Citation for published version

Loh, Jonathan (2017) Indicators of the Status of and Trends in Global Biological, Linguistic and Biocultural Diversity. Doctor of Philosophy (PhD) thesis, University of Kent,.

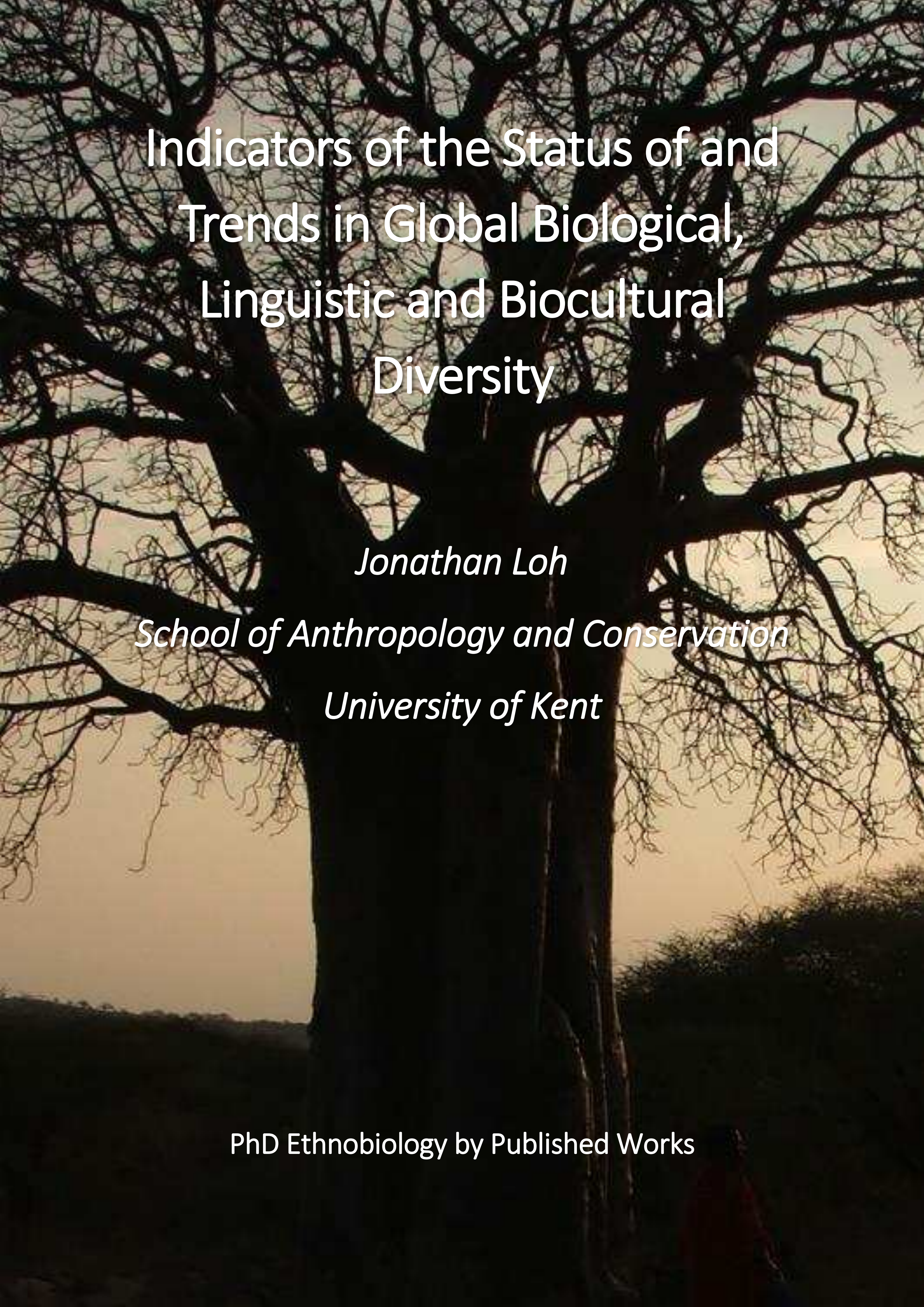
DOI

Link to record in KAR

https://kar.kent.ac.uk/61424/

Document Version

UNSPECIFIED



Indicators of the Status of and Trends in Global Biological, Linguistic and Biocultural Diversity

Jonathan Loh

School of Anthropology and Conservation

University of Kent

PhD Ethnobiology by Published Works

Contents

Summary and Supporting Statement

Jonathan Loh

PhD by Published Work

University of Kent

School of Anthropology and Conservation

Rev. 31 March 2017

Contents

Abstract

Biodiversity is in global decline and around 19% of the world's vertebrate species are listed as threatened on the IUCN Red List (Baillie et al. 2010; IUCN 2013). Linguistic diversity is also in decline and it is believed that as many as 90% of the world's 7,000 languages are threatened with extinction this century (Krauss 1992; Nettle and Romaine 2000). It has also been noted that there is a strong similarity in the distributions of terrestrial species diversity and linguistic diversity at the global scale, with the greatest richness found in the humid tropics and the lowest richness in the cold temperate zones (Mace and Pagel 1995; Sutherland 2003; Gavin et al. 2013). The term biocultural diversity has come into use to describe the collective diversity of species, languages and cultures around the world and their ongoing declines (Maffi 2001b; Harmon 2002). One of the papers presented here develops the first national index of biocultural diversity, which confirms the pattern of greatest richness in the tropics, particularly in Southeast Asia (Loh and Harmon 2005).

However, measures of the state of biological, linguistic and biocultural diversity based on richness alone simply record the number of species or languages present and ignore underlying trends in abundance or populations of species or speakers of languages. Extinction risk has been the most widely-used measure of the status of both species and languages, but indicators based on time-series population data offer an alternative and more responsive measure of status and trends. The other papers presented here describe the development of Living Planet Index (Loh et al. 2005; Collen et al. 2009), an indicator which aggregates trends in populations of several thousand vertebrate species worldwide and shows an overall decline of about 30% over four decades since 1970, and the Index of Linguistic Diversity (Harmon and Loh 2010; Loh and Harmon 2014), a closely-related indicator based on trends in speaker numbers of around a thousand languages worldwide, and which also shows a decline of about 30% over the same period. At the regional level, the respective trends diverge. For biodiversity, there was a greater rate of decline in the tropics compared with temperate regions, whereas for linguistic diversity, there was a far higher rate of decline in the Americas, Australia and the Pacific compared with Africa, Asia and Europe. An analysis of the threat status of 1,500 languages using the IUCN Red List criteria reveals that 27% languages are threatened with extinction and confirms the regional pattern in the status of languages apparent in the Index of Linguistic Diversity. The differing regional patterns between the declines in languages and species reflect differences in the proximate drivers of diversity loss, where habitat loss or degradation are the major causes of species population declines (Millennium Ecosystem Assessment 2005), while linguistic diversity is lost primarily through language shift, a process whereby a politically, socially or economically dominant language displaces local or indigenous languages either as a result of colonialization, industrialization or migration (Nettle 1999).

Abbreviations

CAFF	Conservation of Arctic Flora and Fauna
CBCD	Centre for Biocultural Diversity
CBD	Convention on Biological Diversity
CBMP	Circumpolar Biodiversity Monitoring Program
COP	Conference of the Parties
CPI	Consumer Prices Index
EBV	Essential Biodiversity Variable
ELF	Ethnolinguistic Fractionalization
GBO	Global Biodiversity Outlook
GDP	Gross Domestic Product
GEO	Global Environment Outlook
GEOBON	Group on Earth Observations - Biodiversity Observation Network
IBCD	Index of Biocultural Diversity
ILD	Index of Linguistic Diversity
IPBES	Intergovernmental Platform on Biodiversity and Ecosystem Services
ISE	International Society for Ethnobiology
IUCN	International Union for Conservation of Nature
IVIC	Instituto Venezolano de Investigaciones Cientificas
LPI	Living Planet Index
LPR	Living Planet Report
MUIENR	Makerere University Institute of Environment and Natural Resources
NDBD	National Biodiversity Data Bank
NINA	Norwegian Institute for Nature Research
OED	Oxford English Dictionary
RL	Red List
RLI	Red List Index
RPI	Retail Prices Index
RSPB	Royal Society for the Protection of Birds
SIL	Summer Institute of Linguistics
SSC	Species Survival Commission
TCF	The Christensen Fund
TdV	Tour du Valat
UNDP	United Nations Development Programme
UNEP	United Nations Environment Programme
UNESCO	United Nations Educational Scientific and Cultural Organization
VITEK	Vitality Index of Traditional Environmental Knowledge
WBT	Wycliffe Bible Translators
WCMC	World Conservation Monitoring Centre
WCS	Wildlife Conservation Society
WET	Wetland Extent Trends
WWF	Worldwide Fund for Nature (World Wildlife Fund in US and Canada)
ZSL	Zoological Society of London

The Submitted Publications and Author Contributions

Five publications, dating from 2005 to 2014, are submitted for this PhD by Published Work. Three of the publications were co-authored with David Harmon, executive director of the George Wright Society, a professional association of people who work in parks, protected areas and cultural sites. The other papers were co-authored with colleagues from conservation organizations including WWF, UNEP-WCMC, RSPB, TdV and ZSL. The relative contributions of the authors are listed below. In chronological order, the published works are:

1. Jonathan Loh, Rhys E Green, Taylor Ricketts, John Lamoreux, Martin Jenkins, Valerie Kapos and Jorgen Randers (2005). The Living Planet Index: using species population time series to track trends in biodiversity. *Phil. Trans. Royal Society B*, 360, 289-295.

   Loh conceived the Living Planet Index. Jenkins, Kapos, Loh and staff at WCMC collected the data. Loh and Jenkins carried out preliminary data processing. Loh conducted the data analysis using the chain method. Green carried out the data analysis using the linear modelling method. Lamoreux collected data on species distributions for the representativeness analysis and Ricketts carried out the representativeness analysis. Loh and Green wrote the text with input from the other authors. This paper was presented at a Royal Society discussion meeting entitled *Beyond extinction rates: monitoring wild nature for the 2010 target*, 19-20 July 2004. The Living Planet Index was originally devised for the WWF Living Planet Report in 1998 (Loh et al. 1998) and developed over the subsequent years (Loh et al. 1999; Loh 2000; Loh 2002; Loh and Wackernagel 2004). Until 2004, the LPI in the Living Planet Report employed the chain method for calculating the index, devised and calculated by Loh in collaboration with Jenkins. Randers came up with the original idea to create an index and provided comments on the manuscript.

2. Jonathan Loh and David Harmon (2005). A global index of biocultural diversity. *Ecological Indicators*, 5, 231-241.

   Loh and Harmon collected the data. Loh devised the methods and conducted the analyses. Loh and Harmon co-wrote the text. The paper was based in part on a report by Harmon and Loh produced for Terralingua and presented at the 9th International Convention on Ethnobiology at the University of Kent, Canterbury, in 2004 (Harmon and Loh 2004a).

3. Ben Collen, Jonathan Loh, Sarah Whitmee, Louise McRae, Rajan Amin and Jonathan E M Baillie (2009). Monitoring change in vertebrate abundance: the Living Planet Index. *Conservation Biology*, 23, 317-327.

   Collen and Loh wrote the text with input from the other authors. McRae, Whitmee and staff at ZSL collected the data and entered the data into a database. Loh developed the revised chain method. Collen and Amin developed the generalized additive modelling method and wrote software to carry out the calculations. Collen carried out the representativeness analysis. Until 2006, the Living Planet Index was calculated using the chain method (Loh and Goldfinger 2006), and is described in the earlier paper by Loh et al. (2005); this paper documents significant changes and improvements made in the method for calculating the LPI.

4. David Harmon and Jonathan Loh (2010). The Index of Linguistic Diversity: a new quantitative measure of trends in the status of the world's languages. *Language Documentation and Conservation*, 4, 97-151.

   The authors contributed equally to this work. Harmon collected the data and entered the data into a database. Loh developed the method, processed and analysed the data. Harmon and Loh co-wrote the text of the paper. Harmon wrote the text of Appendix A. Loh wrote Appendix B. The paper was based on a study carried out for Terralingua, funded by the Christensen Fund and presented at the 12th International Congress on Ethnobiology in Tofino, British Columbia, in May 2010. Preliminary methods and findings were presented at a conference at the American Museum of Natural History entitled *Sustaining Cultural and Biological Diversity in a Rapidly Changing World*, 2-5 April 2008, some of which was later published in a book chapter (Harmon, Woodley and Loh 2010), and at the IUCN World Conservation Congress in Barcelona, September 2008.

5. Jonathan Loh and David Harmon (Loh and Harmon 2014). *Biocultural Diversity: threatened species, endangered languages*. WWF Netherlands, Zeist, the Netherlands. 60 pages.

   Loh and Harmon collected the data. Loh processed the data, developed the methods and carried out the analyses. Loh wrote the text with input from Harmon. Results were presented at the 36th International LAUD Symposium *Endangerment of Languages across the Planet: the Dynamics of Linguistic Diversity and Globalization* at the University of Koblenz-Landau, Germany, 31 March – 3 April 2014, and at *Studium Generale* at the University of Utrecht, 3 February 2015.

Introduction

The publications presented here are concerned with indicators of global biological, linguistic and biocultural diversity. An indicator is a statistic which summarizes some aspect of the state of a system in a single variable. GDP, for example, was established as a measure of the productivity of an economy (Dickenson 2011), and growth in GDP is the most widely-used indicator of the economic development (Jackson 2011). An indicator does not contain information about every aspect of the state of an entire system. Average life expectancy at birth is used as an indicator of the overall health of a country's population (UNDP 2015). Although there is more to a nation's health than life expectancy, it is nevertheless a useful indicator for comparing countries at a glance or tracking trends in a country's development.

There is sometimes confusion about the difference between an indicator and an index. An index is a metric which does not have units. Many indicators are expressed as an index. The CPI is an index which tracks the prices of a selection of goods. Some prices may go up and others may go down, but the CPI is designed to reflect prices in general as experienced by a typical consumer. The index is not expressed in currency units, but given an arbitrary value in a certain year (100 in 2005). Inflation is expressed as the CPI's annual percentage change (Office for National Statistics 2015).

The submitted papers are about the development of three indicators, all of which are expressed as indices. The papers describe exactly what these indicators measure and how the indices are constructed mathematically. But first it is helpful to be clear about the broad global entities that they are trying to measure: what it is that they are indicators of. Biodiversity is a term that can mean different things to different people, and linguistic diversity and biocultural diversity are no more precise.

What is biodiversity?

Biodiversity means different things depending on its context. The OED entry for biodiversity reads simply: "Diversity of plant and animal life, esp. as represented by the number of extant species" (OED Online 2016a). Kevin Gaston (1996b) describes three different meanings of the term biodiversity: one generic, one scientific and one political. The first is a concept synonymous with the "variety of life" or "life on Earth". This concept is extremely broad, and biologists generally recognize diversity at three levels: genetic, species or organism, and community or ecosystem (Groombridge 1992; Harper and Hawksworth 1994). The most widely used definition of biodiversity, which comes from the UN Convention on Biological Diversity, captures this breadth and is interpreted to refer explicitly to those three levels:

> "*Biological diversity*" means the variability among living organisms from all sources including, *inter alia*, terrestrial, marine and other aquatic ecosystems and the ecological complexes of which they are part; this includes diversity within species, between species and of ecosystems. (CBD Secretariat 1994)

Gaston's second, more scientific meaning of the term refers to biodiversity as a measurable entity. Biodiversity as a generic concept does not represent a single coherent unity which can be measured by a single variable (Gaston 1996b). Any measure of biodiversity, therefore, can only quantify some aspect of biodiversity and not its totality in the sense of "Life on Earth". Anne Magurran (2010) highlights this point:

> In practice, to assess how much diversity we have and what it does, we need to be more specific about the aspect of biodiversity we are concerned with, and the area and time frame over which we want to measure it. For example, we might consider the types and relative abundances of species of trees in a forest, or the genetic diversity associated with the individuals of those species, or even how the number and composition of forests across a biogeographic region have changed over the past century.

In *Measuring Biological Diversity*, Magurran (2004) focuses on biodiversity as "the variety and abundance of species in a defined unit of study", making species "the common currency of diversity". Gaston (1996a) similarly argues that the species can be viewed as "…the fundamental unit of biodiversity, species richness as the fundamental meaning of biodiversity, and the high level of species extinction as the main manifestation of the biodiversity crisis."

Because species richness is the most readily measurable aspect of biodiversity, it is sometimes treated as the only measure of biodiversity, and species extinction and extinction risk are widely used indicators of biodiversity loss (Smith et al. 1993; IUCN 1996; IUCN and WCMC 1998; Butchart et al. 2004; Butchart et al. 2005; IUCN 2013). While species are generally discrete entities which are easier to count and quantify than genes or ecosystems, those lower and higher levels of biodiversity are not captured by species richness.

Gaston's third meaning of biodiversity as a social or political construct (Gaston 1996b) describes not a neutral scientific concept, but a value-laden one. Biodiversity was originally coined and gained popularity as a term for a public good which should be conserved. Oxford Dictionaries online gives a definition of biodiversity which captures this sense: "The variety of plant and animal life in the world or in a particular habitat, a high level of which is usually considered to be important and desirable" (Oxford Dictionaries n.d.).

Indeed, the CBD definition of biodiversity in the broad sense is also a political construct, not only because it has been accepted by 196 governments, but because the very purpose of the CBD is "…the conservation of biological diversity, the sustainable use of its components and the fair and equitable sharing of the benefits arising out of the utilization of genetic resources…" (CBD Secretariat 1994).

If conserving biodiversity in its broadest sense, encompassing genetic, species and ecosystem diversity, is considered important and desirable, then it raises the question of how this can be measured. How, for example, can we know whether the objective of the CBD has been achieved? This is a relevant question here because two of the indicators described in the submitted publications have been adopted by the CBD and developed in the context of its 2010 and 2020 targets (Biodiversity Indicators Partnership 2013a; Biodiversity Indicators Partnership 2013b; CBD Secretariat 2014; UNEP-WCMC 2016a; UNEP-WCMC 2016b; UNEP-WCMC 2016c).

What is Linguistic Diversity?

One of the themes running through this work is the analogy between species and languages, which share a number of similarities. Darwin observed that languages, like species, lend themselves to a hierarchical system of classification in "groups within groups" as a result of evolutionary descent from a common ancestor (Darwin 1871). If a language is analogous to a species, then a speaker is analogous to an organism, a linguistic community to a population, a dialect to a subspecies, a language family to a higher taxonomic grouping of species such as a family or order, and a language superfamily or stock to a class or phylum (Fig. 1).

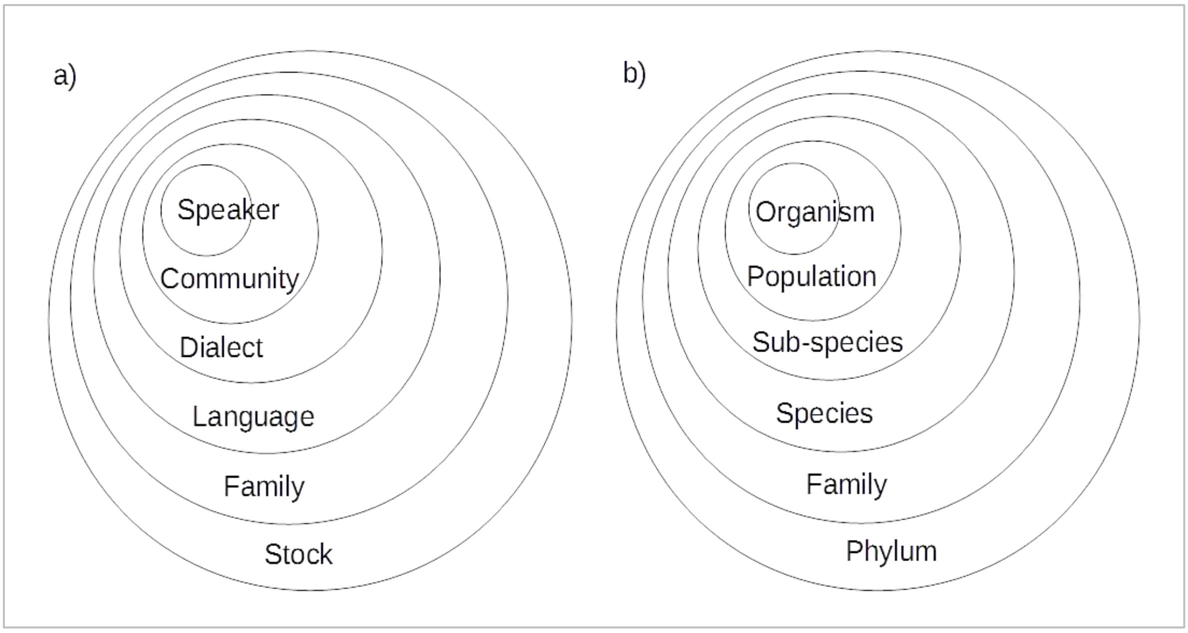


Figure 1: Parallel hierarchical classification of a) linguistic and b) biological diversity, after Harmon (1996)

Like biodiversity, linguistic diversity too can mean a generic concept, a measurable entity or a political construct. Under the generic concept, Daniel Nettle (1999) defines three kinds of linguistic diversity. The first is the number of languages in a given area or population, which he calls language diversity, and is equivalent to species richness in biology. It is referred to here as language richness.

