

Language death

DAVID CRYSTAL



CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS

The Edinburgh Building, Cambridge CB2 2RU, UK www.cup.cam.ac.uk
40 West 20th Street, New York, NY 10011-4211, USA www.cup.org
10 Stamford Road, Oakleigh, Melbourne 3166, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain

© David Crystal 2000

This book is in copyright. Subject to statutory exception and
to the provisions of relevant collective licensing agreements,
no reproduction of any part may take place without
the written permission of Cambridge University Press.

First published 2000

Printed in the United Kingdom at the University Press, Cambridge

Typeface Adobe Minion 11/14pt *System* QuarkXPress™ [SE]

A catalogue record for this book is available from the British Library

Library of Congress cataloguing in publication data

Crystal, David, 1941–
Language death / David Crystal.
p. cm.

Includes bibliographical references and indexes.

ISBN 0 521 65321 5 (hardback)

1. Language obsolescence. I. Title.

P40.5.L33 C79 2000

4177–dc21 99-053220

ISBN 0 521 65321 5 hardback

Contents

Preface vii

1 What is language death? 1

2 Why should we care? 27

3 Why do languages die? 68

4 Where do we begin? 91

5 What can be done? 127

Appendix: some useful organizations 167

References 170

*Index of dialects, languages, language families,
and ethnic groups* 182

Index of authors and speakers 185

Subject index 188

1 *What is language death?*

The phrase ‘language death’ sounds as stark and final as any other in which that word makes its unwelcome appearance. And it has similar implications and resonances. To say that a language is dead is like saying that a person is dead. It could be no other way – for languages have no existence without people.

A language dies when nobody speaks it any more. For native speakers of the language in which this book is written, or any other thriving language, it is difficult to envision such a possibility. But the reality is easy to illustrate. Take this instance, reported by Bruce Connell in the pages of the newsletter of the UK Foundation for Endangered Languages (FEL), under the heading ‘Obituaries’:¹

During fieldwork in the Mambila region of Cameroon’s Adamawa province in 1994–95, I came across a number of moribund languages . . . For one of these languages, Kasabe (called Luo by speakers of neighbouring languages and in my earlier reports), only one remaining speaker, Bogon, was found. (He himself knew of no others.) In November 1996 I returned to the Mambila region, with part of my agenda being to collect further data on Kasabe. Bogon, however, died on 5th Nov. 1995, taking Kasabe with him. He is survived by a sister, who reportedly could understand Kasabe but not speak it, and several children and grandchildren, none of whom know the language.

There we have it, simply reported, as we might find in any obituary column. And the reality is unequivocal. On 4 November 1995, Kasabe existed; on 5 November, it did not.

Here is another story, reported at the Second FEL Conference in

¹ Connell (1977: 27). The newsletters of this organization changed their name in early issues. The name was *Iatiku* for Numbers 2–4, and *Ogmios* for No. 6 on. Issues 1 and 5 had no distinctive name, and in this book these are referred to as *FEL Newsletter*.

Edinburgh in 1998 by Ole Stig Andersen.² This time, 8 October 1992 is the critical day:

The West Caucasian language Ubuh . . . died at daybreak, October 8th 1992, when the Last Speaker, Tevfik Esenç, passed away. I happened to arrive in his village that very same day, without appointment, to interview this famous Last Speaker, only to learn that he had died just a couple of hours earlier. He was buried later the same day.

In actual fact, Kasabe and Ubykh (a widely used alternative spelling) had effectively died long before Bogon and Tevfik Esenç passed away. If you are the last speaker of a language, your language – viewed as a tool of communication – is already dead. For a language is really alive only as long as there is someone to speak it to. When you are the only one left, your knowledge of your language is like a repository, or archive, of your people's spoken linguistic past. If the language has never been written down, or recorded on tape – and there are still many which have not – it is all there is. But, unlike the normal idea of an archive, which continues to exist long after the archivist is dead, the moment the last speaker of an unwritten or unrecorded language dies, the archive disappears for ever. When a language dies which has never been recorded in some way, it is as if it has never been.³

The language pool

How many languages are at the point of death? How many are endangered? Before we can arrive at an estimate of the scale of the problem, we need to develop a sense of perspective. Widely quoted

² Andersen (1998: 3).

³ There is, of course, always the possibility that other speakers of the same dialect will be found. In the Ubykh case, for instance, there were at the time rumours of two or three other speakers in other villages. Such rumours are sometimes found to be valid; often they are false, with the speakers being found to use a different dialect or language. But even if true, the existence of a further speaker or two usually only postpones the real obituary by a short time. For some Aboriginal Australian examples, see Wurm (1998: 193). Evans (forthcoming) provides an excellent account of the social and linguistic issues which arise when working with last speakers, and especially of the problem of deciding who actually counts as being a 'last speaker'.

figures about the percentage of languages dying only begin to make sense if they can be related to a reliable figure about the total number of languages alive in the world today. So how many languages are there? Most reference books published since the 1980s give a figure of between 6,000 and 7,000, but estimates have varied in recent decades between 3,000 and 10,000. It is important to understand the reasons for such enormous variation.

