize new species has come from indigenous traditions of taxonomic naming. The large and striking Oenpelli python, long known to Kunwinjku speakers as *nawaran*, is one such species: it was first incorporated into western scientific taxonomy in 1977 as *Morelia oenpellensis*.[28]

Meanwhile a host of native bees of great economic importance to local Aboriginal people, each with their own honey and wax types, remain unidentified by entomologists. For the time being their English identification in bilingual dictionaries is limited to the inadequate designation *trigona species*.

Traditional cultures also contain detailed knowledge of the healing properties of plants, transmitted in local languages. The recent discovery of a drug, prostarin, effective against HIV-type 1 goes back to a conversation between Samoan tribal healer Epenesa Mauigoa and ethnobotanist Paul Allen Cox about traditional medicinal uses of the stem of a particular tree, *Homalanthus nutans*. The fact that Cox had learned Samoan as a missionary's son was a key fact in enabling this conversation. Comparable curative potentials abound across the immensely variegated world of traditional ethnobotanics, and their

full investigation requires the collaboration of traditional healers, ethnobotanists, and linguists.

Vocabularies of indigenous languages often also show the ecological links between particular plant and animal species. Throughout Arnhem Land the spangled grunter fish bears the same name as the native white apple tree, *Syzygium eucalyptoides*, because this fish eats the fruits that fall from this tree into creeks and billabongs:[29] in Kunwinjku both are called *bokorn*. Knowledge of this link is of obvious value to anyone who happens to be out fishing for spangled grunter: look for the tree, and in the water below you are likely to find its fish "mate." The languages of Central Arnhem Land abound in such pairings, making them a veritable fisherman's guide to the area.

As another example of traditional ecological knowledge, consider the Mparntwe Arrernte language of the Alice Springs area, where various types of grub are an important source of food. Mparntwe Arrernte has a special method of naming grubs after the bushes where you can find them: *tnyeme* ("witchetty bush") yields the *tnyematye* ("witchetty grub"), *utnerrenge* ("emu bush") yields the grub known as *utnerrengatye*, and you can work out for yourself the name of the grub found in *thenge*, the ironwood tree.

A host of examples like these lead out into the vast ethnobiological wings of our Library of Babel. But all such knowledge is at great risk, as long as it is only available in little-known languages spoken by just a few hundred people, since a shift to another language can cut off its transmission. Once we go over to calling the *bokorn* fish a "spangled grunter," and the *bokorn* tree a "white apple," our words no longer deliver the ecological link between them.

Further reading

Trigger (1987) and Garde (in press) contain detailed discussions of locally appropriate language uses in Aboriginal speech communities. Feil (1987) gives demographic figures for the New Guinea Highlands showing the correlation between intensification of agriculture and language size. The fascinating multilingual situation in the Amazonian Vaupés is discussed in Jackson (1983) and Aikhenvald (2002). Moore (2004) examines how children become multilingual in upcountry Cameroon, and Thurston (1987, 1992) gives a rare and entertaining watch-it-happen study of the processes and results of "esoterogeny" in New Britain. Sutton (1978) discusses Cape York clan lects, and the origins of Iwaidja mutation are in Evans (1998); Evans (2003a) sets out a model for how ideologies of difference can drive the emergence of clan and other variation. For the situation in pre-Roman Italy see Pulgram (1958), and Robb (1993) on ancient Europe more generally. Ostler (2005) gives a wide-ranging discussion of the fates of languages and past empires in the historical period. On the Taino language of Cuba see Álvarez Nazario (1996).

Collard and Foley (2002) also compare the distribution of biological and linguo-cultural diversity, pointing out the need to take historical forces into account on top of the basic ecological determinants. Maffi (2001, 2005) treat the links between language and biodiversity, while Nettle (1999) is a pioneering essay on the causes of language diversity. The best worldwide listing of the world's languages is Ethnologue (www.ethnologue.com), maintained by the missionary organization Summer Institute of Linguistics; although this is constantly being updated, it is nonetheless far from accurate or consistent.[30] The best current worldwide classification is that by Dryer (2005), and

The World Atlas of Linguistic Structures, in which it appears (Haspelmath et al. 2005), is a mine of information on the geographical distribution of over 140 linguistic features; its information-packed CD, by Hans-Jörg Bibiko, allows you to compile your own maps and tables from a vast database.

On Seri ethnobotany see Felger and Moser (1973, 1985); the latter includes descriptions of how eelgrass seeds are harvested. Arvigo and Balick (1993), Balick and Cox (1996), Cox and Balick (1994), and Sofowora (1982) discuss ethnobotany and traditional herbalism. The Arrernte examples are from Wilkins (1993), and the book this is in (Williams and Baines 1993) contains an interesting collection of papers on the now-burgeoning topic of traditional ecological knowledge. A good discussion of other highly specific indigenous biological knowledge, and a sustained argument linking the loss of linguistic and biological diversity, are in Nettle and Romaine (2000).