Linguists refer to Nettle's second kind of diversity as genetic diversity, which is analogous to phylogenetic diversity in biology. It reflects the degree of difference between languages in terms of their sound systems (phonology), word formation (morphology) and grammar, and is a function of the evolutionary relationships between languages. Most of the languages of Europe, for example, belong to a single language family and are more closely related to each other than are the languages of South America, which belong to many different families, some of which have only a single member (a language isolate, equivalent to a monophyletic species). South America therefore can be said to have more phylogenetic linguistic diversity than Europe (Nettle 1999).

The third kind of diversity Nettle calls structural diversity, and refers to the variation in structures such as typology – the order in which subject, verb and object appear in a sentence – or the size of phoneme inventory, which can vary from as low as 11 to more than 100 phonemes in some African languages (Nettle 1999). Interestingly there is no clear relationship between phylogenetic and structural diversity. Structural diversity in linguistics would, in this respect, appear to be similar to karyotype diversity in biology, the variation between species in the number of chromosomes, which ranges from one (the ant *Myrmecia*) to 630 pairs (*Ophioglossum* fern). Even closely related species differ, for example Atlantic salmon *Salmo salar* has 29 pairs of chromosomes compared with 40 in the sea trout *Trutta trutta* (McVean 2002).

When it comes to treating linguistic diversity as a measurable entity it is necessary to focus on one of these levels of diversity, and language richness is the most widely used (Gavin et al. 2013). Phylogenetic diversity is very difficult to measure in languages, and linguists disagree about the relationships between languages. While most linguists agree, for example, with Joseph Greenberg's classification of African languages into four main families (Greenberg 1963), they disagree about his classification of the indigenous languages of the Americas (Greenberg 1987; Greenberg 1996) as belonging to a single Amerind superfamily or stock (Nettle 1999). Identifying high-level linguistic groupings is problematic because of an important difference between languages and species. While species other than bacteria rarely exchange genetic material with another species, languages commonly borrow words, phrases and even grammar from neighbouring languages. English, for example, belongs to the Germanic language family but has incorporated a large part of its vocabulary from French, a member of the Italic family. This borrowing or inheriting material from non-ancestral languages makes it difficult to construct phylogenies. Another problem is that languages evolve much faster than species, so similarities between languages due to common ancestry will have almost completely disappeared if the ancestral language existed at great time depth (Pagel et al. 2013). Beyond 9,000 years it becomes almost impossible to discern common ancestry, although Mark Pagel and colleagues (2013) used relationships between “ultraconserved” words to derive a date of about 15,000 years ago for the ancestor of a Eurasiatic superfamily.

Linguistic structure varies in many dimensions, so structural diversity is no easier to measure than phylogenetic diversity. Even if it could be measured, it would say nothing about the other two kinds of linguistic diversity, and would not be particularly useful. Language richness therefore has become the default measure of linguistic diversity and has been used in most quantitative studies of the subject (Mace and Pagel 1995; Nettle 1998; Nettle 1999; Moore et al. 2002; Sutherland 2003; Fincher and Thornhill 2008; Gavin and Sibanda 2012).

Linguistic diversity is also a political construct. Language is a central component of social and political identity, particularly for minority ethnolinguistic groups (Bernard 1992; Nettle and Romaine 2000; Reyhner 2007; Collin 2010) and conserving linguistic diversity is sometimes but not universally deemed to be desirable. In addition, linguistic diversity may be used as an indicator of cultural diversity which, similarly, is a value-laden term.

The CBD’s strategic plan includes a target on the conservation of “traditional knowledge, innovations and practices of indigenous and local communities” (CBD Secretariat 2004), creating a need for new cultural indicators, alongside biodiversity indicators. While there are some indicators available that can measure such traditional knowledge, innovations and practices at the local scale, such as the Vitality Index of Traditional Environmental Knowledge (Zent and Maffi 2007), these indicators are not suitable for measuring broader national, regional or global trends. However, it may be possible to do so with proxy indicators. Given that traditional knowledge is often maintained and transmitted in indigenous and local languages (Harmon 2001; Maffi 2001a; Maffi 2001b; Skyhawk 2012), indicators of the state of global linguistic diversity are the best proxies for traditional knowledge and practices at the national to global scale currently available. Although the indicators described in the submitted publications were not developed specifically for the CBD, they have been used by the CBD Secretariat for tracking progress towards the 2010 and 2020 targets (Biodiversity Indicators Partnership 2013b; CBD Secretariat 2014).

What is Biocultural Diversity?

If the broad concepts of biodiversity and linguistic diversity are hard to define, then biocultural diversity is broader and harder. For biocultural diversity, the distinction between generic concept and political construct is blurred. The concept of biocultural diversity arose directly as a consequence of the fact that something was being lost, and those who promoted the concept also argued for its conservation, hence the concept began as a value-laden one.

Biocultural diversity was first discussed in the 1980s when linguists and anthropologists began reporting on the decline in indigenous languages and traditional cultures; they made comparisons with threatened species and biodiversity loss by way of analogy, but without making deeper connections between biological and linguistic diversity (Maffi 2001a). The International Society of Ethnobiology (ISE) was among the first international fora at which biocultural diversity was discussed, and issued the Declaration of Belém at its initial meeting in 1988, which stated: "As ethnobiologists, we are alarmed that….many species, both plant and animal, are threatened with extinction; and indigenous cultures around the world are being disrupted and destroyed…[and] that there is an inextricable link between cultural and biological diversity…" (ISE n.d.).

In an article entitled *Maintaining Biocultural Diversity*, Adela Baer (1989) argued for the conservation of "biocultural human diversity", by which she was referring not to the global diversity of biological species and human cultures, but to genetic and cultural diversity within the human species alone. Nevertheless, Baer argued that because isolated and endangered ethnic groups are both genetically and culturally adapted to life in diverse physical and biotic environments all around the world, and because the most sustainable long-term strategy for the conservation of threatened habitats and species is to work with local populations rather than try to exclude them, "…the conservation of human diversity is closely tied to the conservation of all living diversity". Baer also pointed out that the extinction of endangered human cultures is irreversible, just as it is for species, and yet receives far less attention.

As the term biodiversity gained popularity, the notion behind biocultural diversity was that the "variety of life on Earth" should explicitly include human linguistic and cultural diversity too. Luisa Maffi's definition of biocultural diversity conveys the idea that biological and cultural diversity interact and influence one another over time; that human cultures evolve over time through close and constant association with other species present in the ecosystems which humans inhabit, and that many of those other species have become adapted to living with humans: "Biocultural diversity comprises the diversity of life in all of its manifestations: biological, cultural, and linguistic, which are interrelated (and possibly coevolved) within a complex socio-ecological adaptive system" (Maffi 2007). Terralingua, the NGO of which Maffi is president, describes biocultural diversity as:

> …diversity in both nature and culture. It's a living network made up of the millions of species of plants and animals that have evolved on Earth, and of the thousands of human cultures and languages that have developed over time. Languages, cultures, and ecosystems are interdependent. They're bound together through the myriad ways in which people have interacted with the natural environment. (Terralingua n.d.)

The ISE established the Global Coalition for Biocultural Diversity in 1990 "to encourage the permanent and meaningful dialogue between indigenous peoples, scientists and environmentalists in order to develop a unified strategy to defend the biological and cultural diversity of planet Earth." Biocultural diversity carried the idea that the areas of highest biological and linguistic diversity in the world are places inhabited by indigenous or traditional people (Cocks 2010). Biocultural diversity also became associated with hotspots, particularly in tropical forests, where both linguistic diversity and terrestrial biodiversity are high (Harmon 1996; Gorenflo et al. 2012).

The identification of hotspots and their association with indigenous peoples led to the perception firstly that biocultural diversity is a concept primarily concerned with those areas and peoples, and secondly that conservationists seeks to preserve biocultural diversity in a fixed, primordial state (Cocks 2006; Brosius and Hitchner 2010; Cocks 2010). While Brosius and Hitchner challenge the idea that culture is fixed in time and space, Michele Cocks argues that: "…the theory of bio-cultural diversity should extend the term 'indigenous, local' people to include more varied social groups" (Cocks 2010), and that it "…fails to take into account the multiple dimensions of culture, for example, how aspects of culture can be modified, adapted, and maintained despite changes a community might experience in its social and material context and its removal from precolonial residence areas" (Cocks 2006).

Biocultural diversity is not of course restricted to tropical hotspots and is global in scope. Elands et al. (2015), for instance, discuss urban biocultural diversity in European cities, particularly in relation to the use of green open spaces. The Florence declaration on the links between biological and cultural diversity in the European context (UNESCO and CBD Secretariat 2014) recognises that the European landscape is "predominantly a biocultural multifunctional landscape... [resulting] from the combination of historical and on-going environmental and land use processes and cultural heritage…" Nor should it be true that biocultural diversity conservation implies suspended animation. Culture and language evolve as species do, and without evolutionary processes biocultural diversity would not exist.

Part of the difficulty of discussing and defining biocultural diversity is that the subject matter crosses the divide between the natural and social sciences (Harmon 2001). Interdisciplinarity is necessary in trying to define, understand and measure biocultural diversity (Newing 2010). Loh and Harmon (2005) provide a transboundary definition of biocultural diversity as: "…biological diversity at all its levels, from genes to populations to species to ecosystems; cultural diversity in all its manifestations (including linguistic diversity), ranging from individual ideas to entire cultures; and, importantly, the interactions among all of these."

Measuring Biodiversity, Linguistic Diversity and Biocultural Diversity

The submitted publications describe methods for measuring biodiversity, linguistic diversity and biocultural diversity. All three are multi-dimensional entities and therefore no single measure can capture all dimensions. Nevertheless, the intention behind the indicators was to provide a proxy for the state of the entity in its sense of an overall, generic concept. Professional conservationists and concerned members of the public are interested to know how biological or linguistic diversity is changing over time, at a global scale, as well as to be able to compare regions or countries. Another important characteristic of a useful generic indicator is that it must be easily understandable by a non-technical audience (Gregory et al. 2005). Clearly there will be no perfect measure, but the challenge is to come up with something that will be of use.

Species and language richness have been the most widely used measures of biological and linguistic diversity. There are several short-comings associated with the use of richness data as an indicator. First, richness gives very little information about diversity at the genetic level, either variation within a species or language (intraspecific/linguistic diversity) or variation between species (phylogenetic or interspecific/linguistic diversity). A group comprising lions, tigers, leopards and jaguars (or Danish, Norwegian, Swedish and Icelandic speakers) is equivalent in richness to a group consisting of lions, swans, anacondas and whale sharks (or Danish, Basque, Javan and !Kung speakers), but clearly one is more phylogenetically diverse than the other. Until recently, phylogenetic relationships had not been sufficiently well mapped to develop a quantitative genetic diversity indicator. However, advances in generating phylogenetic trees in both biology (Isaac et al. 2007; Bininda-Emonds et al. 2008; McCormack et al. 2013; Prum et al. 2015) and linguistics (Gray and Jordan 2000; Gray and Atkinson 2003; Gray, Atkinson and Greenhill 2011) may mean that this will change.

Secondly, richness gives little information about ecological diversity. There is no direct equivalent of ecosystem diversity with languages, although one could argue that there is an analogy with types of human society or traditional livelihood that have evolved in different ecosystems, such as tropical coastal fisher, arid-zone nomadic pastoralist, mountain pastoralist, temperate lowland farmer, and so forth, each of which are associated with particular languages. By this analogy, a group of four indigenous Amazonian languages, even though they may be distantly related, would be less diverse than a group consisting of languages representing four very different cultures or lifestyles.

A third problem with species or language richness as indicators is that they are rather static, and only change very slowly with time. They are not suited to conveying the population declines over annual to decadal time-scales. The richness of a group of species or languages remains constant until one or more of its members goes extinct, even if members of the group have declined in number to the point of near extinction. Extinction has been the principal way in which the loss of species and languages is measured and discussed since the 1980s. Both biologists and linguists have argued that we are facing an extinction crisis (Krauss 1992; Wilson 1994; Heywood and Watson 1995; Leakey 1996). Biologists estimate that the current extinction rate of amphibian, bird and mammal species is between 100 and 1,000 times higher than the average background rate seen in the fossil record, and the rate for all species is expected to rise 10-100 times higher in the future (Millennium Ecosystem Assessment 2005; Baillie et al. 2010). Projections based on future habitat loss due to climate change estimate that 18-35% of species will be ‘committed to extinction’ by 2050 (Thomas et al. 2004). The introduction to the Convention on Biological Diversity describes its genesis as a response to the fact that “Species extinction caused by human activities continues at an alarming rate” (CBD Secretariat 1994).

Similar arguments have been made about linguistic diversity. Michael Krauss (1992) estimated that 90% of the world’s languages will go extinct this century, a much-cited statistic (Pinker 1995; Nettle and Romaine 2000). This projection was based on the numbers of languages that Krauss believed to be either moribund, meaning no longer being learnt by children, or endangered, which he defined as having fewer than 100,000 speakers, about 90% of the world’s 6,000 languages known at the time. Given that most of the world’s languages are spoken by fewer than 10,000 people (Lewis 2009), and have always been so, it is not fair to assume that a small population necessarily means a language is endangered. If we take Krauss’s estimate of the percentage of languages which are moribund, this would mean that 20-50% of all languages are committed to extinction.

The problem with statements about extinction projections is that they say nothing about the current rate of biodiversity or linguistic diversity loss. Biologists cannot say how many species went extinct last year, or even in the last 20 years. A species cannot be declared to have recently gone extinct in the wild because, like Karl Popper’s black swans (Popper 1959), absence of evidence is not evidence of absence. Occasionally a species which was believed to have been extinct has been rediscovered – so-called Lazarus species – such as the yellow-tailed woolly monkey (Mittermeier, de Macedo and Luscombe 1975) or the Lord Howe Island stick insect (Priddel et al. 2003).

The main feature of two of the indicators described in these publications – the Living Planet Index and the Index of Linguistic Diversity – is that they are based on trends in population (or number of speakers) across a large sample of species or languages, rather than on richness. This has two significant advantages over measures based of richness alone, or extinction. Firstly, it is possible to reveal short to medium-term trends. A population change of a few percent per year is easily discernible over annual to decadal time scales. Population-based indicators can therefore be described as being more sensitive than richness indicators. Secondly, provided that the sample of species or languages is representative, a measure based on average or aggregate population trend is indicative of diversity as a broader, generic concept. This is because population trends are not actually a measure of diversity in the strict mathematical sense, but can be considered an indicator of the state or health of “life on Earth” in a general sense. Unlike the LPI and ILD, the Index of Biocultural Diversity is based on richness data, and as a consequence it is static and does not track changes over time. However, it combines measures of biological and cultural richness to create what remains the only quantitative indicator of global biocultural diversity.

A drawback of focusing on quantitative indicators is that they steer target-setting towards readily-measurable objectives, which are not necessarily the most important. Not all objectives are easily quantified and progress towards them can be monitored only by qualitative means. The CBD Aichi Target 18, for example, which most closely relates to biocultural diversity, reads: “By 2020, the traditional knowledge, innovations and practices of indigenous and local communities relevant for the conservation and sustainable use of biodiversity, and their customary use of biological resources, are respected…at all relevant levels.” (Biodiversity Indicators Partnership 2013a). The active word in the target is “respected”, which is hard to quantify. Parties to the CBD must rely on narrative accounts of progress to Target 18. Quantitative indicators are a complement to detailed qualitative knowledge of biocultural diversity, not a substitute for it. It is recognized that global or regional indices such as the IBCD, LPI or ILD are useful when used to provide a simplified overview of a complex picture, and contextual data for informing policy frameworks, but not for guiding locally-specific policy decisions.

Development of the Indicators

The Living Planet Index (LPI)

The LPI was conceived as part of WWF's Living Planet Campaign which ran from 1997 to 2000. The original idea of Jorgen Randers, then with WWF-International and one of the co-authors of the paper submitted (Loh et al. 2005), was to develop an index which would answer the question, "how fast are we losing nature?" At the time, the most widely used measure of the state of global biodiversity was the number or proportion of known species listed as threatened (World Resources Institute 1996; UNEP 1999; OECD Environment Directorate 2001; Prescott-Allen 2001) on the IUCN Red List (IUCN 1996; IUCN and WCMC 1998). The first LPI was designed to show the changing state of forests, freshwater and marine ecosystems worldwide, and was based on trends in global forest cover and populations of freshwater species and marine species (Fig. 2). It was published in WWF's *Living Planet Report 1998* in collaboration with WCMC as a measure of "the health of the world's natural ecosystems and biodiversity" (Loh et al. 1998) .

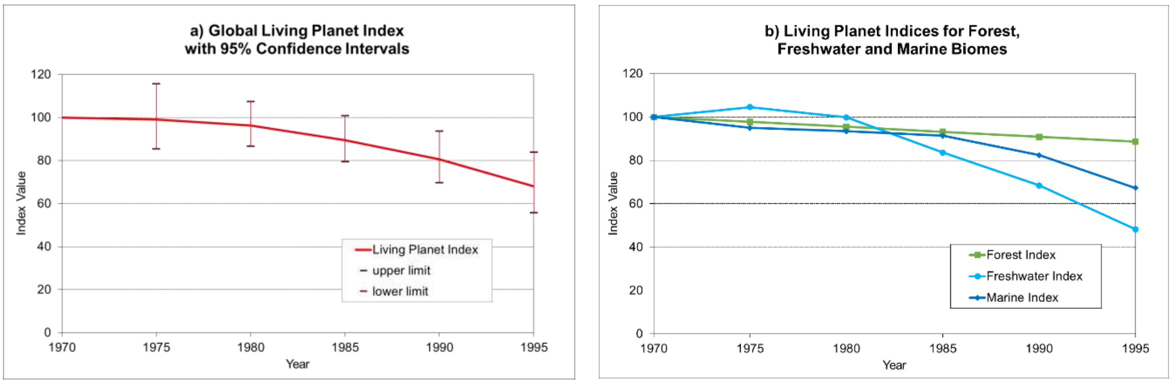


Figure 2: The first Living Planet Index from WWF Living Planet Report 1998: a) the aggregated global index, b) forest, freshwater and marine biome indices

Subsequently, global forest cover was replaced by average trends in populations of forest species worldwide (Loh 2000) and, later, populations of any terrestrial species (Loh and Wackernagel 2004). With these changes, the LPI became entirely an index of trends in species abundance.

In July 2004 the Royal Society held a discussion meeting entitled *Beyond extinction rates: monitoring wild nature for the 2010 target*. The first publication submitted here is the paper on the LPI from that meeting (Loh et al. 2005). Defined as the average of the three biome indices, giving equal weight to marine, freshwater and terrestrial species, the global LPI showed a 24-48% decline over 30 years (Fig. 3). One important advance in the LPI method made for that paper was the calculation of confidence intervals by bootstrapping.