The most obvious reason is an empirical one. Until the second half of the twentieth century, there had been few surveys of any breadth, and the estimates which were around previously were based largely on guesswork, and were usually far too low. William Dwight Whitney, plucking a figure out of the air for a lecture in 1874, suggested 1,000.⁴ One language popularizer, Frederick Bodmer, proposed 1,500; another, Mario Pei, opted for 2,796.⁵ Most early twentieth-century linguists avoided putting any figure at all on it. One of the exceptions, Joshua Whatmough, writing in 1956, thought there were 3,000.⁶ As a result, without professional guidance, figures in popular estimation see-sawed wildly, from several hundred to tens of thousands. It took some time for systematic surveys to be established. *Ethnologue*, the largest present-day survey, first attempted a world-wide review only in 1974, an edition containing 5,687 languages.⁷ The Voegelins' survey, published in 1977, included around 4,500 living languages.⁸ Since the 1980s, the situation has changed dramatically, with the improvement of information-gathering techniques. The thirteenth edition of *Ethnologue* (1996) contains 6,703 language headings, and about 6,300 living languages are classified in the *International encyclopedia of linguistics* (1992).⁹ There are 6,796 names listed in the index

⁴ See Silverstein (1971: 113).

⁵ Bodmer (1944: 405). Pei (1952: 285); in a later book (1954: 127), this total decreased by 1.

⁶ Whatmough (1956: 51).

⁷ See now the 13th edition, Grimes (1996); also www.sil.org/ethnologue. The first edition in fact dates from 1951, when Richard S. Pittman produced a mimeographed issue of ten pages, based on interviews with people attending the Summer Institute of Linguistics.

⁸ Voegelin and Voegelin (1977). I used their total in the first (1987) edition of my *Cambridge encyclopedia of language* (Crystal 1997a).

⁹ Bright (1992); the files of *Ethnologue* (then in its 11th edition) were made available for this project, hence the similarity between the totals.

to the *Atlas of the world's languages*.¹⁰ The off-the-cuff figure most often heard these days is 6,000, with the variance sometimes going below, sometimes above.¹¹ An exceptionally high estimate is referred to below.

A second reason for the uncertainty is that commentators know that these surveys are incomplete, and compensate for the lack of hard facts – sometimes by overestimating, sometimes by underestimating. The issue of language loss is itself a source of confusion. People may be aware that languages are dying, but have no idea at what rate. Depending on how they estimate that rate, so their current global guess will be affected: some take a conservative view about the matter; some are radical. (The point is considered further below.) Then there is the opposite situation – the fact that not all languages on earth have yet been ‘discovered’, thus allowing an element of growth into the situation. The ongoing exploration of a country’s interior is not likely to produce many fresh encounters, of course, given the rate at which interiors have already been opened up by developers in recent years; but in such regions as the islands of Indonesia and Papua New Guinea, or the South American or Central African rainforests, reports do come in from time to time of a previously unknown community and language.¹² For example, in June 1998 two such nomadic tribes (the Vahudate and the Aukedate, comprising 20 and 33 families, respectively) were found living near the Mamberamo River area, 2,400 miles east of Jakarta in Irian Jaya. This is a part of the world where the high mountains and deep valleys can easily hide a community, and

¹⁰ This is my count of Mosely and Asher (1994).

¹¹ Dixon (1997: 143) cites 5,000–6,000, as do Grenoble and Whaley (1998a), in their preface; Wardhaugh (1987: 1) cites 4,000–8,000, and settles on 5,000; Ruhlen (1987) goes for 5,000; Wurm (1991: 1) says ‘well over 5,000’; Krauss consulted a number of linguists in writing his article on ‘The world’s languages in crisis’ (1992: 5), and found widespread agreement that 6,000 was a reasonable estimate; Crystal (1997a: 287) also cites 6,000. Other major surveys are in progress: a ‘World Languages Report’, supported by UNESCO and Linguapax, and financed by the Basque Country, is scheduled for publication in 2001; see also the Global Language Register below.