In 2006, a new LPI database, data collection and data management system were established at ZSL. The volume of species population time-series data increased significantly (Fig. 4). The method for calculating the LPI was also improved by shifting from a five-yearly to an annual basis, using log-linear interpolation to estimate annual species abundance values, and fitting general additive models to time-series comprising more than six data points. Software was written in R to enable the index and confidence limits to be calculated automatically. A second paper on the LPI was published to document these changes to the method and update the results (Collen et al. 2009).

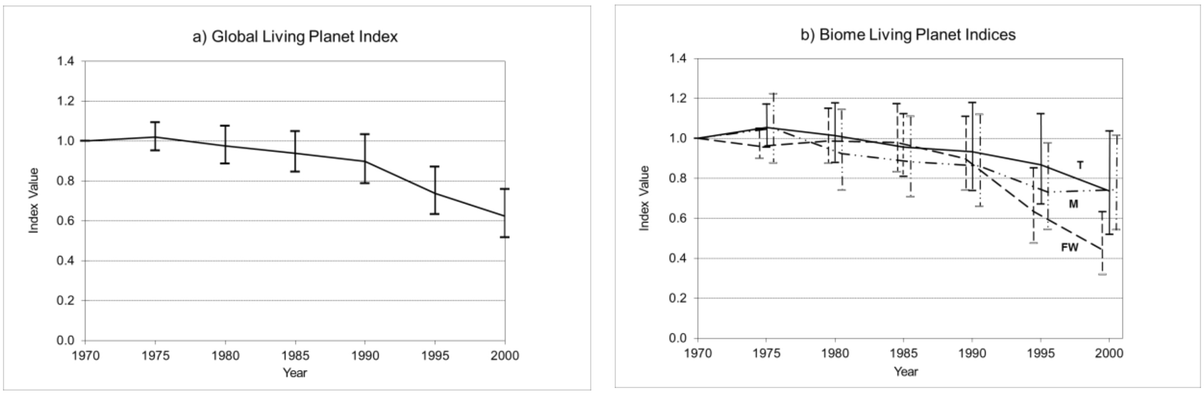


Figure 3: Living Planet Index (from Loh et al. 2005) a) Global LPI with 95% confidence intervals and b) LPIs for terrestrial (T), freshwater (FW) and marine (M) systems with 95% confidence intervals

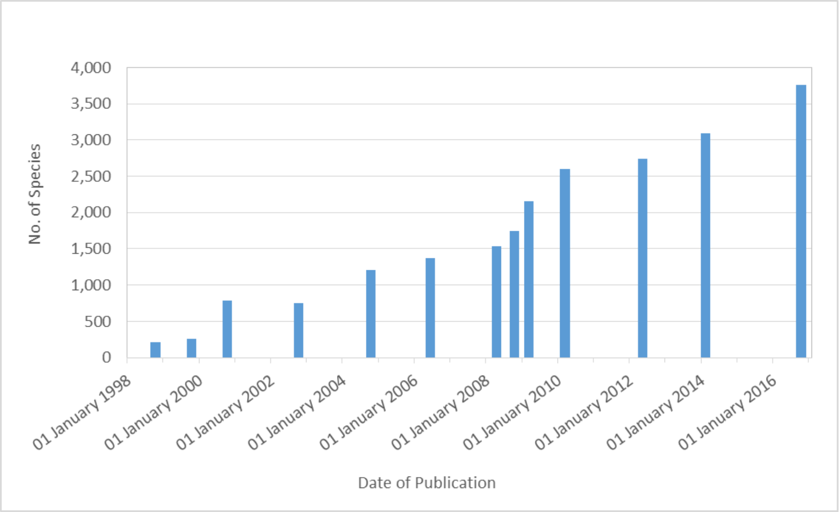


Figure 4: Number of species in the LPI in successive editions of Living Planet Report and other reports

The Index of Biocultural Diversity (IBCD)

The IBCD (Loh and Harmon 2005) was an attempt to create the first measure of biocultural diversity at the national level. The nature of the data available meant that the indicator would measure biocultural diversity at a single point in time, approximately the end of the 20th century. The index made use of data from multiple sources on linguistic, religious and ethnic diversity as well as (wild) plant, bird and mammal species diversity to score 221 countries and territories in terms of their overall biocultural diversity. Insufficient data were available to include domesticated species or cultivated varieties. Cultural and biological diversity were equally weighted in the index. Three different indices were produced, one based on unadjusted richness (IBCD-RICH), a second adjusted for land area (IBCD-AREA) and a third adjusted for population (IBCD-POP). Overall, Indonesia and Papua New Guinea were the countries with the highest IBCD scores (Fig. 5).

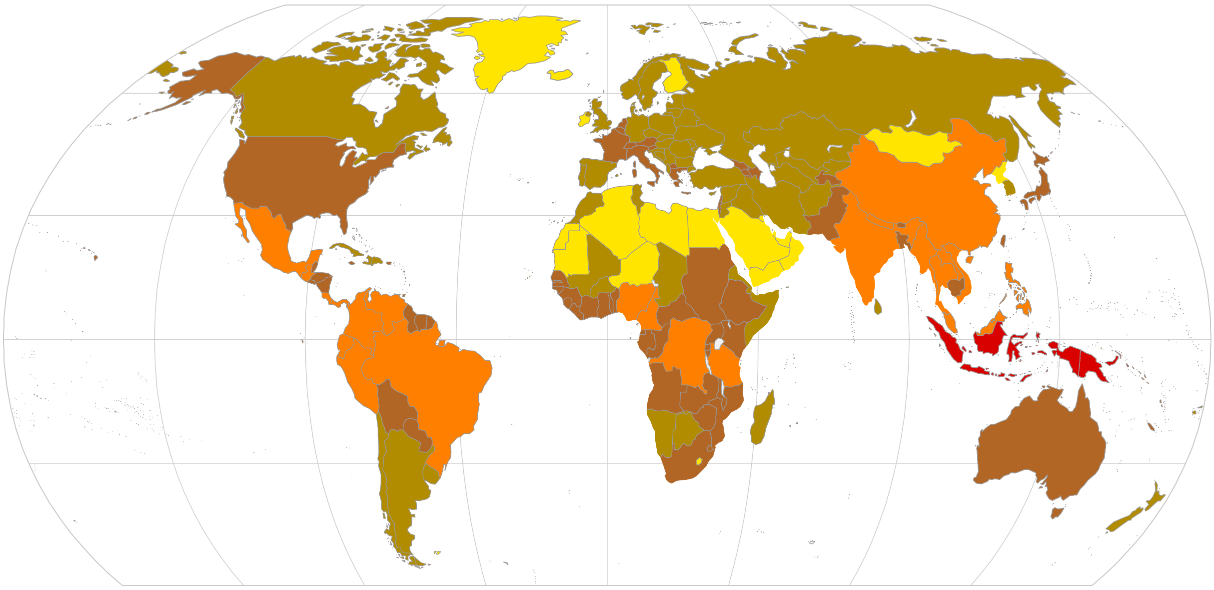


Figure 5: Index of Biocultural Diversity, area-adjusted (IBCD-AREA). Red (0.7-1.0) > orange (0.6-0.7) > brown (0.5-0.6) > green (0.4-0.5) > yellow (0.0-0.4); 1.0 is the highest possible score and 0.0 the lowest (Loh and Harmon 2005). Indonesia and Papua New Guinea have the world's highest scores at 0.76 and 0.75 respectively.

The Index of Linguistic Diversity (ILD)

The ILD measures trends in linguistic diversity by applying the LPI method to data on numbers of speakers of languages. The ILD was part of a larger project funded by the Christensen Fund, *Global Indicators of the Status and Trends of Linguistic Diversity and Traditional Knowledge*, which began in 2007 along with work on another indicator, VITEK, developed by Stanford Zent of IVIC in Venezuela. One objective of the project was to compare global trends in linguistic diversity with trends in biodiversity (Harmon and Loh 2009).

Time-series data were compiled on numbers of speakers of 1,500 languages, around one fifth of the world's total. For each language, the ILD calculates how its share of the global or regional population changes over time. For a selection of languages, the ILD calculates how their average share of the global or regional population changes over time. The global ILD declined by about 20% between 1970 and 2005, but the regional trends varied between +7% (Eurasia) and -64% (Americas) over the same period (Fig. 6).

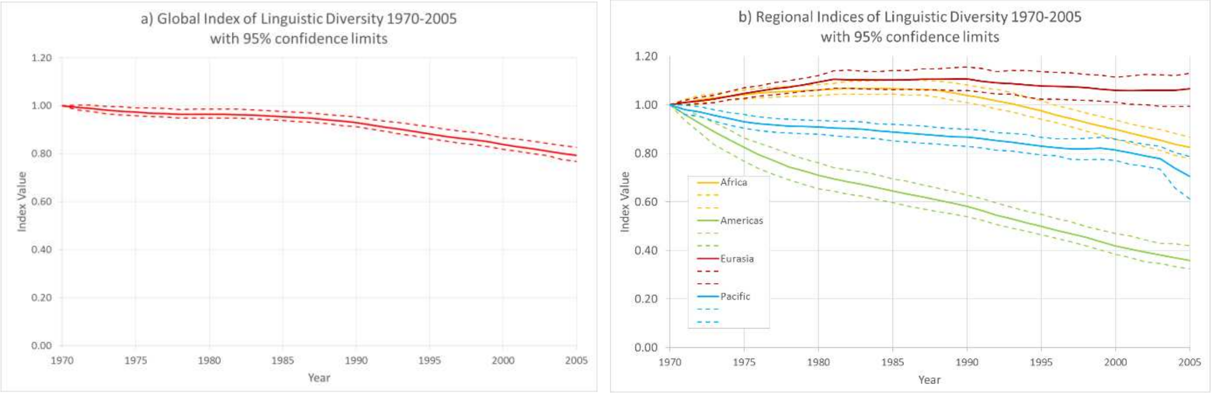


Figure 6: Index of Linguistic Diversity (Harmon and Loh 2010); a) the global index declined by about 20% from 1970 to 2005 while b) the regional indices varied between a positive trend (Europe), a 17% decline (Africa), a 30% decline (Pacific) and a 64% decline (the Americas).

Biocultural Diversity: threatened species, endangered languages

The most recent publication (Loh and Harmon 2014) used a two-pronged approach to compare status and trends in linguistic diversity with biodiversity. Firstly it brought together the LPI and an updated ILD to compare their trends at both global and regional levels. Secondly, to complement the comparison of trends, it used the IUCN Red List criteria to assess the threat status of the world's languages and then compare languages with mammals, birds, reptiles and amphibians (Fig. 7). See Appendix (p. 30) for a description of the methods used. There are other systems of assessing the threat status of endangered languages, such as UNESCO's *Atlas of the World's Languages in Danger* (Moseley 2010), but they rely on different criteria to those used by biologists to assess species. By using the same indicators and criteria it was possible directly to test assertions such as linguist Michael Krauss's, that "The biological species endangerment rates of the very worst biological alarmist estimates, e.g. that 50 per cent (not just 5 per cent) of mammals may be endangered (or threatened), are still better than the very best language endangerment scenarios I can imagine" (Krauss 2007).

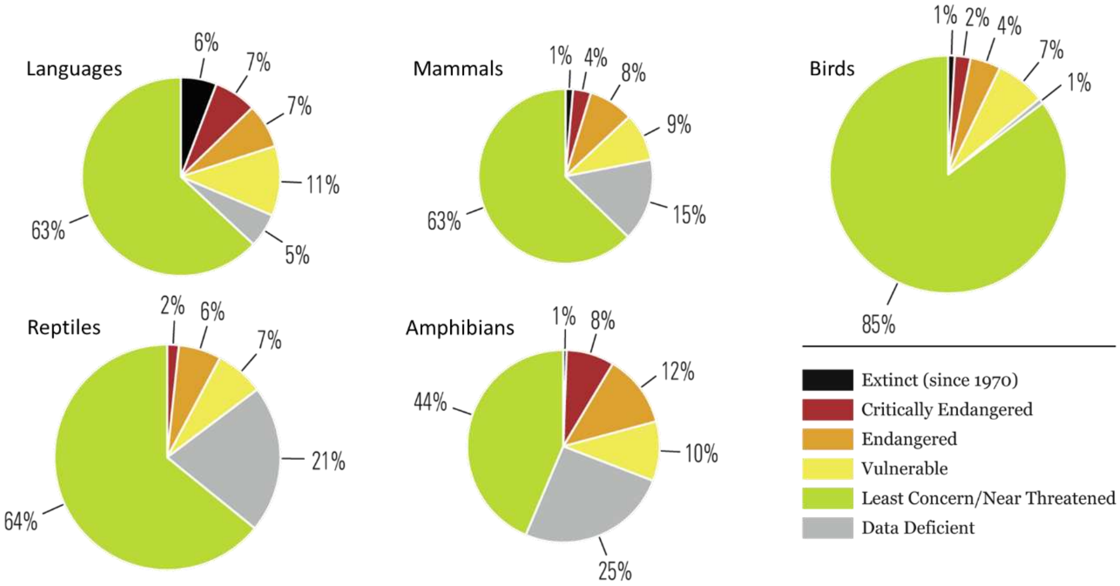


Figure 7: Red List status of languages and four vertebrate classes (size of each pie chart is proportional to the global number in each group. Mammal, bird and amphibian data, IUCN (2013); reptile data, Bohm et al. (2013). The status of languages is at least as seriously threatened as the vertebrate groups.

Both LPI and ILD show similar trends at the global scale, and the Red List assessment suggests that the overall threat status of the world's languages is at least as serious as that of vertebrate species. At the regional level, however, a different picture emerges (Fig. 8). The LPI indicates a north-south split in trends in biodiversity, while the ILD shows an east-west split in trends in linguistic diversity. Specifically, the LPI declined rapidly in the tropics but displayed little overall change in temperate realms. The ILD by contrast declined very rapidly in the Americas, and more slowly in the Africa, Eurasia and the Pacific. In Australia, however, the ILD fell even faster than in the Americas. This pattern in the regional ILDs is borne out by the regional Red Listing of languages.

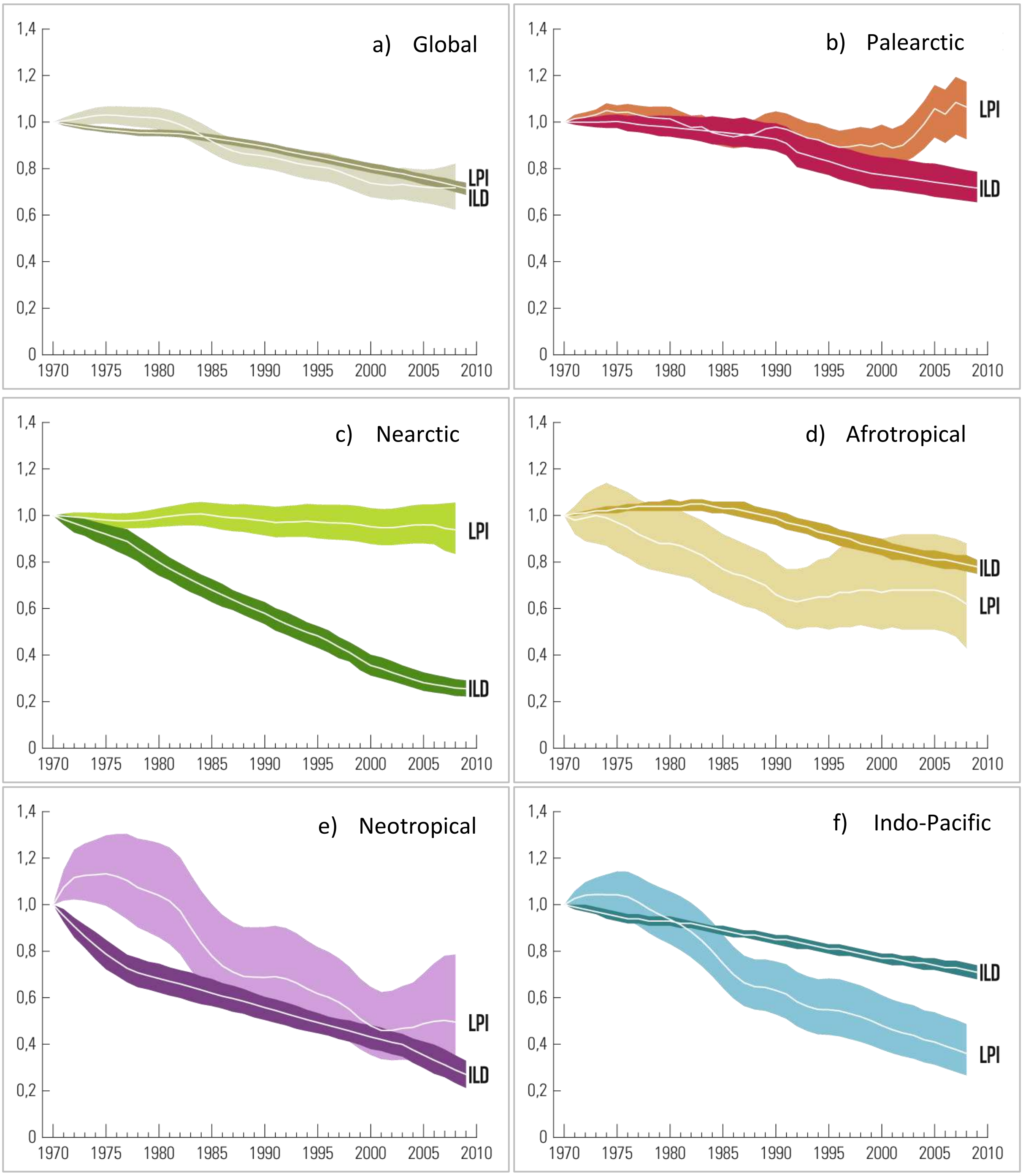


Figure 8: Comparison of the Living Planet Index with the Index of Linguistic Diversity from 1970 to 2007/8 (Loh and Harmon 2014). Both indices are set equal to 1.0 in 1970. While the LPI and ILD decline at a similar rate globally (a), the indices for b) the Palearctic, c) Nearctic, d) Afrotropical, e) Neotropical and f) Indo-Pacific biogeographic realms show different patterns.

The *Biocultural Diversity* report adopted an evolutionary and ecological approach to species and languages in an attempt to provide a more integrated concept of biocultural diversity than that of the IBCD. It is not necessary to think about languages and species in this way in order to compare their conservation status and trends, but it does provide a unified theoretical framework for biocultural diversity as a generic concept and sets its current status in an evolutionary context.

Impacts of the Indicators

The Living Planet Index

International Policy Targets and Indicators

At the World Summit on Sustainable Development in Johannesburg, 2002, parties to the CBD agreed “to achieve by 2010 a significant reduction of the current rate of biodiversity loss at the global, regional and national level as a contribution to poverty alleviation and to the benefit of all life on Earth” (CBD Secretariat 2006). At the seventh Conference of the Parties (CoP7) of the CBD in Kuala Lumpur in 2004, governments agreed a set of quantifiable goals and indicators to measure progress towards the 2010 target. One of the indicators identified for immediate testing was “Trends in abundance and distribution of selected species” (CBD Secretariat 2004). The LPI was adopted as an indicator for the 2010 target by the CBD under the Biodiversity Indicators Partnership (BIP) process, and reported in its *Global Biodiversity Outlook* series (CBD Secretariat 2006; CBD Secretariat 2010). A paper in *Science* analysed 31 indicators of progress towards the 2010 target, including the LPI, and found no significant reduction in the rate of loss on biodiversity (Butchart et al. 2010).

Following the world’s failure to achieve the 2010 target, CoP10 of the CBD held in Nagoya, Japan, adopted the 2011-2020 Strategic Plan for Biodiversity which included a new set of “Aichi” targets. These relate to the five strategic goals of the plan, one of which is “to improve the status of biodiversity by safeguarding ecosystems, species and genetic diversity”. Again, the LPI was adopted as an indicator and reported in *Global Biodiversity Outlook 4* (CBD Secretariat 2014) and *Aichi Targets Passport* (Biodiversity Indicators Partnership 2013a). A follow-up paper in *Science* reported on mid-term progress towards the Aichi targets (Tittensor et al. 2014). The authors used 55 indicators including the LPI to report against the 20 targets, and concluded that current policy and management efforts to stem biodiversity loss are unlikely to result in an improvement in trends by 2020. UNEP’s international environmental assessment report the *Global Environment Outlook* also published the LPI (Ash et al. 2007; Armenteras et al. 2012), and another international biodiversity assessment and reporting process, IPBES, has also selected the LPI as a global and regional biodiversity indicator.

Other international indicators based on species population trends have been developed over the last decade which relate to single taxa such as the Global Wild Bird Index (Stattersfield, Bennun and Jenkins 2008; BirdLife International 2013) or the European Grassland Butterfly Indicator (European Environment Agency 2013). Birds and butterflies have well-developed monitoring schemes, especially in Europe and North America, and lend themselves to multi-species population indices. The LPI however remains the largest and longest-established global dataset on population trends for vertebrates.

National LPIs

The first national application of the LPI was in Norway, where WWF produced a Norwegian Nature Index (WWF Norway 2005). This index was later adopted, developed and published by the Norwegian Environment Agency (Nybø, Certain and Skarpaas 2011) and is included in the government’s official indicators on sustainable development. The methodology has evolved and been adapted over time by researchers at the Norwegian Institute for Nature Research (Aslaksen and Garnåsjordet 2012). The second implementation of a national LPI was by WWF Canada in their Canadian *Living Planet Report* (Mitchell, Loh and Goldfinger 2007; McRae et al. 2007).

There have also been some unsuccessful attempts by WWF national organizations to produce a national LPI. South Africa, Turkey and the Guianas tried but failed to produce a national index because of insufficient data. The LPI demands data on species populations going back at least two decades, covering a representative sample of habitats, taxa and geographic regions. This data demand presents a barrier for the LPI as a national biodiversity indicator which many countries cannot overcome, although the LPI database at least provides a starting point and a set of criteria for future data collection.

A successful example of an LPI for a developing country was the Uganda Biodiversity Index. This indicator was based on data collected by the National Biodiversity Data Bank (NBDB) at Makerere University. The *State of Uganda's Biodiversity* reports (Arinaitwe, Pomeroy and Tushabe 2000; Pomeroy and Mwima 2002) included data on a large number of species population trends, some of which dated back as far as the 1960s. Data from the 2000 report was used to construct a *Living Uganda Index* using the LPI method (Jenkins, Kapos and Loh 2004) which was then incorporated into the *State of Uganda's Biodiversity 2004* report (Pomeroy and Tushabe 2004). The 2006 and 2008 reports (Pomeroy, Lutaaya and Tushabe 2006; Pomeroy and Tushabe 2008) took the LPI method a stage further. As well as species population trend data, NDBD collected data on species richness and habitat cover. These datasets included trends on, for example, the number of fish species in Lake Victoria, trees species in protected areas, area of forest cover and area of wetlands. The LPI methodology was used to create a Species Richness Index, a Species Population Trends Index and a Habitat Cover Index, which were then combined into an overall Uganda Biodiversity Index, and Ugandan species population trends were compared with global trends. Data for an updated index are currently being collated.

Regional and Thematic LPIs

In collaboration with ZSL, the Circumpolar Biodiversity Monitoring Program (CBMP) produced an Arctic LPI. Data were collected by the CBMP, particularly from Russia, and the resulting index included trends for approximately one-third of all Arctic vertebrate species (McRae et al. 2010; McRae et al. 2012). An LPI of Mediterranean Wetlands was produced in collaboration with Tour du Valat (TdV), a research institute in France, using data collected from 27 countries around the Mediterranean (Galewski et al. 2011). This indicator formed a part of the Mediterranean Wetlands Observatory project. Thematic analyses include LPIs of migratory species for the Convention on Migratory Species (Latham et al. 2010), estuaries (Deinet et al. 2010), vertebrates (Baillie et al. 2010), protected areas (Craigie et al. 2010; Milligan et al. 2014) and recovering species for Rewilding Europe (Deinet et al. 2013).

New indicators based on the LPI method

The LPI method has been applied to datasets other than species populations to produce average trends. One example is an attempt to create an index of trends in habitat extent based on incomplete global data on changes in wetland area worldwide, the WET index (Dixon et al. 2016). Here the LPI method was applied to a patchy dataset to show trends in wetland area within defined geographic units and aggregated these to regional and global scales. Another example is the ILD.

Index of Biocultural Diversity

The IBCD was intended to be useful as a policy tool for measuring and monitoring the level of biocultural diversity at the national level, and it was described in the IUCN journal Policy Matters as with that purpose in mind (Harmon and Loh 2004b). However, its utility was limited in two ways which meant that it has had more impact in raising the profile of biocultural diversity in academic circles than in policy circles. Firstly it was not capable of showing trends, making it a purely static measure, and secondly it was not capable of finer resolution at a sub-national level; consequently it became more useful to those interested in a global overview rather than to authorities responsible for implementing conservation policy. Nevertheless it remains the only indicator which set out specifically to be a measure of biocultural diversity for nearly all countries in the world.

Index of Linguistic Diversity

In comparison with the LPI, there has been limited uptake of the ILD by the media. Following its publication in 2010 and update in 2014 there was some coverage of the ILD in the press (Braun 2011; Vidal 2014).

The ILD has had some impact in policy fora. One of the CBD's 2010 targets was to protect the traditional knowledge, innovations and practices of indigenous and local communities, a target which remains in the 2020 Aichi framework. One indicator selected for this target was trends in linguistic diversity and numbers of speakers of indigenous languages. Linguistic diversity is a proxy for traditional knowledge, and more direct indicators have been proposed including the VITEK (Zent and Maffi 2007) and others (Anderson and Poppel 2002; Lasimbang 2008; Statistics New Zealand 2008), but data for these are only available from a limited number of locations and are inadequate at the global or regional scale.

UNESCO was chosen as the partner organization to develop the indicator for the CBD, but has yet to publish it. The ILD however was published in 2010 and was subsequently incorporated as an Aichi target indicator through the BIP (Biodiversity Indicators Partnership 2013b). The biogeographic realm ILDs from *Biocultural Diversity: threatened species, endangered languages* (Loh and Harmon 2014) were incorporated into recent regional versions of *Global Biodiversity Outlook 4* on mid-term progress towards the Aichi targets (UNEP-WCMC 2016a; UNEP-WCMC 2016b; UNEP-WCMC 2016c). Given that many indigenous people who care about the continuation of their traditional culture believe that maintaining their language is essential (Bernard 1992; Reyhner 2007; Skyhawk 2012), the continual decline of the global ILD from the 1970s to the 2000s suggests that the CBD target to safeguard traditional knowledge, innovations and practices has not been met.

Critique of the Indicators

The Living Planet Index

Richard Gregory and others (2005) presented a set of 14 qualities of an effective biodiversity indicator for scientific and policy purposes. The LPI performs quite well against many of these criteria, being:

a) regularly updated;
b) transparent and easy to interpret;
c) easily understood by non-experts, including policy-makers and members of the public;
d) quantitative and showing rate of change over time;
e) responsive to environmental change over relatively short time-scales
f) able to be disaggregated to help understand patterns and potential causes of trends;
g) based on available or easily-collected, quantitative data, not requiring excessive financial resources;
h) indicative of attributes of biodiversity, ideally reflecting ecosystem health;
i) user-driven in response to the needs of stake-holders;
j) policy-relevant, to help develop and review policy measures;
k) stable, buffered from irregular, large natural fluctuations; and
l) susceptible to human influence and change.

Arco van Strien and colleagues (2012) further described a set of desirable mathematical properties of indicators based on species trends, derived from economic theory relating to price indices. They are:

a) if all species are declining, then the index declines, and vice versa;
b) if all species change by a common factor, the index also changes by that factor;
c) if all abundances in one year are the same as another year, the index is also the same;
d) index is insensitive to base year;
e) index is not dominated by species appearing or disappearing from the ecosystem;
f) index is not sensitive to spatial scale.

According to these criteria they concluded that out of the indicator methods they evaluated, including the Shannon and Simpson indices, arithmetic mean and mean species abundance, the geometric mean of species' populations is the most favourable. The geometric mean is the basis of averaging the species trends in the LPI (and language trends in the ILD) and so puts the LPI on a sound mathematical basis as an indicator.

However, two limitations of the LPI, where it falls short of Gregory *et al*.'s criteria, are its timeliness – the ability to identify trends rapidly and give an early warning of issues – and its representativeness of all species in a group or taxon, which in the case of the LPI means global biodiversity. Timeliness and representativeness are problematic issues, both of which relate to data availability. Another possible weakness is the quality of the available data.

The LPI makes use of available population trend data on vertebrate species. The data come from published studies in the scientific literature, reports from governmental and non-governmental organizations, online databases and some wide-scale multi-species surveys such as the Pan-European Common Bird Monitoring Scheme (European Bird Census Council). One of the limitations of the LPI is that there is a time-lag between a survey being carried out and the results being published in a journal or online, and then another delay between the data becoming available and their incorporation into the LPI database and the eventual publication of an updated LPI. There has always been at least a three-year lag between the most recent index year in any LPI and its publication date. Furthermore, the time-lag means that the most recent years in any LPI have fewer data points than earlier years, and therefore higher uncertainty around the most recent index values. This means the LPI cannot act as an early warning signal for emerging issues.

Perhaps the most important question and criticism of the LPI is how representative are the LPI data geographically (Pereira and David Cooper 2006; Proença et al. 2016) and, to widen the question out more broadly, how representative is the LPI of global biodiversity? Geographically, the coverage of time-series data included in the LPI database is skewed towards Europe and North America, where species populations are surveyed most intensively (Proença et al. 2016). This bias is particularly strong for bird species, for which large datasets are available from the North American Breeding Bird Survey (Patuxent Wildlife Research Center and Canadian Wildlife Service 2001) and the Pan-European Common Bird Monitoring Scheme (European Bird Census Council 2016). To compensate for the geographic bias to some extent, the LPI calculates trends for either biogeographic realms (Loh et al. 2005) or temperate and tropical zones (Collen et al. 2009) (or ocean basins in the case of marine species) independently, regardless of the number of time-series they include, and then combines those trends with equal weighting in the aggregated global LPI. This reduces but does not eliminate the geographic bias because tropical species are far more numerous than temperate (and, according to the LPI data, declining more rapidly than temperate species).

Imagine a *perfect LPI*, an indicator based on annual population trends of every species living on Earth: vertebrates and invertebrates; plants and animals; prokaryotes and eukaryotes. The aggregated trend would not be a proxy but a direct measure of the overall state of global biodiversity. How would it look compared with the actual, imperfect LPI? We can only speculate because we cannot count every species of life on Earth, let alone collect data on their populations. The actual LPI is very limited, containing less than 0.1% of all species. It is restricted to vertebrates only, for reasons of data availability. Expanding the LPI to incorporate invertebrates has been attempted in the past, but the only available long-term time-series data, on butterflies and commercially-harvested crustaceans, were highly unrepresentative of invertebrates as a whole. Yet we can ask whether population trends in vertebrate species are likely to be representative of all species of life. Because most of the species at the highest trophic level in any ecosystem are vertebrates, we could assume that they are a reasonable proxy for everything going on beneath them. Conservation organizations have focused their attention on vertebrates with the justification that protecting those species in their natural habitat will protect other species too.

If we accept that an index of vertebrate species trends is the best we are able to deliver, then we can ask whether the species included in the LPI database are representative of all vertebrate biodiversity. The answer is that birds and mammals are over-represented at the expense of reptiles, amphibians and fish. Populations of birds, the most over-represented class, are declining more slowly on average than other classes, which makes the LPI under estimate the overall rate of decline. The most recent LPI, in *Living Planet Report 2014*, weighted trends for each class according to the estimated number of species in each class, as suggested by Loh *et al.* (2005) and Collen *et al*. (2009). The resulting global index declined by over 50% from 1970 to 2010 (McRae, Freeman and Deinet 2014). This may well be a more accurate indicator of global trends in vertebrate populations, but it puts the greatest weight on fishes, the group with the poorest representation in the data.

The LPI can only be as accurate as its underlying time-series data. Data therefore have to meet certain criteria for inclusion in the LPI database: population estimates from at least two points in time using the same method must be given for a named species and location. The species name, location, year, population estimate, method and source of the data were entered into the LPI database along with any other relevant information such as known threats or management. A system was devised to score the quality of each time-series, which could then be used to test the sensitivity of the calculated indices to data quality. It was found that lower quality data tended to estimate population changes positively compared with higher quality data, and therefore the effect of including them in the index was to reduce the average rate of decline. Because the lower quality data did not have a strong effect on the indices, all data were included in the LPI.

The Index of Biocultural Diversity

The IBCD gives a snapshot of biocultural diversity at the national level, approximately at the turn of the millennium. Biological and cultural diversity were measured using several indicators of species richness and cultural richness. At the time of publication, there were few datasets on biological and cultural richness available at the national-level and covering most of the world's countries (Harmon and Loh 2004a). For species richness, near-complete national inventories were available for birds, non-marine mammals and vascular plants (Groombridge and Jenkins 2002), while for cultural richness, inventories were available for languages from *Ethnologue* (Grimes 2000), and for religions and ethnic groups from the *World Christian Encyclopaedia* (Barrett, Kurian and Johnson 2001). There was no international standard or UN agency responsible for the collection of these data, and they were assembled from national censuses and biological inventories. Any such global dataset may therefore be incomplete, inconsistent or of uneven quality (Harmon and Loh 2004a).

One can question whether language, religion and ethnic group richness are the most appropriate indicators of cultural diversity. The OED defines culture, in the sense that it is meant here, as "The distinctive ideas, customs, social behaviour, products, or way of life of a particular nation, society, people, or period. Hence: a society or group characterized by such customs, etc." (OED Online 2016b). Richerson and Boyd (2005) define culture in a way that stresses the idea of information transmission between individuals: "Culture is information capable of affecting individuals' behavior that they acquire from other members of their species through teaching, imitation and other forms of social transmission." The information may take the form of ideas, knowledge, beliefs, skills, values or attitudes (Richerson and Boyd 2005). Cultural indicators, therefore, may be based on language, religion, agriculture and food systems, traditional knowledge or ceremonies, among many other possibilities (Harmon, Woodley and Loh 2010). Cultural diversity has more dimensions than language, religion and ethnicity but, given the availability of suitable datasets, these were the best proxies. Subsequent studies have used the same three measures of cultural diversity (Patsiurko, Campbell and Hall 2012). As cultural practices such as farming systems and traditional knowledge are often transmitted vertically down the generations (Richerson and Boyd 2005), language and ethnicity which are predominantly passed from parent to offspring are reasonable proxies for other cultural traits or values.

Ethnic groups can be considered to be an aspect of cultural diversity. Currie and Mace (2012) define them as: "…groups of individuals that share a common, self-ascribed identity based on the belief in common descent, and/or a shared culture, distinct in some respect from those of other groups. This belief is culturally inherited from one generation to the next. Symbolic markers of these groups such as differences in dress, language, and institutions are themselves culturally inherited." Ethnolinguistic fractionalization (ELF), a measure based on the linguistic and/or ethnic composition of a population, has been widely used as a measure of diversity at the national level (Alesina et al. 2003; Fedderke, Luiz and Kadt 2008; Bossert, D' Ambrosio and La Ferrara 2011; Patsiurko, Campbell and Hall 2012;

Sturm and De Haan 2015). ELF indicators are based on the probability that two individuals chosen at random from a population belong to different ethnolinguistic groups, and are analogous to the Simpson diversity index in ecology (Harmon, Woodley and Loh 2010). Their applicability in the analysis of ethnic politics has been questioned on the grounds that ethnic groups are notoriously difficult to define and take no account of the cultural distance between groups (Fearon 2003) or that not all ethnic groups are politically relevant (Posner 2004). Nevertheless, fractionalization is a useful, intuitive indicator and superior to richness alone because it takes population shares into account (Fearon 2003). Ideally, the IBCD would be calculated as the fractionalization of both ethnolinguistic groups and species but, because this measure requires data on numbers of individuals belonging to each group or species, there are insufficient data to calculate national species fractionalization indices, even for the best studied taxa, birds and mammals.

Religion is an important dimension of cultural diversity within a population, but has some particular limitations as an indicator. Religious groupings are hard to define and therefore difficult to count, although an analysis of national data in the *World Christian Database*, which superseded the *World Christian Encyclopaedia*, found that “religious composition estimates…are generally plausible and consistent with other datasets” (Hsu et al. 2008). The nation-state has been questioned as the appropriate level for the analysis of religious diversity because globalization, the internet and migration have spread world religions and new religious movements, so increasing religious diversity globally as well as at the national level (Beyer 2010; Bouma and Ling 2011). However, because of the importance of religion to cultural identity, it was included as a component of the IBCD.

The IBCD is a national-level indicator, but species, languages and cultures do not respect borders. It would be preferable to use a system of dividing the world into more meaningful biocultural units, perhaps along the lines of ecoregions, but data on languages and cultures are collected at the national level and are not readily available in any other way. That being said, if the object is to influence policy-making, national-level indicators are more useful than ones based on non-political entities.

Finally, because it is based only on richness counts, the IBCD is not suited to tracking change in diversity over time. Even if populations of some species or ethnolinguistic groups were to decline rapidly, the richness count remains unchanged until one or more extinctions occurs. For this reason, the IBCD would be greatly improved by incorporating population data. An initial step towards this goal was the development of the ILD.

The Index of Linguistic Diversity

The original idea behind the ILD was to apply the LPI methodology to languages, using numbers of speakers of languages in place of populations or abundance of species. In many ways, the task was more straight-forward for languages than for species. The number of languages in the world is known reasonably well and there are data on numbers of speakers for most of them. Furthermore, historic estimates going back several decades exist for many languages, meaning that it is possible to create an index of trends that is reasonably representative of global linguistic diversity. These data come largely from a single source, *Ethnologue: Languages of the World*, a catalogue of the world’s languages which lists information on their geographic location, number of speakers, phylogenetic classification, viability and availability of literature. It has been published and periodically updated since the 1950s and is now in its 19th edition online (Lewis, Simons and Fennig 2016).

The first edition of *Ethnologue* (Wycliffe Bible Translators 1951) included only 46 languages. The editors described it as an attempt to start listing and classifying languages in the way that the world’s flora and fauna had been. By the eighth edition (Grimes 1974), *Ethnologue* listed 5,587 languages and was the first complete catalogue of known living languages. The number of languages increased in subsequent editions, largely through reclassification and splitting of known languages, and now stands at 7,097 (Lewis, Simons and Fennig 2016). Although it is criticized for being published by a Christian organization, *Ethnologue* remains the only comprehensive source of data on the world's languages and speaker numbers, nor is there reason to suppose that the data likely to be less reliable because of its evangelical origins.

The first version of the ILD (Harmon and Loh 2010) was based on data extracted from the first to the 15th editions of *Ethnologue* (Gordon 2005) while the second version of the ILD (Loh and Harmon 2014) incorporated additional data from the much-improved 16th edition of *Ethnologue* (Lewis 2009) which contained 6,909 living languages. Because the ILD is based on a random sample of 1,500 languages taken from Gordon (2005), the trends in the ILD are reasonably representative of the world's languages globally and regionally, as well as of language families. This is a major advantage over the LPI which is not based on a random sample of the world's vertebrate species.

The world's human population approximately doubled in the period that the ILD covers. The default trend in the number of speakers of most languages is to increase at the same rate as the human population of the region in which the language is spoken. Without any change in diversity, therefore, the global ILD would approximately double between 1970 and 2005. In order to cancel out this effect the ILD is adjusted for overall population growth, both at global and regional levels. This is an important difference between the ILD and LPI, and has implications for what it is that the ILD actually measures. Instead of looking at trends in numbers of speakers of languages, and calculating the average trend, as the LPI method does, the ILD looks at trends in the share of the total population represented by each language, and calculates that average trend. This is, mathematically, a truer measure of diversity, although a less intuitive concept. If the LPI is analogous to a stock market index based on the market capitalization of listed companies, then the analogy for the ILD is market share. If average market share of a sample of languages is declining, it means that a few languages are increasing their share while a greater number are losing theirs.