¹² The world’s languages have a highly uneven distribution: c. 4% are in Europe; c. 15% in the Americas; c. 31% in Africa; c. 50% in Asia and the Pacific. The countries mentioned have the highest distributions: Papua New Guinea and Indonesia alone have 25% (1,529 languages) between them (according to the 1996 edition of *Ethnologue*).

it is likely that their speech will be sufficiently different from that of other groups to count as a new language. The social affairs office in the region in fact reports that its field officers encounter new groups almost every year.¹³

Even in parts of the world which have been explored, however, a proper linguistic survey may not have been carried out. As many as half the languages of the world are in this position. Of the 6,703 languages listed in the thirteenth edition of *Ethnologue*, 3,074 have the appended comment – ‘survey needed’. And what a survey chiefly does is determine whether the speakers found in a given region do indeed all use the same language, or whether there are differences between them. If the latter, it then tries to decide whether these differences amount only to dialect variations, or whether they are sufficiently great to justify assigning the speakers to different languages. Sometimes, a brief preliminary visit assigns everybody to a single language, and an in-depth follow-up survey shows that this was wrong, with several languages spoken. Sometimes, the opposite happens: the initial visit focuses on differences between speakers which turn out not to be so important. In the first case, the number of languages goes up; in the second case, it goes down. When decisions of this kind are being made all over the world, the effect on language counts can be quite marked.

To put some flesh on these statistics, let us take just one of those languages where it is said a survey is needed: Tapshin, according to *Ethnologue* also called Tapshinawa, Suru, and Myet, a language spoken by ‘a few’ in the Kadun district of Plateau State, Nigeria. It is said to be unclassified within the Benue-Congo broad grouping of languages. Roger Blench, of the Overseas Development Institute in London, visited the community in March 1998, and sent in a short report to the Foundation for Endangered Languages.¹⁴ He stressed the difficulty of reaching the settlement: Tapshin village is a widely dispersed settlement about 25 km north of the

¹³ The report is reproduced in *Ogmios* 9. 6. For similar discoveries in South America, see Adelaar (1998: 12); Kaufman (1994: 47) reports that about 40 languages have been discovered in South America during the past century. ¹⁴ Blench (1998).

Pankshin–Amper road, reached by a track which can be traversed only by a four-wheel drive, and which is often closed during the rainy season. The Tapshin people call themselves Ns'r, and from this derives Blench's name for them, Nsur, and presumably also the name Suru in *Ethnologue*; but they are called Dishili by the Ngas people (referred to as the Angas in *Ethnologue*). The name Myet derives from a settlement, Met, some distance west of Tapshin. The Tapshin people claim that the Met people speak 'the same' language as they do, but Blench is cautious about taking this information at face value (for such judgements may be no more than a reflection of some kind of social or historical relationship between the communities). No data seems previously to have been recorded on Nsur. From his initial wordlists, he concludes that there has been substantial mutual influence with the Ngas language. He estimates that there are some 3–4,000 speakers, though that total depends on whether Met is included along with Nsur or not.

This small example illustrates something of the problem facing the linguistic analyst. There is a confusion of names which must be sorted out, in addition to the observable similarities and differences between the speakers.¹⁵ The Nsur situation seems fairly manageable, with just a few alternatives to be considered. Often, the problem of names is much greater. Another Plateau State language, listed as *Berom* in *Ethnologue*, has 12 alternative names: *Birom*, *Berum*, *Gbang*, *Kibo*, *Kibbo*, *Kibbun*, *Kibyen*, *Aboro*, *Boro-Aboro*, *Afango*, *Chenberom*, and *Shosho*. The task then is to establish whether these are alternative names for the same entity, or whether they refer to different entities – the name of the people, the name of an individual speaker, or the name of the language as known by its speakers (a European analogy would be *Irish*, *Irishman/woman*, and *Gaelic/Irish/Erse*, respectively). Then there is the question of what the language is called by outsiders. There could of course be several 'outsider' names (*exonyms*), depending on how many other groups the language is in contact with (cf.

¹⁵ For a discussion of the problem of naming, with particular reference to China, see Bradley (1998: 56 ff.).

deutsch being equivalent to *allemand*, *German*, *Tedesco*, etc.), and these might range from friendly names through neutral names to offensive names (cf. ‘He speaks French’ vs ‘He speaks Frog’). Shosho, in the above list, is apparently an offensive name. But all this has to be discovered by the investigator. There is no way of knowing in advance how many or what kind of answers will be given to the question ‘What is the name of your language?’, or whether a list of names such as the above represents 1, 2, 6, or 12 languages. And the scale of this problem must be appreciated: the 6,703 language headings in the *Ethnologue* index generate as many as 39,304 different names.

Many of these names, of course, will refer to the dialects of a language. But this distinction raises a different type of difficulty: does a name refer to the whole of a language or to a dialect? The question of whether two speech systems should be considered as separate languages or as dialects of the same language has been a focus of discussion within linguistics for over a century. It is crucial to have criteria for deciding the question, as the decisions made can have major repercussions, when it comes to language counting. Take, for example, the Global Language Register (GLR), in the process of compilation by the Observatoire Linguistique:¹⁶ in a 1997 formulation by David Dalby, this project proposed a three-fold nomenclature – of *tongue* (or *outer language*), *language* (or *inner language* – or *idiom*, in a further proposal), and *dialect* – to avoid what it considered to be the oversimplified dichotomy of *language* and *dialect*. Early reports related to this project suggested that, using these criteria, an order of magnitude of 10,000 languages was to be expected – a surprisingly large total, when compared with the totals suggested above. The explanation is all to do with methodology. The GLR total is derived from the *tongues* and *idioms* of their system, and includes as languages many varieties which other approaches would consider to be dialects. One

¹⁶ The following details are taken from a Logosphere Workshop held at the School of Oriental and African Studies, London, September 1997, specifically from Dalby (1997), and his follow-up paper subsequently circulated.

example will illustrate the ‘inflationary’ effect of this approach. The orthodox approach to modern Welsh is to consider it as a single language, with the notable differences between (in particular) north and south Welsh referred to as dialects. On grounds of mutual intelligibility and sociolinguistic identity (of Wales as a nation-principality), this approach seems plausible. The GLR analysis, however, treats the differences between north and south Welsh as justifying the recognition of different languages (each with their own dialects), and makes further distinctions between Old Welsh, Book Welsh, Bible Welsh, Literary Welsh, Modern Standard Welsh, and Learners’ Normalized Welsh (a pedagogical model of the 1960s known as ‘Cymraeg Byw’). Excluding Old Welsh, in their terms a total of six ‘inner languages’ can be recognized within the ‘outer language’ known as modern Welsh. One can see immediately how, when similar cases are taken into account around the world, an overall figure of 10,000 could be achieved.

The language/dialect issue has been addressed so many times, in the linguistics literature, that it would be gratuitous to treat it in any detail here.¹⁷ In brief, on purely linguistic grounds, two speech systems are considered to be dialects of the same language if they are (predominantly) mutually intelligible. This makes *Cockney* and *Scouse* dialects of English, and *Quechua* a cover-name for over a dozen languages. On the other hand, purely linguistic considerations can be ‘outranked’ by sociopolitical criteria, so that we often encounter speech systems which are mutually intelligible, but which have nonetheless been designated as separate languages. A well-recognized example is the status of Swedish, Danish, and Norwegian, which are counted as separate languages despite the fact that the members of these communities can understand each other to an appreciable extent. A more recent example is *Serbo-Croatian*, formerly widely used as a language name to encompass a set of varieties used within former Yugoslavia, but following the

¹⁷ Standard accounts are to be found in Chambers and Trudgill (1980: ch. 1) and Crystal (1997a: ch. 47).

civil wars of the 1990s now largely replaced by the names *Serbian*, *Croatian*, and *Bosnian*. In 1990 there was a single language spoken in these countries; now there are three. The linguistic features involved have changed hardly at all; but the sociopolitical situation has changed irreversibly.

It is of course likely that the linguistic differences between these languages will increase, as their respective communities strive to maximize them as symbols of local identity. This process is already happening. If it continues, then one day it is conceivable that Serbian and Croatian could become mutually unintelligible – a further example of something that has happened repeatedly and normally in linguistic evolution. Indeed, it is possible that a significant increase in the world's languages may one day emerge as an evolutionary consequence of the contemporary trend to recognize ethnic identities. Even global languages could be affected in this way. The point has been noted most often in relation to English, where new varieties have begun to appear around the world, as a consequence of that language's emerging status as a world lingua franca. Although at present Singaporean, Ghanaian, Caribbean, and other 'New Englishes' continue to be seen as 'varieties of English', it is certainly possible for local sociopolitical movements to emerge which would 'upgrade' them to language status in due course. Books and articles are already appearing which (in their nomenclature, at least) anticipate such outcomes.¹⁸ After all, if a community wished its way of speaking to be considered a 'language', and if they had the political power to support their decision, who would be able to stop them doing so? The present-day ethos is to allow communities to deal with their own internal policies themselves, as long as these are not perceived as being a threat to others. The scenario for the future of English is so complex and unpredictable, with many pidgins, creoles, and mixed varieties emerging and gradually acquiring prestige, that it is perfectly possible that in a few generations time the degree of local distinctiveness in a speech

¹⁸ McArthur (1998), Rosen (1994), and the journal *World Englishes*. See also Crystal (1998).

system, and the extent of its mutual unintelligibility with other historically related systems, will have developed to the extent that it will be given a name other than 'English' (as has happened already – though not yet with much success – in the case of Ebonics). At such a time, a real evolutionary increase in the number of 'English languages' would have taken place. A similar development could affect any language that has an international presence, and where situations of contact with other languages are fostering increased structural diversity. The number of new pidgins and creoles is likely to be relatively small, compared with the rate of language loss, but they must not be discounted, as they provide evidence of fresh linguistic life.