SIL, *Ethnologue*'s publisher, does not collect primary data on speaker numbers itself but collates and makes use of data from many sources. These include government censuses and academic articles, as well as reports from missionaries. There is no consistent method of data collection. Government censuses may over-report or under-report the number of speakers of some languages for political reasons (Collin 2010). Unlike the time-series data in the LPI, successive editions of *Ethnologue* are not necessarily comparable. The ILD data, therefore, are susceptible to the problem of non-genuine changes. For example, linguist A might assess the number of speakers of a language to be N. Years later, linguist B estimates it to be N/2. This may be a genuine change because the number of speakers halved, or it may be a non-genuine change because A and/or B estimated incorrectly. To avoid potentially non-genuine changes in speaker numbers, data points which implied very rapid rates of change (halving or doubling within three years) in large populations (N>1000) were removed as a precaution.

While an animal belongs to only one species, a person may speak more than one language. This presents a challenge for the ILD, which treats each speaker as being monolingual, and counts only "mother tongue speakers". Multilingualism is the norm in many societies where an individual may, for example, speak one language at home, another at work or school and perhaps other regional or foreign languages as well. A decline in mother tongue speakers may not mean that a language is in trouble if there is a large pool of fluent second language speakers. Nevertheless, language loss often starts with a bilingual phase where children are equally able to use their mother tongue alongside another, more widely-spoken language outside the home. The following generation becomes less fluent in the mother tongue. Bilingualism is not necessarily a buffer for declining languages, as has been the case for Spanish and Asian languages among second and third-generation immigrants in the United States (Portes and Hao 1998), Celtic languages in Britain and Ireland (Kandler, Unger and Steele 2010) and may now be happening among Bantu languages in South Africa (Posel and Zeller 2016), all of which have to compete with English as a politically and economically-dominant language.

The Red List of Languages

UNESCO and SIL have both undertaken assessments of the world's languages using a Red List-type of approach to place languages into categories such as *Extinct*, *Critically endangered*, or *Vulnerable* (Lewis and Simons 2010; Moseley 2010). They both used intergenerational transmission as the principal criterion for determining status, which relies on comparing language competence between successive generations, from grandparents to parents to children. Loh and Harmon (2014) used data from the sample of 1500 languages in the ILD to assess their threat status according to the IUCN Red List criteria. This approach provided a direct comparison between the status of languages and that of species and another perspective on the trends shown by the LPI and the ILD, as well as an update to an earlier analysis based on population size only (Sutherland 2003).

One disadvantage of Red Listing is that threat categories are slow to change as a population dwindles. For example, a language with millions of speakers in a gradual decline of 30% in three generations (equivalent to less than 0.5% per year assuming 25-year generations) would remain in the *Vulnerable* category unless either its rate of decline increased to 50% or the number of speakers dropped below a threshold of 2,500. It takes a big change to shift from *Vulnerable* to *Endangered*. The LPI and ILD are better-suited to tracking gradual trends.

IUCN criterion A, the rate of population decline, is applied over three generations and, because human generation times are long, many languages fall into a threat category because of criterion A alone. To be cautious in estimating threat status, therefore, only criteria C (population size *and* rate of decline) and D (population size alone), were used to assess languages: had criterion A been used, languages overall would be considered more threatened than vertebrate species. It is possible that avoiding A was unduly cautious. Language shift is capable of threatening languages spoken by millions. Northern Khmer, for example, with 1.6 million speakers in Thailand in 2006 (Lewis, Simons and Fennig 2016), is losing speakers to Thai as a result of mass media, increased mobility, government policy and political upheavals in Cambodia (Vail 2006). However, criteria C and D alone were used to avoid possible non-genuine changes (criterion B, geographic range, was not used).

Conclusion

The submitted publications present new indicators of the state of biological, linguistic and biocultural diversity, which made a number of advances over previously available indicators, yet further improvements are undoubtedly possible. The most important innovation was to look at population trends across a wide sample of species or languages, as opposed to species or language richness, and then calculate the average trend of either the global set or a particular subset such as a region or taxonomic group. These indicators were the first attempts to show trends at global or regional scale in biological or linguistic diversity in the generic sense of “life on earth”. Population trend indicators are complementary to indicators based on richness or extinction rates, and for this reason a Red List analysis of languages was carried out alongside a population-based analysis. The major advantage of population-based indices over richness, extinction rates or extinction risk is the possibility of showing recent change over a relatively short time-scale. For species, global Red Lists are more comprehensive than the LPI because more species have been assessed than there are species for which population trend data are available. For languages this is not the case, although there remains the problem of identifying non-genuine changes in the language time-series data. An important goal for the LPI has been to increase the coverage of species in the database both biogeographically and taxonomically. For the ILD, it will be hard to resolve non-genuine changes by improving historic population estimates, but future data on speaker numbers will probably have fewer non-genuine changes with respect to the most recent data.

The IBCD is an indicator based on richness data only, but was the first attempt to create a measure of biocultural diversity. The natural next step to improve the IBCD would be to incorporate population trends. This is unlikely to be a possibility for plant species, but may be possible for some birds and mammals using data available from the LPI database. For languages, ethnic groups and possibly religions it might be feasible to incorporate population trend data from *Ethnologue* or national censuses. Another good reason for incorporating population data is that richness alone does not give any information about the distribution or evenness of species or ethnolinguistic diversity within a country or region, and calculating a diversity index in the mathematical sense along the lines of the Simpson/ELF or Shannon indices could be a possibility. These would have to be modified however as they are poor at tracking population trends (Buckland et al. 2005; van Strien, Soldaat and Gregory 2012).

Trends based on populations, be they biological or linguistic, provide no information about genetic diversity. The LPI and ILD could both be improved if they incorporated a measure of the degree of difference between species or languages, such that a decline in population of a particularly unique species or language, such as a monophyletic species or a language isolate, carried more weight than the decline in a species or language from a large family of close relatives. This could be achieved by incorporating phylogenetic data on the evolutionary branch length of species and languages as a genetic diversity weighting. Similarly, the IBCD could incorporate phylogenetic data to show the degree of difference represented in a country’s biocultural diversity. Phylogenetic trees are available for some language families and vertebrate classes, but for many languages measuring evolutionary branch length may prove impossible. Even without data on genetic distinctiveness from phylogenetic trees, however, estimations might be made based on existing taxonomic classification alone.

The spatial and temporal patterns seen in the overall trends in populations of species and languages worldwide reveal interesting global parallels and regional contrasts which help to shed light on both the decline in biological and linguistic diversity and the geographic patterns in the drivers of those trends. While the global LPI and ILD are both driven ultimately by the unsustainable consumption and production of resources, the contrasting regional pattern between them reflects a difference in the immediate drivers of diversity loss. The greatest pressures on biodiversity come from habitat loss and degradation, over-fishing or hunting, and alien invasive species (Millennium Ecosystem Assessment 2005). Pressure on linguistic diversity comes either from one human population displacing another – whether as a result of war, genocide or disease – or, more commonly, from language shift as a dominant language displaces a smaller one within the same population (Nettle and Romaine 2000). This process is analogous in some ways to invasive species: a single alien language expands at the expense of many indigenous ones.

Biodiversity loss has been most rapid in the tropics over the last few decades because that is where the pressures of human population growth, deforestation, habitat conversion, urbanization, pollution and over-exploitation of species have been strongest (Millennium Ecosystem Assessment 2005). Linguistic diversity loss shows a different pattern, where the greatest rate of loss has taken place in those parts of the world colonized and settled by Europeans. The languages which have declined most rapidly since the 1970s are indigenous languages spoken in the Americas and Australia where English, Spanish or Portuguese have become dominant. Interestingly, in Africa, Asia and Europe the decline is not as rapid. However, language shift also takes place within those regions when a national language displaces a minority language, either with or without deliberate government intervention.

This pattern of language loss is consistent with Nettle's description of the decline in global linguistic diversity resulting from the after-shock of the Neolithic Revolution, followed by the after-shock of the Industrial Revolution (Nettle 1999). The Neolithic Revolution saw many phylogenetically diverse languages spoken by thousands of small groups of hunter-gatherers living at low population density displaced by languages spread around the world by the expansion of farming, which then diversified into large, more closely-related language families such as Indo-European, Niger-Congo and Austronesian (Renfrew 1987; Gray and Jordan 2000; Gray and Atkinson 2003). The Neolithic after-shock was the eventual arrival of Indo-European-speaking farming peoples in the Americas and Australia, and the consequent loss of indigenous languages from genocide, disease, deliberate policy and subsequently language shift. The Industrial Revolution resulted in enormous changes in lifestyles and livelihoods to populations formerly dependent on agriculture, and its after-shock spread around the world over the last 250 years, bringing about a steady decline in linguistic diversity in the developing regions the world, including in Europe, where indigenous populations were not overwhelmed but gradually shifted to the socially and politically dominant languages of a more developed economy (Nettle 1999; Nettle and Romaine 2000).

Acknowledgements

Many thanks are due to David Harmon and the other co-authors of the papers presented here for their agreement and encouragement to submit them for a PhD by published works. I am grateful to the School of Anthropology and Conservation and the University of Kent for giving me the opportunity to do so. Many colleagues at the Global Footprint Network, Terralingua, WWF International, WWF Netherlands, and the Zoological Society of London – some of whom are also co-authors – have provided wonderful collaborations and interesting conversations. My supervisors Dr Rajindra Puri and Dr Peter Bennett provided invaluable help and clear insights. My examiners Professors Patricia Howard and Kate Jones provided very useful comments on this manuscript. Finally I would like to thank my family for their total support and encouragement.

Appendix: Methods used in *Biocultural Diversity: Threatened species, endangered languages*

Loh and Harmon (2014) compared trends in the Living Planet Index (LPI) from 1970 to 2007 with trends in the Index of Linguistic Diversity (ILD) from 1970 to 2008 at the global and regional scales. The LPI and ILD are based on similar methods for calculating average trends across a large selection of species or languages, and are described in detail in three of the present papers: Loh *et al*. (2005), Collen *et al*. (2009) and Harmon and Loh (2010).

Loh and Harmon (2014) also compared the status of global biodiversity and linguistic diversity using the IUCN Red List categories. Red Listing provides a means of using a very different methodology from the LPI and ILD by which to compare the status of the world's languages with that of different taxonomic groups of species. The Red List is a catalogue of species that have been assessed and categorized as *extinct* (EX), *extinct in the wild* (EW), *critically endangered* (CR), *endangered* (EN), *vulnerable* (VU), *near threatened* (NT) or *least concern* (LC) according to their risk of extinction. Species assessed as CR, EN or VU are considered to be threatened. Where insufficient information is known to make an assessment, a species is categorized as *data deficient* (DD).

A set of well-defined criteria are used to assess and place a species in a Red List category (IUCN 2012). Criteria A1-A4 relate to the rate of decline in a species' population. Criteria B1 and B2 refer to reduction or fragmentation of a species' geographic range. Criteria C1 and C2 relate to a both a species' population size and its rate of decline. Criterion D relates to population size alone, regardless of its rate of change. Some of those criteria can also be applied to languages. Criteria A2/A4, C1 and D were applied to data on the number of mother tongue speakers of a language (in place of species population data) to assess its extinction risk (see Table 1). This method ignores those speakers who can speak a language that is not their mother tongue.

IUCN Criteria			Population Decline						
			A			**C1**			**D**
			≥80% in three generations	≥50% in three generations	≥30% in three generations	≥25% in one generation	≥20% in two generations	≥10% in three generations	Any
Population Size	**A**	Any	CR	EN	VU				
	B	not used							
	C1	1-249				CR			
		250-2499					EN		
		2500-10000						VU	
	C2	not used							
	D	0							EX
		1-49							CR
		50-249							EN
		250-1000							VU

Table A1: IUCN Red List criteria (IUCN 2012) used in the assessment of languages. EX = extinct, CR = critically endangered, EN = endangered, VU = vulnerable. For criterion C, both size and decline criteria must be met. If a language did not meet any criteria it was categorized as least concern/near threatened (LC/NT); if there were insufficient data to assess a language it was categorized as data deficient (DD). Generation time was taken to be 25 years.

Criterion A2 refers to a population decline of 80% or more in three generations which has already occurred; criterion A4 refers to a population decline that began in the past and is projected to decline by at least 80% within three generations. The human generation time was taken to be 25 years. Criteria A1 and A3 were not used as they refer to populations where the causes of decline are known, have ceased and are reversible (A1) or where population decline is not yet known to have occurred but is suspected or projected in future (A3). Criterion B was not used as there are insufficient data available on trends in the geographic ranges of languages. Criterion C2 was not used as it requires additional data on either the size of subpopulations (C2a) or the number of adults in a population (C2b).

Criteria A2/A4, C1 and D were used to assess a sample of 1,500 languages, the same sample that was used to generate the ILD (see Harmon and Loh 2010). For each language, the number of mother tongue speakers and the year in which the number was recorded were entered into an Excel spreadsheet. For those languages with two or more data points, ie. speaker numbers from two or more years (not necessarily consecutive), the average annual rate of change was calculated between the most recent pair of years.

$$r = \left(\frac{p_2}{p_1}\right)^{\left(\frac{1}{y_2 - y_1}\right)}$$

Where r is the annual average rate of change, y_1 is the first year, y_2 is the second year, p_1 is the number of speakers in y_1 and p_2 is the number of speakers in y_2.

If a language's annual rate of decline was greater than 10% per year ($r < 0.9$), and the initial number of speakers was 1000 or more ($p_1 > 999$), then r was not used in the assessment of that language on the grounds that speaker numbers are unlikely to decline that fast in a population of that size, and therefore there was a possibility that the data may be anomalous. In such cases, only criterion D was used.

For those languages with two or more data points, criteria A2/A4, C1 and D were applied. For those languages with a single data point, criterion D only was applied.

Under criterion A2/A4, a language was assessed to be CR if $r < 0.9788$, which would lead to a decline of at least 80% within 75 years or three generations. It was assessed to be EN if $r < 0.9908$, which would lead to a decline of at least 50% in 75 years, or VU if $r < 0.9953$, which would lead to a decline of at least 30% in 75 years.

Under criterion C1, a language was assessed to be CR if $p_2 < 250$ and $r < 0.9886$, which would lead to a decline of at least 75% within 25 years. It was assessed to be EN if $p_2 < 2500$ and $r < 0.9955$, which would lead to a decline of at least 20% in 50 years, or VU if $p_2 < 10000$ and $r < 0.9986$, which would lead to a decline of at least 10% in 75 years.

Under criterion D, a language was assessed to be CR if the most recently recorded number of speakers was less than 50, $p < 50$. It was assessed to be EN if $p < 250$ or VU if $p < 1000$. If the most recently recorded number of speakers of a language was zero, the language was assessed to be EX.

Where the criteria gave different assessments of threat status for the same language, for example EN under C1 and VU under D, the higher or more severe status was assigned to that language. Where a language did not meet any of the criteria for EX, CR, EN or VU status, it was assessed to be least concern or near threatened (LC/NT). If there were no population data available for a language, it was evaluated as data deficient (DD). Because there is a degree of uncertainty in the reliability of some earlier estimates of speaker numbers, meeting the A criterion alone was not considered to be sufficient for a language be reliably assessed as threatened. Therefore a language was assessed to be threatened only if it met either or both criteria C1 and D.

References

Alesina, A., et al. (2003). Fractionalization. *Journal of Economic Growth* [Online], 8(2), 155-194.

Anderson, T. and Poppel, B. (2002). Living Conditions in the Arctic. *Social Indicators Research*, 58, 191-216.

Arinaitwe, H., Pomeroy, D. and Tushabe, H. (2000). *The State of Uganda's Biodiversity 2000*. Kampala, Uganda: Makerere University Institute of Environment and Natural Resources.

Armenteras, D., et al. (2012). Biodiversity. In: United Nations Environment Programme ed. *Global Environment Outlook GEO-5: Environment for the Future we Want*. Progress Press Ltd, pp. 133-166.

Ash, N., et al. (2007). Biodiversity. In: United Nation Environment Programme ed. *Global Environment Outlook GEO-4: Environment for Development*. Valletta, Malta: Progress Press Ltd, pp. 157-192.

Aslaksen, I. and Garnåsjordet, P. A. (2012). The Norwegian Nature Index. *Norsk Geografisk Tidsskrift - Norwegian Journal of Geography* [Online], 66(5), 239-240.

Baer, A. (1989). Maintaining Biocultural Diversity. *Conservation Biology* [Online], 3(1), 97-98. Available from: http://www.jstor.org/stable/2385998.

Baillie, J. E. M., et al. ed. (2010). *Evolution Lost: Status and Trends of the World's Vertebrates*. [Online]. Zoological Society of London.

Barrett, D. B., Kurian, G. T. and Johnson, T. M. eds. (2001). *World Christian Encyclopedia*. 2nd edn. Oxford, UK: Oxford University Press.

Bernard, H. (1992). Preserving Language Diversity. *Human Organization* [Online], 51(1), 82.

Beyer, P. (2010). Religious Diversity and Globalization. In: Meister, C. ed. *The Oxford Handbook of Religious Diversity*. Oxford, UK: Oxford University Press.

Bininda-Emonds, O. R. P., et al. (2008). The delayed rise of present- day mammals. *Nature* [Online], 456(7219), 274.

Biodiversity Indicators Partnership (2013a). *Aichi Targets Passport*. Cambridge, UK: UNEP-WCMC.

Biodiversity Indicators Partnership (2013b). *Index of Linguistic Diversity* [Online]. Available from: http://www.bipindicators.net/ild [Accessed 5/24/2016].

BirdLife International (2013). *State of the World's Birds: Indicators for our Changing World*. Cambridge, UK: BirdLife International.

Bossert, W., D' Ambrosio, C. and La Ferrara, E. (2011). A Generalized Index of Fractionalization. *Economica* [Online], 78(312), 723-750.

Bouma, G. D. and Ling, R. (2011). Religious Diversity. In: Clarke, P. B. ed. *The Oxford Handbook of the Sociology of Religion*. [Online]. Oxford, UK: Oxford University Press.

Braun, D. M. (2011). *Language Diversity Index Tracks Global Loss of Mother Tongues | National Geographic (Blogs)* [Online]. Last updated: 1 March 2011. Available from: http://voices.nationalgeographic.com/2011/03/01/language_diversity_index_tracks_global_loss_of_mother_tongues/ [Accessed 3/17/2016].

Brosius, J. P. and Hitchner, S. L. (2010). Cultural diversity and conservation. *International Social Science Journal* [Online], 61(199), 141-168.

Buckland, S. T., et al. (2005). Monitoring change in biodiversity through composite indices. *Philosophical Transactions of the Royal Society B: Biological Sciences* [Online], 360(1454), 243-254.

Butchart, S. H. M., et al. (2005). Using Red List Indices to measure progress towards the 2010 target and beyond. *Philosophical Transactions of the Royal Society of London.Series B, Biological Sciences* [Online], 360(1454), 255.

Butchart, S., et al. (2004). Measuring global trends in the status of biodiversity: red list indices for birds. *PLoS Biol* [Online], 2(12), e383.

Butchart, S. H. M., et al. (2010). Global biodiversity: indicators of recent declines. *Science (New York, N.Y.)* [Online], 328(5982), 1164-1168.

CBD Secretariat (2014). *Global Biodiversity Outlook 4*. Montréal, Canada: Secretariat of the Convention on Biological Diversity.

CBD Secretariat (2010). *Global Biodiversity Outlook 3*. Montréal, Canada: Secretariat Convention on Biological Diversity.

CBD Secretariat (2006). *Global Biodiversity Outlook 2*. Montreal, Canada: Secretariat of the Convention on Biological Diversity.

CBD Secretariat (2004). *Strategic Plan: Future Evaluation of Progress*. UNEP/CBD/COP/7/L.27 edn. Montreal, Canada: Convention on Biological Diversity.

CBD Secretariat (1994). *Convention on Biological Diversity: Text and Annexes*. Geneva, Switzerland: UNEP.