Estimates about the number of languages in the world, therefore, must be treated with caution. There is unlikely to be any single, universally agreed total. As a result, it is always problematic translating observations about percentages of endangered languages into absolute figures, or vice versa. If you believe that 'half the languages in the world are dying', and you take one of the middle-of-the-road totals above, your estimate will be some 3,000 languages. But if you then take this figure out of the air (as I have seen some newspaper reporters do), and relate it to one of the higher estimates (such as the Global Language Register's 10,000), you would conclude that less than a third of the world's languages are dying – and, as a consequence, that the situation is not as serious as has been suggested. The fact that this reasoning is illegitimate – the criteria underlying the first total being very different from those underlying the second – is disregarded. And, as I read the popular press, I see all kinds of claims and counter-claims being made, with the statistics used to hold a weight of argument they cannot bear.

At the same time, despite the difficulties, we cannot ignore the need for global measures. As so much of the situation to be described below is bound up with matters of national and international policy and planning, we have to arrive at the best estimates we can, in order to persuade governments and funding bodies about the urgency of the need. Accordingly, I will opt for the range

of 5,000–7,000 as my lower and upper bounds, for the year 2000 – $6 \pm 1K$ – and will relate any further talk of percentages to this.¹⁹

The size of the problem

A language is said to be dead when no one speaks it any more. It may continue to have existence in a recorded form, of course – traditionally in writing, more recently as part of a sound or video archive (and it does in a sense ‘live on’ in this way) – but unless it has fluent speakers one would not talk of it as a ‘living language’. And as speakers cannot demonstrate their fluency if they have no one to talk to, a language is effectively dead when there is only one speaker left, with no member of the younger generation interested in learning it. But what do we say if there are two speakers left, or 20, or 200? How many speakers guarantee life for a language?

It is surprisingly difficult to answer this question. One thing is plain: an absolute population total makes no sense. The analysis of individual cultural situations has shown that population figures without context are useless. In some circumstances, such as an isolated rural setting, 500 speakers could permit a reasonably optimistic prediction; in others, such as a minority community scattered about the fringes of a rapidly growing city, the chances of 500 people keeping their ethnic language alive are minimal. In many Pacific island territories, a community of 500 would be considered

¹⁹ As an endnote to this section, it is worth remembering that the languages we have today are only a fraction of all the languages there have ever been. There are too many unknowns for estimates to be other than highly speculative, but we can make some guesses using two criteria. First, we have some evidence from the known span of recorded Western history about the number of languages (and civilizations) that have died; and from historical linguistics we know something about the rate at which languages change – for example, the rise of the Romance languages from Vulgar Latin. We also have a vague idea about the age of the language faculty in humans, which probably arose between 100,000 and 20,000 years ago. Combining these variables is a daring task, but some people have attempted it. Pagel (1995: 6) concludes that there may have been as many as 600,000 languages spoken on earth, or as few as 31,000; his ‘middle of the road’ estimate is 140,000. Even if we take his lowest estimate, it is plain that far more languages have died, in the history of humankind, than now remain. For the question of whether the rate of decline has increased in recent times, see below; for the issue of what we may have lost, see chapter 2.

quite large and stable; in most parts of Europe, 500 would be minuscule. Speaker figures should never be seen in isolation, but always viewed in relation to the community to which they relate. Thus, in one survey, by Akira Yamamoto,²⁰ languages which had between 300 and 500 speakers included the Santa Ana dialect of Keresan (USA), Ulwa (Nicaragua), and Sahaptin (USA); but the first of these localities had a community population of only 600, the second had about 2,000, and the third had about 12,000. Plainly, the figure 500 tells a different story in each case, when it comes to evaluating the level of endangerment. Yamamoto concludes his survey with the comment that population size alone is not an accurate indicator of a language situation. He gives an example of a language which at the time of the survey had just 185 speakers of all ages – Karitiana (Brazil). Though this seems small, he points out that the total size of the community was only 191 – in other words, we have to say that over 96% of the people speak the language. And as the children are apparently continuing to learn Karitiana as their first language (with Portuguese coming later, as a second language), Yamamoto asks pertinently, is this really an endangered language?

The presumption is that any language which has a very small number of speakers is bound to be in trouble, and common sense tells us that this should usually be the case.²¹ Perhaps only in places where the circumstances are especially favourable could such a language survive (see, further, chapter 3). So, notwithstanding the exceptions, most people would accept that a language spoken by less than 100 is in a very dangerous situation. They would then probably think in terms of a 'sliding scale' whereby languages with less than 500 would be somewhat less endangered, those with 1,000 even less so, and so on. What is unclear is the level at which we would stop automatically thinking in terms of danger. The figures

²⁰ Yamamoto (1997: 12).

²¹ Many articles on endangered languages reflect this point: for example, Norris (1998: 3) says: 'There are a number of factors which contribute to a language's ability to survive. First and foremost is the size of the population with an Aboriginal mother tongue or home language. Since a large base of speakers is essential to ensure long-term viability, the more speakers a language has, the better its chances of survival.' See, further, chapter 4.

suggested for this level are higher than we might expect. A total of 10,000 suggests safety in the short term, but not in the medium term.²² In the savannah zone in Africa, for example, some linguists consider a language to be endangered if it has less than 20,000 speakers.²³ And in parts of West Africa, where English and French creoles in particular are attracting huge numbers of new speakers, many local languages are felt to be endangered – even though they are currently spoken by several hundred thousand. This is what surprises people – that languages with such large numbers of speakers can nonetheless be in danger. Yet, within the twentieth century, we have seen many languages fall from very large numbers: for example, in 1905 one estimate of Breton gave 1.4 million speakers; today, depending on the kind of fluency criteria used, the figure may be as low as 250,000.²⁴ And when we consider the causes of language death (chapter 3), it is evident that the factors involved are so massive in their effect that even a language with millions of speakers may not be safe. Even Yoruba, with 20 million speakers, has been called ‘deprived’ because of the way it has come to be dominated by English in higher education.²⁵ And during a visit to Southern Africa in 1998, speakers of several of the newly recognized official languages of South Africa expressed to me their anxiety for their long-term future, in the face of English – including several Afrikaners (whose language, Afrikaans, is spoken by around 6 million). The same reaction was observed in Zimbabwe, where not only speakers of Ndebele (1.1 million) but even of Shona (7 million) professed the same anxiety. One experience illustrates the trend that these people find so worrying: engaging a Johannesburg driver in conversation, it transpired that he was conversant with all 11 of his country’s official languages – an ability which he did not think at all unusual. However, his main ambition was to earn enough to enable all his children to learn English. None of the other languages ranked highly in his esteem.

Although concerns have been expressed about some languages

²² For example, Dixon (1991: 231).

²³ Footnote to a field report on Kagoro (Mali) by Vydrine (1998: 3).

²⁴ Total given for 1991 in the Breton entry in Price (1998: 38). ²⁵ Brenzinger (1998: 93).

with relatively large populations, it is the ones with the smallest totals which have inevitably captured the most attention. Yamamoto also recognizes this (see fn. 20 above): ‘the number of speakers is an immediate index for its endangered situation’. It is difficult to see how a community can maintain its identity when its population falls beneath a certain level. Hence there is some force behind the statistics of language use which scholars have been compiling in recent years – though these surveys have not been taking place long enough for one to see long-term trends (e.g. whether there is an increase in the rate at which languages are being lost). An updated table in *Ethnologue* (February 1999) recognizes 6,784 languages, with data available for 6,059. Using this latter figure – and inevitably disregarding the question-marks which accompany several of the estimates – we can obtain the totals in Table 1, all for first language speakers.

There are many observations which can be made from a scrutiny of a summary table of this kind, and of the fuller table which underlies it. Beginning with the largest totals: it is evident that a very small number of languages account for a vast proportion of the world’s population (thought to have passed 6 billion in mid 1999). The 8 languages over 100 million (Mandarin, Spanish, English, Bengali, Hindi, Portuguese, Russian, Japanese) have nearly 2.4 billion speakers between them; and if we extend this count to include just the top 20 languages, we find a total of 3.2 billion – over half the world’s population. If we continued the analysis downwards, we would eventually find that just 4% of the world’s languages are spoken by 96% of the population.

Turning this statistic on its head: 96% of the world’s languages are spoken by just 4% of the population. That is the perspective within which any discussion of language death must be seen. And, at the bottom end of the table, there are some sobering deductions. From the rightmost column, we can see that a quarter of the world’s languages are spoken by less than 1,000 people; and well over half by less than 10,000. The median number of speakers for all languages in the list is 6,000. If the figure of 20,000 (referred to above as a danger-level in some parts of the world) were taken as a universal datum, this would correspond to exactly two-thirds of

Table 1

	<i>N</i>	%	Cumulative downwards %	Cumulative upwards %
more than 100 million	8	0.13		99.9
10–99.9 million	72	1.2	1.3	99.8
1–9.9 million	239	3.9	5.2	98.6
100,000–999,999	795	13.1	18.3	94.7
10,000–99,999	1,605	26.5	44.8	81.6
1,000–9,999	1,782	29.4	74.2	55.1
100–999	1,075	17.7	91.9	25.7
10–99	302	5.0	96.9	8.0
1–9	181	3.0	99.9	

the world's languages. Then, using the leftmost column, we can see that nearly 500 languages have less than 100 speakers; around 1,500 have less than 1,000; and 3,340 have less than 10,000. If a population of 20,000 is again taken as a danger-level datum, we are talking about 4,000 languages. Most of these will be found in those parts of the world where languages are most numerous – notably in the equatorial regions everywhere (see fn. 12 above). The underlying table also lists 51 languages with just a single speaker – 28 in Australia, 8 in the USA, 3 in South America, 3 in Africa, 6 in Asia, 3 in the Pacific islands.