Cocks, M. (2010). What is Biocultural Diversity? A theoretical review. In: Daniel G. Bates, J. T. ed. *Human Ecology: Contemporary Research and Practice*. [Online]. DE: Springer Verlag, pp. 67-77.

Cocks, M. (2006). Biocultural Diversity: Moving Beyond the Realm of ‘ Indigenous’ and ‘Local’ People. *Human Ecology; an Interdisciplinary Journal* [Online], 34(2), 185-200.

Collen, B., et al. (2009). Monitoring change in vertebrate abundance: the Living Planet Index. *Conservation Biology* [Online], 23(2), 317-327.

Collin, R. O. (2010). Ethnologue. *Ethnopolitics; Formerly Global Review of Ethnopolitics* [Online], 9(3-4), 425-432.

Craigie, I. D., et al. (2010). Large mammal population declines in Africa’s protected areas. *Biological Conservation* [Online], 143(9), 2221-2228.

Currie, T. E. and Mace, R. (2012). The evolution of ethnolinguistic diversity. *Advances in Complex Systems* [Online], 15(1-2).

Darwin, C. (1871). *The Descent of Man, and Selection in Relation to Sex*. [Online]. London: J. Murray.

Deinet, S., et al. (2010). *The Living Planet Index for Estuarine Species*. London, UK: ZSL.

Deinet, S., et al. (2013). *Wildlife Comeback in Europe: The Recovery of Selected Mammal and Bird Species. Final Report to Rewilding Europe.* . London, UK: ZSL.

Dixon, M. J. R., et al. (2016). Tracking global change in ecosystem area: The Wetland Extent Trends index. *Biological Conservation* [Online], 193, 27-35.

Elands, B., et al. (2015). Policy interpretations and manifestation of biocultural diversity in urbanized Europe: conservation of lived biodiversity. *Biodiversity and Conservation* [Online], 24(13), 3347-3366.

European Bird Census Council (2016). *What is Pan-European Common Bird Monitoring Scheme?* [Online]. Available from: http://www.ebcc.info/pecbm.html [Accessed 9/2/2016].

European Bird Census Council *EBCC Home Page* [Online]. Available from: http://www.ebcc.info/index.php?ID=28 [Accessed 4/3/2016].

European Environment Agency (2013). *The European Grassland Butterfly Indicator: 1990–2011*. Luxembourg: Publications Office of the European Union.

Fearon, J. D. (2003). Ethnic and cultural diversity by country. *Journal of Economic Growth* [Online], 8(2), 195-222.

Fedderke, J., Luiz, J. and Kadt, R. (2008). Using fractionalization indexes: deriving methodological principles for growth studies from time series evidence. *Social Indicators Research; an International and Interdisciplinary Journal for Quality-of-Life Measurement* [Online], 85(2), 257-278.

Fincher, C. L. and Thornhill, R. (2008). A parasite-driven wedge: infectious diseases may explain language and other biodiversity. *Oikos* [Online], 117(9), 1289-1297.

Galewski, T., et al. (2011). Long-term trends in the abundance of Mediterranean wetland vertebrates: From global recovery to localized declines. *Biological Conservation* [Online], 144(5), 1392-1399.

Gaston, K. J. (1996a). Species richness: measure and measurement. In: Gaston, K. J. ed. *Biodiversity: A Biology of Numbers and Difference*. Oxford, UK: Blackwell, pp. 77-113.

Gaston, K. J. (1996b). What is biodiversity?. In: Gaston, K. J. ed. *Biodiversity: A Biology of Numbers and Difference*. Oxford, UK: Blackwell, pp. 1-9.

Gavin, M. C., et al. (2013). Toward a Mechanistic Understanding of Linguistic Diversity. *Bioscience* [Online], 63(7), 524-535.

Gavin, M. C. and Sibanda, N. (2012). The island biogeography of languages. *Global Ecology and Biogeography* [Online], 21(10), 958-967.

Gordon, R. G. ed. (2005). *Ethnologue : Languages of the World*. [Online]. 15th edn. Dallas, Texas: SIL International.

Gorenflo, L. J., et al. (2012). Co-occurrence of linguistic and biological diversity in biodiversity hotspots and high biodiversity wilderness areas. *Proceedings of the National Academy of Sciences of the United States of America* [Online], 109(21), 8032-8037.

Gray, R. D. and Atkinson, Q. D. (2003). Language-tree divergence times support the Anatolian theory of Indo- European origin. *Nature* [Online], 426(6965), 435.

Gray, R. D., Atkinson, Q. D. and Greenhill, S. J. (2011). Language evolution and human history: what a difference a date makes. *Philosophical Transactions of the Royal Society of London.Series B, Biological Sciences* [Online], 366(1567), 1090.

Gray, R. D. and Jordan, F. M. (2000). Language trees support the express-train sequence of Austronesian expansion. *Nature* [Online], 405(6790), 1052.

Greenberg, J. H. (1987). *Language in the Americas*. [Online]. Stanford University Press.

Greenberg, J. H. (1963). *The Languages of Africa*. [Online]. Vol. 25. Indiana Univ.

Greenberg, J. H. (1996). In Defense of Amerind. *International Journal of American Linguistics* [Online], 62(2), 131-164.

Gregory, R. D., et al. (2005). Developing indicators for European birds. *Philosophical Transactions of the Royal Society of London.Series B, Biological Sciences* [Online], 360(1454), 269-288.

Grimes, B. F. ed. (2000). *Ethnologue. Languages of the World*. 14th edn. Dalla, USA: SIL.

Grimes, B. F. ed. (1974). *Ethnologue*. 8th edn. Huntingdon Beach, California, USA: Wycliffe Bible Translators.

Groombridge, B. and Jenkins, M. D. (2002). *Wolrd Atlas of Biodiversity: Earth's Living Resources in the 21st Century*. Berkeley, USA: University of California Press.

Groombridge, B. ed. (1992). *Global Biodiversity: Status of the Earth's Living Resources*. [Online]. London: Chapman & Hall.

Harmon, D. (1996). Losing Species, Losing Languages: Connections between Biological and Linguistic Diversity. *Southwest Journal of Linguistics*, 15(1-2), 89-108.

Harmon, D. (2002). *In Light of our Differences: How Diversity in Nature and Culture Makes Us Human*. Washington and London: Smithsonian Institution Press.

Harmon, D. (2001). On the meaning and moral imperative of diversity. In: Maffi, L. ed. *On Biocultural Diversity*. Washingon, USA: Smithsonian Institution Press, pp. 53-70.

Harmon, D. and Loh, J. (2010). The index of linguistic diversity: a new quantitative measure of trends in the status of the world's languages. *Language Documentation & Conservation* [Online], 4, 97-151.

Harmon, D. and Loh, J. (2009). *Global Indicators of the Status and Trends of Linguistic Diversity and Traditional Knowledge. Indicator #1: The Index of Linguistic Diversity (ILD). Technical Report to the Christensen Fund*. Palo Alto, California, USA: The Christensen Fund.

Harmon, D. and Loh, J. (2004a). *Draft Framework for an Index of Biocultural Diversity*. [Online]. Washington, DC: Terralingua.

Harmon, D. and Loh, J. (2004b). The IBCD: a measure of the world's biocultural diversity. *Policy Matters* [Online], 13, 271-280.

Harmon, D., Woodley, E. and Loh, J. (2010). Measuring status and trends in biological and cultural diversity. *Nature and Culture.Rebuilding a Lost Connection, Eds.S.Pilgirm and JN Pretty, Earthscan* [Online], 41-64.

Harper, J. L. and Hawksworth, D. L. (1994). Biodiversity: measurement and estimation. *Philosophical Transactions of the Royal Society of London.Series B, Biological Sciences* [Online], 345(1311), 5.

Heywood, V. H. and Watson, R. T. (1995). *Global Biodiversity Assessment*. [Online]. Cambridge: United Nations Environment Programme: Cambridge University Press.

Hsu, B., et al. (2008). Estimating the Religious Composition of All Nations: An Empirical Assessment of the World Christian Database. *Journal for the Scientific Study of Religion* [Online], 47(4), 678-693.

Isaac, N. J. B., et al. (2007). Mammals on the EDGE: Conservation Priorities Based on Threat and Phylogeny (Mammals on the EDGE). *PLoS ONE* [Online], 2(3), e296.

ISE (n.d.). *Declaration of Belém - International Society of Ethnobiology* [Online]. Available from: http://www.ethnobiology.net/declaration-of-belem/ [Accessed 7/25/2016].

IUCN (2013). *The IUCN Red List of Threatened Species. Version 2013.2* [Online]. Available from: www.iucnredlist.org

IUCN (2012). *IUCN Red List Categories and Criteria: Version 3.1*. 2nd edn. Gland, Switzerland and Cambridge, UK: IUCN.

IUCN (1996). *1996 IUCN Red List of Threatened Animals*. Gland, Switzerland: IUCN - World Conservation Union.

IUCN and WCMC (1998). *1997 IUCN Red List of Threatened Plants*. [Online]. Gland, Switzerland ; Cambridge, UK: IUCN - The World Conservation Union.

Jackson, T. (2011). *Prosperity without Growth: Economics for a Finite Planet*. [Online]. Paperback edition edn. London: Earthscan.

Jenkins, M., Kapos, V. and Loh, J. (2004). *Rising to the Biodiversity Challenge. Discussion Paper for CBD CoP7*. Cambridge, UK: UNEP-World Conservation Monitoring Centre.

Kandler, A., Unger, R. and Steele, J. (2010). Language shift, bilingualism and the future of Britain's Celtic languages. *Philosophical Transactions of the Royal Society of London.Series B, Biological Sciences* [Online], 365(1559), 3855.

Krauss, M. E. (2007). Keynote-mass language extinction and documentation: The race against time. *The Vanishing Languages of the Pacific Rim* [Online], 3-24.

Krauss, M. (1992). The world's languages in crisis. *Language* [Online], 68(1), 4-10. Available from: http://www.jstor.org/stable/416368.

Lasimbang, J. (2008). Participatory Arrangements and the Future of Traditional Occupations in Relation to Contemporary Policy Framework: A Case Study of the kadazans, Sabah, Malaysia. In: Stankovitch, M. ed. *Indicators Relevant for Indigenous People: A Resource Book*. Baguio City, Philippines: Tebtebba Foundation, pp. 387-415.

Latham, J., et al. (2010). *The Living Planet Index for Migratory Species: An Index of Change in Population Abundance*. London, UK: Zoological Society of London.

Leakey, R. E. (1996). *The Sixth Extinction: Biodiversity and its Survival*. [Online]. London: London : Phoenix.

Lewis, M. P. and Simons, G. F. (2010). Assessing endangerment: expanding Fishman’s GIDS. *Revue Roumaine De Linguistique* [Online], 2, 103-119.

Lewis, M. P. ed. (2009). *Ethnologue: Languages of the World*. [Online]. Vol. 16. 16th edn. Dallas, TX: SIL international Dallas, TX.

Lewis, M. P., Simons, G. F. and Fennig, C. D. eds. (2016). *Ethnologue: Languages of the World*. 19th edn. Dallas, Texas: SIL International.

Loh, J. and Harmon, D. (2005). A global index of biocultural diversity. *Ecological Indicators* [Online], 5(3), 231-241.

Loh, J. (2002). *Living Planet Report 2002*. [Online]. Gland, Switzerland: WWF Worldwide Fund for Nature.

Loh, J. (2000). *Living Planet Report 2000*. [Online]. Gland, Switzerland: WWF Worldwide Fund for Nature.

Loh, J. and Goldfinger, S. (2006). *Living Planet Report 2006*. [Online]. Gland, Switzerland: WWF Worldwide Fund for Nature.

Loh, J., et al. (2005). The Living Planet Index: using species population time series to track trends in biodiversity. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* [Online], 360(1454), 289-295.

Loh, J. and Harmon, D. (2014). *Biocultural Diversity: Threatened Species, Endangered Languages*. [Online]. Zeist, The Netherlands: WWF-Netherlands.

Loh, J., et al. (1999). *Living Planet Report 1999*. [Online]. Gland, Switzerland: WWF Worldwide Fund for Nature.

Loh, J., et al. (1998). *Living Planet Report 1998*. [Online]. Gland, Switzerland: WWF Worldwide Fund for Nature.

Loh, J. and Wackernagel, M. (2004). *Living Planet Report 2004*. [Online]. Gland, Switzerland: WWF - World Wide Fund for Nature.

Mace, R. and Pagel, M. (1995). A latitudinal gradient in the density of human languages in North America. *Proceedings of the Royal Society of London B: Biological Sciences* [Online], 261(1360), 117-121.

Maffi, L. (2007). Biocultural Diversity and Sustainability. In: Pretty, J. N. ed. *The SAGE Handbook of Environment and Society*. [Online]. Los Angeles, USA: SAGE, pp. 267-277.

Maffi, L. (2001a). Introduction: on the interdependence of biological and cultural diversity. In: Maffi, L. ed. *On Biocultural Diversity: Linking Language, Knowledge and the Environment*. Washington, DC, USA: Smithsonian Institution Press, pp. 1-50.

Maffi, L. (2001b). *On Biocultural Diversity : Linking Language, Knowledge, and the Environment*. [Online]. Washington, D.C.] ; London: Washington, D.C. ; London : Smithsonian Institution Press.

Magurran, A. E. (2010). Q&A: What is biodiversity?. *BMC Biology* [Online], 8.

Magurran, A. E. (2004). *Measuring Biological Diversity*. [Online]. Oxford, UK: Blackwell.

McCormack, J. E., et al. (2013). A Phylogeny of Birds Based on Over 1,500 Loci Collected by Target Enrichment and High-Throughput Sequencing. *PLoS ONE* [Online], 8(1).

McRae, L., Freeman, R. and Deinet, S. (2014). The Living Planet Index. In: McLellan, R. ed. *Living Planet Report 2014: Species and Spaces, People and Places*. Gland, Switzerland: WWF International, pp. 16-29.

McRae, L., et al. (2007). A Living Planet Index for Canada. [Online].

McRae, L., et al. (2012). The Arctic Species Trend Index: Using vertebrate population trends to monitor the health of a rapidly changing ecosystem. *Biodiversity* [Online], 13(3-4), 144-156.

McRae, L., et al. (2010). *Arctic Species Trend Index 2010. Tracking Trends in Arctic Wildlife.* [Online]. CAFF International Secretariat.

McVean, G. (2002). Chromosomes. In: Pagel, M. ed. *Encyclopedia of Evolution Vol.1*. Oxford, UK: Oxford University Press, pp. 151-154.

Millennium Ecosystem Assessment (2005). *Ecosystems and Human Well-being: Biodiversity Synthesis*. [Online]. Vol. 5. Washington, DC: World Resources Institute.

Milligan, H., et al. (2014). *Protecting Species: Status and Trends of the Earth's Protected Areas. Preliminary Report*. London, UK: Zoological Society of London.

Mitchell, S., Loh, J. and Goldfinger, S. (2007). *Canadian Living Planet Report 2007*. Toronto, Canada: WWF-Canada.

Mittermeier, R. A., de Macedo, R. and Luscombe, A. (1975). A Woolly Monkey Rediscovered in Peru. *Oryx;* [Online], 13(1), 41-46.

Moore, J. L., et al. (2002). The distribution of cultural and biological diversity in Africa. *Proceedings.Biological Sciences / the Royal Society* [Online], 269(1501), 1645.

Moseley, C. ed. (2010). *Atlas of the World's Languages in Danger*. Third edn. Paris, France: UNESCO.

Nettle, D. (1999). *Linguistic Diversity*. [Online]. Oxford University Press.

Nettle, D. (1998). Explaining global patterns of language diversity. *Journal of Anthropological Archaeology* [Online], 17(4), 354-374.

Nettle, D. and Romaine, S. (2000). *Vanishing Voices: The Extinction of the World's Languages*. Oxford University Press.

Newing, H. (2010). Bridging the Gap: Interdisciplinarity, Biocultural Diversity and Conservation. In: Pilgrim, S. and Pretty, J. eds. *Nature and Culture: Rebuilding Lost Connections*. London, UK: earthscan, pp. 23-40.

Nybø, S., Certain, G. and Skarpaas, O. (2011). *The Norwegian Nature Index 2010*. Trondheim, Norway: The Norwegian Directorate for Nature Management.

OECD Environment Directorate (2001). *Key Environmental Indicators*. Paris, France: OECD.

OED Online (2016a). *Biodiversity, N.* [Online]. Last updated: June 2016. Available from: http://www.oed.com/

OED Online (2016b). *Culture, N.* [Online]. Last updated: June, 2016. Available from: http://www.oed.com/

Office for National Statistics (2015). *Consumer Price Indices: A Brief Guide*. Office for National Statistics.

Oxford Dictionaries (n.d.). *Biodiversity* [Online]. Available from: http://www.oxforddictionaries.com/definition/english/biodiversity

Pagel, M., et al. (2013). Ultraconserved words point to deep language ancestry across Eurasia. *Proceedings of the National Academy of Sciences of the United States of America* [Online], 110(21), 8471.

Patsiurko, N., Campbell, J. L. and Hall, J. A. (2012). Measuring cultural diversity: ethnic, linguistic and religious fractionalization in the OECD. *Ethnic and Racial Studies* [Online], 35(2), 195-217.

Patuxent Wildlife Research Center and Canadian Wildlife Service (2001). *North American Breeding Bird Survey* [Online]. Last updated: 10/31/01. Available from: https://www.pwrc.usgs.gov/bbs/

Pereira, H. M. and David Cooper, H. (2006). Towards the global monitoring of biodiversity change. *Trends in Ecology & Evolution* [Online], 21(3), 123-129.

Pinker, S. (1995). *The Language Instinct: The New Science of Language and Mind*. [Online]. Vol. 7529. Penguin UK.

Pomeroy, D., Lutaaya, B. and Tushabe, H. eds. (2006). *The State of Uganda's Biodiversity 2006*. Kampala, Uganda: Makerere University Institute of Environment and Natural Resources.

Pomeroy, D. and Mwima, P. (2002). *The State of Uganda's Biodiversity 2002*. Kampala, Uganda: Makerere University Institute of Envionment and Natural Resources.

Pomeroy, D. and Tushabe, H. eds. (2008). *The State of Uganda's Biodiversity 2008*. Kampala, Uganda: Makerere University Institute of Environment and Natural Resources.

Pomeroy, D. and Tushabe, H. (2004). *The State of Uganda's Biodiversity 2004*. Kampala, Uganda: Makerere University Institute of Environment and Natural Resources.

Popper, K. R. (1959). The logic of scientific discovery. *London: Hutchinson* [Online].

Portes, A. and Hao, L. (1998). E Pluribus Unum: Bilingualism and Loss of Language in the Second Generation. *Sociology of Education* [Online], 71(4), 269-294.

Posel, D. and Zeller, J. (2016). Language shift or increased bilingualism in South Africa: evidence from census data. *Journal of Multilingual and Multicultural Development* [Online], 37(4), 357-370.

Posner, D. N. (2004). Measuring Ethnic Fractionalization in Africa. *American Journal of Political Science* [Online], 48(4), 849-863.

Prescott-Allen, R. (2001). *The Wellbeing of Nations: A Country-by-Country Index of Quality of Life and the Environment*. [Online]. Washington, DC: Island Press.

Priddel, D., et al. (2003). Rediscovery of the 'extinct' Lord Howe Island stick-insect (Dryococelus australis (Montrouzier)) (Phasmatodea) and recommendations for its conservation. *Biodiversity & Conservation* [Online], 12(7), 1391-1403.

Proença, V., et al. (2016). Global biodiversity monitoring: From data sources to Essential Biodiversity Variables. *Biological Conservation* [Online], http://dx.doi.org/10.1016/j.biocon.2016.07.014.

Prum, R. O., et al. (2015). A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. *Nature* [Online].

Renfrew, C. (1987). *Archaeology and Language: The Puzzle of Indo-European Origins*. [Online]. London, UK: Penguin Books.

Reyhner, J. (2007). Rationale and Needs for Stabilizing Indigenous Languages. In: Cantoni, G. ed. *Stabilizing Indigenous Languages*. Revised edn. Flagstaff, Arizona, USA: Northern Arizona University, pp. 3-14.