As we have already seen, conditions vary so much around the world that it is impossible to generalize from population alone about the rate at which languages die out. That is why there is so much variation in the claims that are currently being made, that 'x% of the world's languages are going to die out in the next 100 years' – x here has been anything from 25% (a conservative estimate which correlates with the 'less than 100' criterion) to 80% or more (a radical estimate which correlates with the 'less than 100,000' criterion). It is impossible, in our present state of knowledge, to say more about these deductions other than that they are well-informed guesswork. Most available demographic data (on death-rate, fertility-rate, etc.) is country-based, and not language-related. On the other hand, there have been enough micro-studies

of specific locations carried out over a period of time to indicate the rate at which a downward trend operates. One report, on Dyirbal (Australia), found some 100 speakers in 1963, with everyone over about 35 speaking it as a first language; by 1993, there were just 6 speakers, all over about 65, with comprehension by some younger people.²⁶ Another report showed that in 1990 there were 60 fluent speakers of Aleut in Atka (USA), the main village where it survives; but by 1994 this number was down to 44, with the youngest speakers in their twenties.²⁷ At that rate of attrition, the language could stop being used by 2010.²⁸ (The factors which can influence the rate of decline are reviewed in chapter 3.)

Here is a more detailed example of the nature of a downwards trend. A Canadian census-based study²⁹ showed that between 1981 and 1996 most of Canada's 50 Aboriginal languages suffered a steady erosion; indeed, by the latter date only 3 of the languages were felt to have large enough populations to be secure from the threat of long-term extinction (Inuktitut, Cree, Ojibway). A superficial look at the census data might suggest the contrary, for in this 15-year period the number of people reporting an indigenous mother-tongue actually increased by 24% (chiefly the result of high fertility rates among the population). However, a closer look at the statistics shows a very different picture. There are four critical points (to each of which I add a general observation).

- The number of people who spoke an indigenous language at home grew by only 6%. In real terms, for every 100 people with an indigenous mother-tongue, the number whose home

²⁶ Dixon (1997: 105).

²⁷ Bergsland (1998: 38). Another example of a language which has gone from vital to moribund within a generation is Cup'ik in Chevak, Alaska: see Woodbury (1998: 239). The suddenness of the change in the languages of the Great Plains is emphasized in Furbee, Stanley, and Arkeketa (1998: 75).

²⁸ Another example of extrapolation is given for Tlingit and Haida in Dauenhauer and Dauenhauer (1998: 72): on the basis of current trends, if the youngest speaker of Tlingit is 45, and lives to be 100, the language will be dead in 2050. It should be noted that a pattern of decline is not always a smooth descending curve. There is evidence of a cyclical process in some places, as a period of loss is followed by one of maintenance. In parts of India, for example, there is evidence of people letting their indigenous language fall into disuse in early childhood, or after moving to a city to find work; but if they join new social networks after marriage, or return to their village with a newfound political awareness, they may then become actively involved in its resuscitation (Annamalai 1998: 25).

²⁹ Norris (1998).

language was most often an indigenous language declined from 76 to 65. (The importance of using the language at home is critical, in parts of the world where a population lives in relative isolation, and where it is unlikely that numbers will be enhanced through immigration. In the present survey, the viability of a language is directly reflected in its proportion of home language use: in the more viable languages, an average of 70 out of every 100 used their indigenous language at home; in the less viable ones, this had fallen to 30 or fewer.)³⁰

- The age trend shows a steady decline: 60% of those aged 85+ used an indigenous mother-tongue, compared with 30% of those aged 40–44, and 20% of children under 5. The average age of speakers of all indigenous languages rose from 28 to 31. (Age is another critical factor, as it shows the extent to which language transmission between generations has been successful. The lower the average language population age, the more successful the parents have been in getting young people to speak it. A rise in average speaker age is a strong predictor of a language's progress towards extinction.)
- The points at which language loss chiefly take place can also be identified: in 1981, 91 out of 100 children under 5 spoke their mother-tongue at home; in 1996, these children had reached their late teens, and only 76 out of 100 now did so. (The ages at which there is a shift in language use are highly significant.³¹ The dependence of very young children on their family means that few have an opportunity to shift from their

³⁰ Some demographers use an *index of continuity*, derived by dividing the number of people who speak an indigenous language at home by the number of those who speak it as a mother-tongue. A figure of less than 100 indicates a decline in the viability of the language. Another measure is an *index of ability*, derived by dividing the number of mother-tongue users by the number of people who have reasonable conversational ability in it. A figure of more than 100 indicates the presence of second-language speakers, and thus the possibility of revival. See Harrison (1997).