Richerson, P. J. and Boyd, R. (2005). *Not by Genes Alone: How Culture Transformed Human Evolution*. [Online]. Chicago: University of Chicago Press.

Skyhawk, S. (2012). *Why should we Keep Tribal Languages Alive?* [Online]. Available from: http://indiancountrytodaymedianetwork.com/2012/04/06/why-should-we-keep-tribal-languages-alive-99182 [Accessed 7/12/2016].

Smith, F. D. M., et al. (1993). Estimating extinction rates. *Nature* [Online], 364(6437), 494.

Statistics New Zealand (2008). Towards a Maori Statistics Framework: A discussion document. In: Stankovitch, M. ed. *Indicators Relevant for Indigenous Peoples: A Resource Book*. Baguio City, Philippines: Tebtebba Foundation, pp. 251-272.

Stattersfield, A., Bennun, L. and Jenkins, M. (2008). *State of the World's Birds: Indicators for our Changing World*. Cambridge, UK: Birdlife International.

Sturm, J. E. and De Haan, J. (2015). *Income inequality, capitalism, and ethnolinguistic fractionalization*. [Online]. Vol. 105., pp. 593-597.

Sutherland, W. J. (2003). Parallel extinction risk and global distribution of languages and species. *Nature* [Online], 423(6937), 276.

Terralingua (n.d.). *Terralingua – Protecting Biocultural Diversity* [Online]. Available from: http://terralingua.org/ [Accessed 8/29/2016].

Thomas, C. D., et al. (2004). Extinction risk from climate change. *Nature* [Online], 427(6970), 145.

Tittensor, D. P., et al. (2014). A mid-term analysis of progress toward international biodiversity targets. *Science (New York, N.Y.)* [Online], 346(6206), 241-244.

UNDP (2015). *Human Development Report 2015*. New York, USA: United Nations Development Programme.

UNEP (1999). *GEO-2000: UNEP's Millennium Report on the Environment*. [Online]. London: Earthscan.

UNEP-WCMC (2016a). *The State of Biodiversity in Africa: A Mid-Term Review of Progress Towards the Aichi Biodiversity Targets.* Cambridge, UK: UNEP-WCMC.

UNEP-WCMC (2016b). *The State of Biodiversity in Asia and the Pacific: A Mid-Term Review of Progress Towards the Aichi Biodiversity Targets.* Cambridge, UK: UNEP-WCMC.

UNEP-WCMC (2016c). *The State of Biodiversity in Latin America and the Caribbean: A Mid-Term Review of Progress Towards the Aichi Biodiversity Targets*. Cambridge, UK: UNEP-WCMC.

UNESCO and CBD Secretariat (2014). *Florence Declaration on the Links between Biological and Cultural Diversity*. Florence, Italy: UNESCO; SCBD.

Vail, P. (2006). Can a language of a million speakers be endangered? Language shift and apathy among Northern Khmer speakers in Thailand. *International Journal of the Sociology of Language* [Online], 178, 135-147.

van Strien, A. J., Soldaat, L. L. and Gregory, R. D. (2012). Desirable mathematical properties of indicators for biodiversity change. *Ecological Indicators* [Online], 14(1), 202-208.

Vidal, J. (2014). As forests are cleared and species vanish, there's one other loss: a world of languages. , .

Wilson, E. O. (1994). *The Diversity of Life*. [Online]. Harmondsworth: Harmondsworth : Penguin Books.

World Resources Institute (1996). *World Resources 1996-97: The Urban Environment*. [Online]. Vol. 9. Oxford, UK: OUP.

WWF Norway (2005). *Naturindeks for Norge 2005: Uftor Bakke Med Norsk Natur*. Oslo, Norway: WWF Norway.

Wycliffe Bible Translators (1951). *Translator’s Ethnologue for Intercessors,Translators, Missionaries, and Mission Councils*. Berwick, Victoria, Australia: Wycliffe Bible Translators.

Zent, S. and Maffi, L. (2007). *Final Report on Indicator no 2: Vitality Index for Traditional Ecological Knowledge (VITEK)*. Salt Spring Island, Canada: Terralingua.

The Living Planet Index: using species population time series to track trends in biodiversity

Jonathan Loh, Rhys E. Green, Taylor Ricketts, John Lamoreux, Martin Jenkins, Valerie Kapos and Jorgen Randers

Philosophical Transactions of the Royal Society
(2005) **360**, 289-295

PHILOSOPHICAL
TRANSACTIONS
OF
THE ROYAL
SOCIETY
B

Phil. Trans. R. Soc. B (2005) **360**, 289–295
doi:10.1098/rstb.2004.1584

The Living Planet Index: using species population time series to track trends in biodiversity

Jonathan Loh[1,*], Rhys E. Green[2], Taylor Ricketts[3], John Lamoreux[3], Martin Jenkins[4], Valerie Kapos[4] and Jorgen Randers[5]

[1] *WWF International, Avenue du Mont-Blanc CH-1196, Gland, Switzerland*
[2] *RSPB and Conservation Biology Group, Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK*
[3] *Conservation Science Program, WWF-US, 1250 Twenty-fourth Street, NW, Washington DC 20037, USA*
[4] *UNEP-WCMC, Huntingdon Road, Cambridge CB3 0DL, UK*
[5] *Norwegian School of Management, Elias Smiths vei 15, Box 580 N-1302, Sandvika, Norway*

The Living Planet Index was developed to measure the changing state of the world's biodiversity over time. It uses time-series data to calculate average rates of change in a large number of populations of terrestrial, freshwater and marine vertebrate species. The dataset contains about 3000 population time series for over 1100 species. Two methods of calculating the index are outlined: the chain method and a method based on linear modelling of log-transformed data. The dataset is analysed to compare the relative representation of biogeographic realms, ecoregional biomes, threat status and taxonomic groups among species contributing to the index.

The two methods show very similar results: terrestrial species declined on average by 25% from 1970 to 2000. Birds and mammals are over-represented in comparison with other vertebrate classes, and temperate species are over-represented compared with tropical species, but there is little difference in representation between threatened and non-threatened species. Some of the problems arising from over-representation are reduced by the way in which the index is calculated. It may be possible to reduce this further by post-stratification and weighting, but new information would first need to be collected for data-poor classes, realms and biomes.

Keywords: Living Planet Index; biodiversity indicator; species population trends; representation; 2010 target

1. INTRODUCTION

In its plan of implementation, the 2002 World Summit on Sustainable Development endorsed the Hague Ministerial Declaration of the Sixth Conference of the Parties to the Convention on Biological Diversity (CBD) that committed them 'to achieve by 2010 a significant reduction of the current rate of biodiversity loss at the global, regional and national level as a contribution to poverty alleviation and to the benefit of all life on earth' (CBD 2000, p. 319). At present, there is no system in place to measure the progress towards this objective by standardized, regularly repeated measurements of the state of all important biomes and their biota both at global and regional levels (Green *et al.* 2005; Kapos *et al.* 2004), although good biodiversity indicator programmes exist in some countries, particularly in the UK and the Netherlands, and significant advances have been made at the pan-European level (Gregory *et al.* 2005; de Heer *et al.* 2005). Instead, there is an effort to use existing information, often collected for other purposes, to gain a rough idea of how the state of nature is changing. While we contend that there is an urgent need to initiate well-designed programmes to measure changes in biodiversity, we accept that less satisfactory approaches should be employed in the short term. In this paper, we describe an existing programme for measuring biodiversity change, the Living Planet Index (LPI).

The LPI began in 1997 as a WWF project to develop a measure of the changing state of the world's biodiversity over time, and this remains its objective. Work on the LPI started in collaboration with the World Conservation Monitoring Centre in 1997. The first index was published in the WWF Living Planet Report 1998 (Loh *et al.* 1998) and has been updated subsequently (Loh *et al.* 1999; Loh 2000, 2002; Loh & Wackernagel 2004).

The LPI aims to measure average trends in populations of vertebrate species from around the world since 1970. Each iteration of the Living Planet Report has involved a new round of data collection, so the sample sizes of species populations in the index have grown with each successive edition. The index is currently based on nearly 3000 population time series for over 1100 species. All species in the index are vertebrates. The restriction of the index to vertebrate

---

[*] Author for correspondence (jonathan@livingplanet.org.uk).

One contribution of 19 to a Discussion Meeting Issue 'Beyond extinction rates: monitoring wild nature for the 2010 target'.

animals, and to years from 1970 onwards, is for reasons of data availability: relatively few time-series data for invertebrate or plant populations exist, and those come from geographically restricted locations. Therefore, the LPI is a measure of global biodiversity only as far as trends in vertebrate species populations are representative of wider trends in all species, genes and ecosystems.

The initial aim was to make the LPI as comprehensive and representative as possible with respect to vertebrate class, geography and biome. We felt that it should be based on the largest possible sample size to give the index the greatest precision possible. Representativeness of the species included could not be guaranteed by accepting as many eligible time series as possible, but efforts were made to allow for unrepresentativeness in the way that the index was calculated (see § 2b). A short time-interval between index values was also a goal. Here, we describe the compilation of data and the methods used to calculate the LPI; present index values at five-year intervals from 1970 to 2000; examine the representativeness of LPI data with respect to class, biogeographic realm, biome and threat status; and propose further development of the index.

2. METHODS

(a) *Collection of time-series and ancillary data*

Published scientific literature and unpublished reports were searched for eligible time-series data on vertebrate populations, as were online databases such as the NERC Imperial College Global Population Dynamics Database (see http://www.sw.ic.ac.uk/cpb/cpb/gpdd.html) and Ransom Myers' Stock Recruitment Database (see http://www.mscs.dal.ca/~mgers/welcome.html). Series were included if they met the following criteria:

1. Estimates available for at least two years from 1970 onwards.
2. Estimates of population size (global or regional), population density (e.g. numbers per unit area of survey plots, density from transects or point counts and numbers recorded per unit length of transects), biomass (e.g. spawning stock biomass from fisheries statistics) or numbers of nests (e.g. marine turtles). Numbers or densities of animals harvested by hunting or fisheries, though sometimes taken to be indicative of population size or density, were not used.
3. Survey methods and area covered were comparable throughout each survey of the series, as far as could be ascertained. Estimates for the same species from different workers or research teams published in different papers were not considered to be comparable unless a special effort had been made to ensure this.
4. Time series with little or no indication of how, where or when the data were collected were not used.

Whether a species was native or non-native was not used as a criterion in the data collection. Our analysis includes data from over one thousand species (table 1).

(b) *Preliminary processing of time-series data*

Because any population time-series data that met the above criteria were used in the calculations, the species included in the index were not necessarily representative of all vertebrate classes, biogeographic realms and ecological biomes. Therefore, before carrying out any calculations, the data were first divided up by biome—terrestrial, freshwater or marine—depending on the principal habitat of the species. Where a species commonly occurs in more than one biome, its breeding habitat was designated as its biome. Then, within each biome, species were divided up either according to the biogeographic realm they inhabit—Afrotropical, Australasian, Indo-Malayan, Nearctic, Neotropical or Palaearctic—for terrestrial and freshwater species, or to the ocean they inhabit—Atlantic/Arctic, Pacific, Indian or Southern Ocean—for marine species. For many species, there were time-series data for two or more populations within a single realm or ocean. For some species, one population occurred within one realm or ocean while another population would inhabit another, in which case the populations would be divided accordingly and those species would occur in more than one realm or ocean. Population time-series were assigned to biogeographic realms, following the system used for WWF ecoregions (WWF 2000), ocean basins and biomes based on expert knowledge of habitat requirements and information in standard reference works.

Table 1. Numbers of species included in the LPI by class and biome.

	terrestrial	freshwater	marine	total
fish		91	110	201
amphibians		49		49
reptiles	8	16	16	40
birds	381	132	95	608
mammals	172	11	46	229
total	561	299	267	1127

Separate indices were first calculated for each biogeographic realm and ocean. Multiple time-series for a single species within a realm or ocean would be treated as a single time-series, using the method described below, so that each species carried equal weight within each realm or ocean. Only estimates for the standard set of years 1970, 1975, 1980, 1985, 1990, 1995 and 2000 were used to calculate the index. Wherever an estimate was available for one of the standard set it was used. If there was no estimate for a standard year, but estimates were available before and after it, a value for the standard year was calculated by log-linear interpolation. Values for standard years were not obtained by extrapolation. For example, if counts were available for 1981 and 1984, values for the standard years 1980 and 1985 were not calculated. However, if counts were available for 1979 and 1986, values for 1980 and 1985 would be interpolated. Only series for which there were real or interpolated data for at least two standard years were included in further analyses. We recognize that it would be preferable to use only actual observations, to use data for all years and to avoid interpolation, but updating of the database for the early years is necessary before this can be done.

(c) *Calculation of the index by the chain method*

For each successive pair of standard years in each series we calculated the logarithm of the ratio of the population measure in one standard year to that of the standard year immediately preceding it. That is,

$$d_t = \log(N_t/N_{t-5}),$$

where the N are the two population measures. If a series

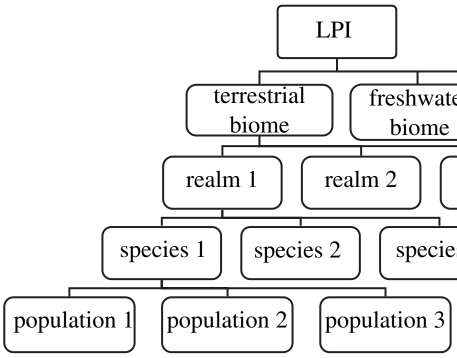


Figure 1. Hierarchy of indices within the LPI.

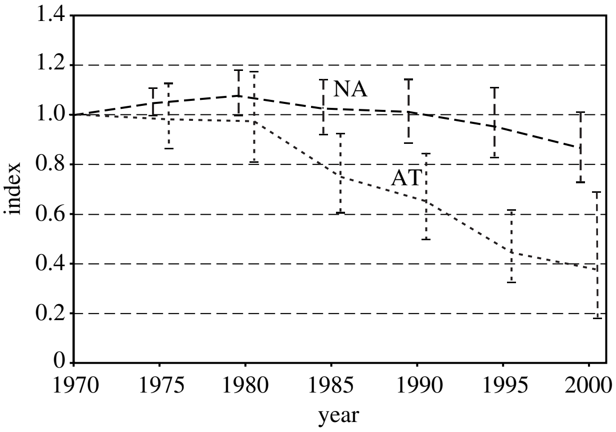


Figure 2. Nearctic terrestrial realm (NA) and Afrotropical terrestrial realm (AT) indices with 95% confidence intervals.

contained one or more standard years in which $N=0$, the mean of N for all standard years with data was calculated and 1% of this mean was added to every observed N before calculating d_t. Where there was more than one population series for a species for a given pair of standard years we calculated the mean value of d_t across all series. Then, given species-specific values of d_t for n_t species, we calculated

$$\bar{d}_t = \frac{1}{n_t}\sum_{i=1}^{n_t} d_{it}.$$

The index for a terrestrial or freshwater realm or an ocean basin in standard year t was calculated as

$$I_t = I_{t-5}10^{\bar{d}_t}.$$

Setting I_{1970} to 1, successive values of I_{1975}, I_{1980}, etc. were calculated.

Indices for terrestrial, freshwater and marine biomes were obtained by calculating the mean of the realm or ocean-specific $\bar{d}_t$ values, and then calculating I as described. The overall LPI was calculated from the mean of terrestrial, freshwater and marine $\bar{d}_t$ values (see figure 1).

Confidence intervals for I were obtained by a bootstrap method. Each bootstrap replicate was calculated by the following procedure. For each interval, $t-5$ to t, a sample of n_t species-specific values of d_t were selected at random from the n_t observed values with replacement. For a given realm or ocean, this was done for each interval, and $\bar{d}_t$ and I_t values were calculated as described in § 2c. The bootstrap procedure was carried out 1000 times and the bounds of the central 950 I values for each standard year were taken to represent the 95% confidence interval for the index in that year for that realm. An extension of this procedure was also used to calculate confidence intervals for terrestrial, freshwater and marine biomes and for the overall index. To do this, bootstrap index values for each realm within a biome were used to calculate bootstrap biome-level index values as described above for a given realm or ocean. Bootstrap biome-level index values were then used to calculate bootstrap overall index values.

(d) *Calculation of the index by a linear modelling method*

In the future it is intended to make better use of all of the annual data collected and to circumvent the need to use data for standard years at five-year intervals by using a linear model to calculate the index. This method was tested on the existing standard year dataset to check the comparability of results with the chain method on the same data. N values were transformed to $X_t = \log(N_t)$. Series in which any N was zero were first modified by adding 1% of the mean, as described in § 2c. For a given realm or basin, a least-squares linear model was fitted with X_t as the dependent variable, and series and year as independent factors acting as main effects. The analysis was weighted by giving all values for species i in standard year of t a weight of $w_{it} = 1/v_{it}$, where v_{it} is the number of series for that species with data available in year t. The analysis yielded coefficients b_t for the main effect of year, representing differences in the dependent variable between each standard year and the reference year (1970). Index values I_t for a given realm or basin were calculated by raising 10 to the power b_t. Indices for terrestrial, freshwater and marine biomes were obtained by calculating the mean of the realm- or basin-specific b_t values and then calculating I as described. The overall index for all biomes was calculated from the mean of biome-specific b_t values.

(e) *Analysis of the representativeness of LPI species*

To quantify the taxonomic representativeness of the species included in the LPI, the numbers of species in each vertebrate class in the LPI dataset were compared with known species totals taken from Groombridge & Jenkins (2000).

To compare the LPI species' representativeness in terms of biogeographic realm and biome, the LPI species were compared with species lists compiled for each of the WWF ecoregions. The 825 WWF ecoregions covering the earth's terrestrial surface are classified according to biome and biogeographic realm. Importantly, all vertebrate species except fish that have been recorded as occurring in each ecoregion have been listed, so the total numbers of non-fish vertebrates living in each realm and biome have been calculated. These lists were used to compare the numbers of species in each realm and each biome in the LPI dataset with the numbers of species recorded as occurring in each realm and biome. To simplify the analysis, the 14 terrestrial ecoregion biomes were reduced to four (table 2), and the non-marine species in the LPI dataset were reclassified into these four simplified ecoregion biomes. It was not possible to carry out this analysis separately for freshwater and terrestrial species and it could not be attempted at all for marine species because species lists have not yet been completed for marine ecoregions.

Finally, the representativeness in terms of threat status of the species in the LPI dataset was evaluated by comparing the number of LPI species in each threat category with the totals given in the 2000 IUCN Red List (Hilton-Taylor 2000). This was done for birds and mammals only because these were the only classes for which threat status had been assessed for all species.

Table 2. Simplified ecoregion biomes used to analyse LPI species representativeness.

simplified ecoregion biome	ecoregion biome
tropical and subtropical forest	tropical and subtropical moist broadleaf forest tropical and subtropical dry broadleaf forest tropical and subtropical coniferous forest mangroves
temperate and boreal forest	temperate broadleaf and mixed forest temperate coniferous forest mediterranean forest, woodland and scrub boreal forest/taiga
tropical grasslands, savannas and deserts	tropical and subtropical grasslands, savannas and shrublands deserts and xeric shrublands flooded grasslands and savannas
temperate grasslands, savannas and tundra	temperate grasslands, savannas and shrublands montane grasslands and savannas tundra

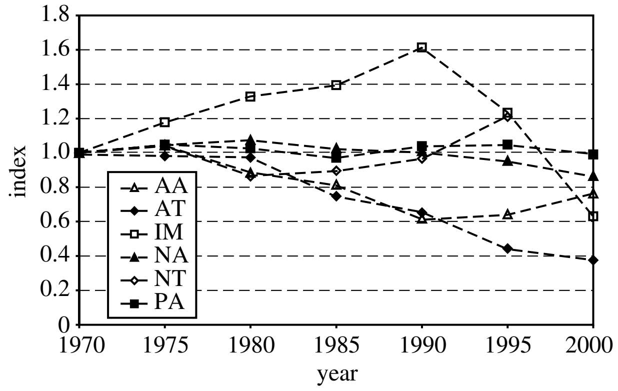


Figure 3. Indices for the six terrestrial realms (AA, Australasian; AT, Afrotropical; IM, Indo-Malayan; NA, Nearctic; NT, Neotropical; PA, Palaearctic).