³¹ *Language shift* is the conventional term for the gradual or sudden move from the use of one language to another (either by an individual or by a group). Other terms frequently encountered in the endangered languages literature include: *language loss*, for a situation where a person or group is no longer able to use a language previously spoken; *language maintenance*, where people continue to use a language, often through adopting specific measures; and *language loyalty*, which expresses the concern to preserve a language when a threat is perceived.

home language. By contrast, the teenage years, characterized by pressure both from peer-group trends and from the demands of the job-market, are a particularly sensitive index of where a language is going.)

- The preceding point takes on fresh significance when people leave the family home. The data show that language loss is most pronounced during the early years of entering the job-market and after marriage (especially among women): between ages 20 and 24, 74 out of 100 women were using an indigenous language; but in the corresponding group 15 years later, this average had fallen to 45. (Such a shift is particularly serious, as these are the years in which women are likely to be bringing up their children. Fewer children are thus going to be exposed to the indigenous language at home.)

There are also several positive signs in the Canadian situation; but the picture of overall decline is very clear, and has its parallels in other census studies, notably in the USA. These studies, however, provide only a very partial picture of the world situation: most countries do not record census data on language use at all, or (when they do) the questions they ask do not throw light on the issue of language endangerment.

It is certainly possible, after immersing yourself in data of this kind, to ‘take a view’ (as lawyers say) about the global situation, and several writers have done so. One of the most widely quoted statistics is that of Michael Krauss, who concludes, after a statistical review:³²

I consider it a plausible calculation that – at the rate things are going – the coming century will see either the death or the doom of 90% of mankind’s languages.

That means only about 600 are ‘safe’. As I have already indicated in my Preface, the groups which have been established to monitor the situation are in total agreement about the seriousness of the situation, though usually avoiding a hard statistic. For example,

³² Krauss (1992: 7).

here are two judgements from the Foundation for Endangered Languages:³³

The majority of the world's languages are vulnerable not just to decline but to extinction.

Over half the world's languages are moribund, i.e. not effectively being passed on to the next generation [see further below].

A middle position would assert 50% loss in the next 100 years. This is the view independently arrived at by three linguists reported by Krauss in 1992.³⁴ 50% is 3,000 languages. 100 years is 1,200 months. To meet that time frame, at least one language must die, on average, every two weeks or so. This cannot be very far from the truth.

Levels of danger

Comparing levels of endangerment is very difficult, in view of the diversity of language situations around the world, and the lack of theoretical models which would allow us to interpret combinations of relevant variables. How should we approach the kind of question raised earlier: which is the more endangered – a language where 400 people out of a community of 500 speak it, or one which has 800 speakers out of 1,000? Plainly, in such cases, the only answer is 'It all depends' – on such factors as the rate of acquisition by the children, the attitude of the whole community to it, and the level of impact of other languages which may be threatening it. At the same time, it is important for people to be able to take such factors into account (intuitively, at least, if surveys have not been made) and arrive at a judgement about just how endangered a language is. Some sort of classification of endangerment needs to be made. Without it, it would be impossible to 'take a view' about the urgency of the need, and thus to allocate scarce resources, in cases where something might be done (chapter 5).

³³ The first is from the preamble to the proposal to establish the Foundation for Endangered Languages, June 1995; the second is from *Iatiku* 2. 3. ³⁴ Krauss (1992: 6).

A common-sense classification recognizes three levels: languages are **safe**, **endangered**, or **extinct**. To this, Michael Krauss adds a notion which has been widely taken up: languages which are no longer being learned as a mother tongue by children are said to be **moribund** (a term originating in the field of medicine).³⁵ This captures the notion of a language well beyond the stage of ‘mere’ endangerment, because it lacks intergenerational transmission; the analogy is with a species unable to reproduce itself. The distinction is illustrated by Krauss with reference to North America, where he identifies a total of 187 indigenous languages. All are, in principle (given the dominant English-language environment), endangered; but major efforts are taking place in some communities to reverse the decline (see chapter 5). The more important statistic is to identify those which are moribund – which Krauss calculates to be 149, or 80%. In Alaska, the percentage is higher: there, only 2 out of the 20 indigenous languages were, in 1992, still being learned by children. A similar percentage is found in Australia. On the other hand, applying his criterion in South America produces a lower figure (27%) and in Central America an even lower one (17%).

Some classifications go a stage further, distinguishing ‘safe’ and ‘not so safe’, as in this five-level system:³⁶

viable languages: have population bases that are sufficiently large and thriving to mean that no threat to long-term survival is likely;

viable but small languages: have more than *c.* 1,000 speakers, and are spoken in communities that are isolated or with a strong internal organization, and aware of the way their language is a marker of identity;

endangered languages: are spoken by enough people to make survival a possibility, but only in favourable circumstances and with a growth in community support;

nearly extinct languages: are thought to be beyond the possibility of survival, usually because they are spoken by just a few elderly people;

³⁵ Krauss (1992: 4). ³⁶ Kincade (1991: 160–3).