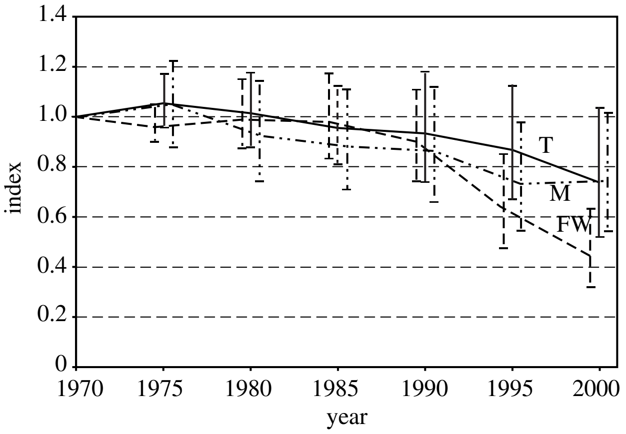


Figure 4. Terrestrial (T), freshwater (FW) and marine (M) indices, with 95% confidence intervals.

3. RESULTS

(a) *Index values obtained by the chain method*

The first step in the calculation of a terrestrial, freshwater or marine index is to calculate indices for each biogeographic realm. Examples for terrestrial species from the Nearctic and Afrotropical realms, based on data for 269 and 71 species, respectively, are shown in figure 2. The Nearctic realm index shows a −13% change from 1970 to 2000, with upper and lower 95% confidence limits at +1% and −27%, respectively, whereas the Afrotropical realm index shows a change of −62% over the same period, within confidence limits at −31% and −82%. The confidence intervals for the Afrotropical realm index are probably wider because it is based on a smaller sample of species.

The next step is to calculate the index for terrestrial species from all the realm-level indices (figure 3). It should be noted here that each realm-level index contributes equally to the calculation of the terrestrial index, regardless of the size of the sample of species and the series upon which they are based. The terrestrial species index declined by 25% from 1970 to 2000, with upper and lower 95% confidence limits at +4% and −48%, respectively (figure 4), which reflects the low precision of the component realm-level indices based upon small samples.

By following the same procedure, we calculated indices for freshwater and marine species from their realm-level components (figure 4). These indices are based on data for 299 species in six realms and 267 species in four oceans respectively. The freshwater species index fell by approximately 55% between 1970 and 2000, within 95% confidence limits ranging from −37% to −68%. The marine species index declined by about 25% over the same period, with upper and lower confidence limits at +2% and −46%.

As the last step in the procedure we calculated the overall LPI from the terrestrial, freshwater and marine indices (figure 5). It should again be noted that these three components are given equal weight in the calculation of the overall index, regardless of the number of species on which they are based. The overall LPI declined by 38% from 1970 to 2000, with upper and lower 95% confidence limits at −24% and −48%, respectively.

(b) *Comparison of indices calculated by the chain method and the linear modelling method*

We used the chain and linear modelling methods to calculate the terrestrial index (figure 6). The two methods give very similar results. The linear modelling

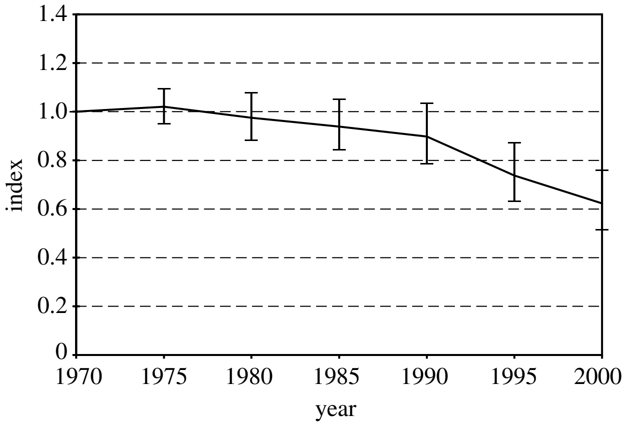


Figure 5. LPI with 95% confidence intervals.

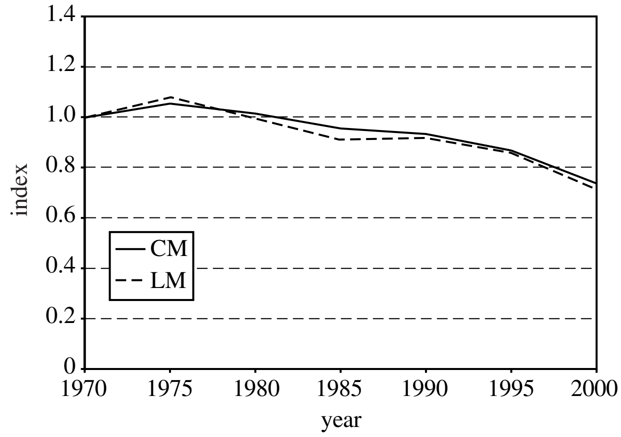


Figure 6. Terrestrial index calculated by both the chain method (CM) and the linear modelling (LM) method.

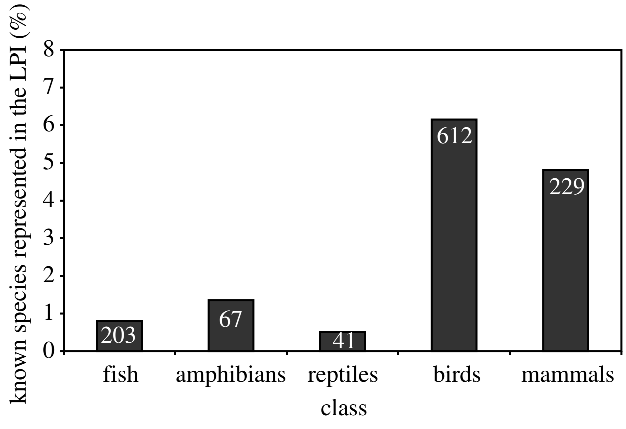


Figure 7. Percentage of known species in each vertebrate class that are represented in the LPI (figures in columns are total numbers of species included in the LPI).

index declined slightly (about three percentage points) more over the 30-year period than the chain index. Confidence intervals have not been calculated for the linear modelling index. A bootstrapping procedure could be implemented to do this, but this is not straightforward because bootstrapping on series would lead to substantial variation among bootstrap samples in the number of series contributing to the calculation of changes across different time-intervals.

(c) *Representativeness of species used to calculate the LPI*

Birds and mammals are better represented in the LPI than fish, amphibians or reptiles. Population time series for approximately 6% of known bird species and nearly 5% of known mammal species are included in the LPI dataset, whereas the figures for fish, amphibians and reptiles are about 1% or less (figure 7).

Nearctic and Palaearctic species are over-represented in the LPI by comparison with species from the Australasian, Afrotropical, Indo-Malayan or Neotropical realms. Approximately 20% of all Nearctic vertebrate species (excluding fish) are included in the LPI, and about 6% of Palaearctic vertebrate species (excluding fish), whereas the statistics for the other realms are only around 1% or less (table 3). This imbalance reflects the facts that, while many more species population data have been collected in temperate North America and Europe, species diversity is far higher in tropical than in temperate regions. The expected distribution of species among classes and realms is compared with the actual numbers of species in each class and realm in table 4. A negative number indicates that a class is under-represented in a given realm and a positive number indicates that a class is over-represented in a given realm.

Species of temperate and boreal forest classes are over-represented in the LPI compared with species of tropical and subtropical forest (6% versus 1%, respectively; figure 8). There is a similar over-representation of temperate grassland and tundra species compared with tropical grassland and desert species (3% versus <1%). This disparity again reflects the reality that data availability is lowest where species diversity is highest.

Threatened bird and mammal species are slightly better represented in the LPI than non-threatened birds and mammals, but the difference is small. Approximately 6% of both threatened and non-threatened bird and mammal species' populations are included in the LPI (figure 9). Within the broad category of threatened species, critically endangered species are slightly better represented (about 10% of LPI bird and mammal species) than endangered species (about 7%), which are slightly better represented than vulnerable species (about 5%).

4. CONCLUSIONS AND FUTURE DEVELOPMENTS

The LPI indicates that populations of wild species of vertebrates have declined overall from 1970 to 2000. The extent to which this is a reflection of trends in global biodiversity as a whole has not been determined. In situations where habitat loss is the primary cause of population declines, it is reasonable to assume that there is a positive correlation between declines in vertebrate and non-vertebrate populations. Where hunting, fishing or indirect exploitation is the cause of a decline in a particular vertebrate species, the decline will not necessarily be indicative of population trends in other species in the same ecosystem. At large scales of entire realms, oceans, regions or biomes, overall declines in vertebrate populations are significant in their own right and may also be seen as indicative of changes in underlying ecosystem processes.

Table 3. Number of non-marine vertebrate species in the LPI dataset, excluding fish, by vertebrate class and biogeographic realm.
(Terrestrial and freshwater biomes only.)

	Australasian	Afrotropical	Indo-Malayan	Nearctic	Neotropical	Palaearctic	total
amphibians	5	0	9	46	2	5	67
reptiles	3	7	6	2	5	2	25
birds	12	21	4	293	17	172	519
mammals	6	54	28	41	7	45	181
total	26	82	47	382	31	224	792
percentage of all recorded species	0.6	1.4	0.9	20.0	0.3	6.3	2.6

Table 4. Actual minus expected number of species in the LPI dataset, excluding fish, in each realm and vertebrate class. (Terrestrial and freshwater biomes only.)

	Australasian	Afrotropical	Indo-Malayan	Nearctic	Neotropical	Palaearctic	total
amphibians	−9	−20	−11	39	−58	−5	−63
reptiles	−31	−37	−30	−10	−60	−18	−186
birds	−36	−36	−47	275	−81	133	208
mammals	−12	24	4	29	−26	22	41
total	−87	−69	−84	333	−225	132	0

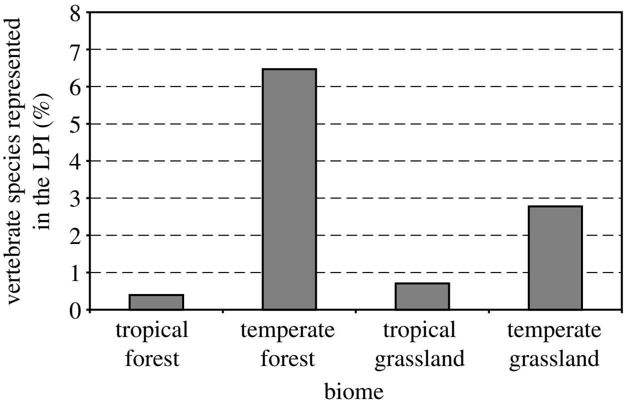


Figure 8. Percentage of all vertebrate species (excluding freshwater fish) recorded in each simplified ecoregion biome that are represented in the LPI.

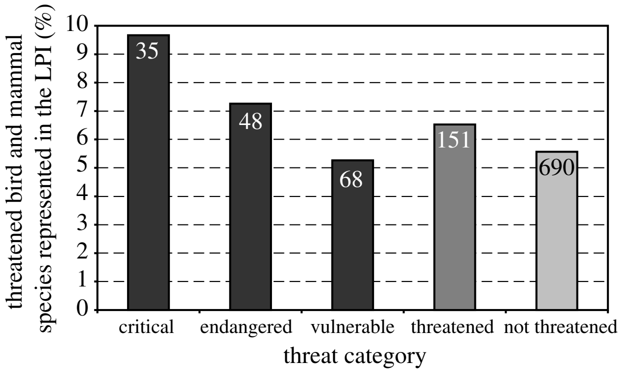


Figure 9. Percentage of all threatened bird and mammal species in each IUCN threat category that are represented in the LPI (figures in columns are total number of species included in the LPI).

The LPI has a number of strengths as an indicator in the context of the CBD 2010 target. First, it is easy to understand and easy to communicate. It resonates with a non-scientific audience because it is analogous to well-known indicators like the Dow Jones Index, which seek to measure the mood of the financial community. Second, and importantly, data exist going back to the 1970s for many species, and even earlier for some. The existence of reasonably long time series is crucial for monitoring progress towards the 2010 target. Third, species population trends can be used as a proxy indicator of the state of the ecosystem that the species inhabit. Finally, species population trend indices can be constructed as indicators of biodiversity at any level: nationally, regionally, globally, or by biome or biogeographic realm, provided that sufficient data exist.

There are also a number of weaknesses with the LPI as a global biodiversity index. These weaknesses all relate to the representativeness of the population data. Data were taken from the literature according to availability and are often not the results of a designed programme of sampling of representative sites within a given species' range, nor of representative species within a biogeographic realm and vertebrate class. For some species, estimates of total population size were available, but more frequently the data were for small parts of a much larger range. These surveys were sometimes for randomly selected sites designed to cover the range, but more often they were not. Hence, it is possible that trends indicated by the series that were available for some species were not representative of those for the species as a whole. It is also possible that trends for the species for which we have data were not typical of species of that class in the biogeographic realm as a whole.

We have not yet attempted to quantify the extent of possible biases. It might be possible to obtain information regarding the motivation of researchers in setting up or publishing their work. It is certainly possible that some monitoring programmes have been carried out because a species or population was thought to be declining, or to monitor a response to beneficial management, either of which could lead to bias. It is difficult to see how these potential sources of bias could be fully quantified and even more difficult to see how they could be allowed for.

Even if the trends of species included in the LPI are taken to be representative of those for the whole of the particular class and realm to which they belong, there remains the problem of how to allow for the fact that a higher proportion of species from some classes and realms is included than from others. The over-representation in the dataset of species from certain realms and ocean basins is already partially allowed for in the method chosen to calculate the index, which gives equal weight to data-poor and data-rich realms. However, there is no equivalent allowance for the over-representation of some classes relative to others within a realm. In principle, it would be possible to post-stratify the data by class and realm and then to calculate the overall index for, say, terrestrial species by using weights of $1/p_{ij}$, where p_{ij} is the proportion of species of class i present in the jth realm that are covered by the LPI database. This would reconstruct the expected index obtained if all class–realm combinations were subject to the same sampling rate. However, this is not possible in practice because the numbers of species for which data are available are too small in many class–realm combinations. Expansion of the database, especially for class–realms with data on few species, might eventually allow this approach to be implemented. The representation analysis (table 4) provides an indication of the realms and classes of species most needed.

Over-representation in the data of threatened species might also be a source of bias although our analysis indicates that this does not occur to any marked extent in the current datasets for birds and mammals. If present, this bias could be corrected by an extension of the method described in § 2b for class and realm, but this would require the expansion of the numbers of species contributing information to data-poor class–realm–threat category combinations.

Other problems caused by the scarcity of data are that (i) not all realms have data for the last time interval 1995–2000, so the aggregated index across all realms is unbalanced for this period, and (ii) the numbers of species in the sample for some years in some realms in some biomes are too small to use the bootstrap method to calculate confidence intervals reliably.

Indices consisting of a chain of estimates of year-to-year change may show drift if series begin and end in different years and do not cover the whole time period (Geissler & Noon 1981). Though our data are of this kind, there was little difference between indices calculated by the chain method and by the least-squares linear model, which allows for staggered entry and termination and missing data in intervening years. Even so, we propose that linear model calculations should replace the chain method because it would then be possible to use the full dataset of annual data without interpolating missing values.

At present the index includes populations of indigenous and non-native species. This may be regarded as undesirable for many purposes because declines in populations of indigenous species represent a decline in the biodiversity value in a region, whereas many would consider declines in alien species to increase biodiversity value. Data for non-native species could be deleted from the LPI in future. The option of producing separate indices for native and non-native species is not feasible at present because the data available for non-native species are insufficient.

In conclusion, although the LPI has faults that cannot easily be remedied, we suggest that, even in its present form, some of the problems arising from the over-representation of some classes and realms are reduced by the way that the index is calculated. It would be possible to carry this reduction further by post-stratification and weighting if sufficient new information can be collected for data-poor classes, realms and biomes. Putting all the LPI population time-series into a searchable interactive database that is accessible on the Internet and allowing remote data input from anywhere in the world might allow the necessary expansion of data holdings, though special efforts such as regional workshops to stimulate the collection and donation of data would probably also be needed. Rigorous and transparent systems for checking the quality of data are also required and might be organized by groups of experts on each vertebrate class.

REFERENCES

de Heer, M., Kapos, V. & ten Brink, B. 2005 Building a policy relevant species trend index for Pan-Europe from disparate data sources. *Phil. Trans. R. Soc. B* **360**.

CBD 2000 Decision VI/26. UNEP/CBD/COP/6/20.

Geissler, P. H. & Noon, B. R. 1981 Estimates of avian population trends from the North American breeding bird survey. *Stud. Avian Biol.* **6**, 42–51.

Green, R. E., Balmford, A. P., Crane, P. R., Mace, G. M., Reynolds, J. D. & Turner, R. K. 2005 A framework for improved monitoring of biodiversity: responses to the World Summit on Sustainable Development. *Conserv. Biol.* **19**, 56–65.

Gregory, R. D., van Strien, A. & Vorisek, P. 2005 Developing indicators for European birds. *Phil. Trans. R. Soc. B* **360**.

Groombridge, B. & Jenkins, M. D. 2000 *Global biodiversity: Earth's living resources in the 21st century.* Cambridge: World Conservation Press.

Hilton-Taylor, C. 2000 *2000 IUCN Red List of threatened species*. Gland, Switzerland: IUCN.

Kapos, V., Jenkins, M. & Loh, J. 2004 *Rising to the biodiversity challenge. The role of species population trend indices like the Living Planet Index in tracking progress towards global and national biodiversity targets*. Washington, DC: World Bank.

Loh, J. (ed.) 2000 *Living Planet Report 2000*. Gland, Switzerland: WWF.

Loh, J. (ed.) 2002 *Living Planet Report 2002*. Gland, Switzerland: WWF.

Loh, J. & Wackernagel, M. (eds) 2004 *Living Planet Report 2004*. Gland, Switzerland: WWF.

Loh, J., Randers, J., MacGillivray, A., Kapos, V., Jenkins, M., Groombridge, B. & Cox, N. 1998 *Living Planet Report 1998*. Gland, Switzerland: WWF.

Loh, J., Randers, J., MacGillivray, A., Kapos, V., Jenkins, M., Groombridge, B., Cox, N. & Warren, B. 1999 *Living Planet Report 1999*. Gland, Switzerland: WWF.

WWF 2000 *The global 200 ecoregions. A user's guide*. Washington, DC: WWF.

A Global Index of Biocultural Diversity

Jonathan Loh and David Harmon

Ecological Indicators (2005) 231-241

ELSEVIER

Ecological Indicators 5 (2005) 231–241

ECOLOGICAL
INDICATORS

This article is also available online at:
www.elsevier.com/locate/ecolind

A global index of biocultural diversity

Jonathan Loh [a,*], David Harmon [b]

[a] Langley Court West, Liss, Hampshire GU33 7JL, UK
[b] George Wright Society, P.O. Box 65, Hancock, MI 49930-0065, USA

Accepted 28 February 2005

Abstract

The relationships between biological and cultural diversity are drawing increasing attention from scholars. Analyses of these relationships are beginning to crystallize around the concept of *biocultural diversity*, the total variety exhibited by the world's natural and cultural systems. Here, we present the first global measure of biocultural diversity, using a country-level index. The index is calculated using three methods: an unadjusted richness measure, one adjusted for land area, and one adjusted for the size of the human population. The adjusted measures are derived from the differences between observed and expected diversity values. Expected diversity was calculated using the species–area relationship. The index identifies three areas of exceptional biocultural diversity: the Amazon Basin, Central Africa, and Indomalaysia/Melanesia.

Keywords: Biological diversity; Cultural diversity; Biocultural diversity

1. Introduction

The relationships between biological and cultural diversity, and the growing threats they face, have drawn increasing attention from scholars over the last decade (Harmon, 2002; Moore et al., 2002; Sutherland, 2003; Maffi, in press). Analyses of these relationships are beginning to crystallize around the concept of *biocultural diversity*, the total variety exhibited by the world's natural and cultural systems (Maffi, 2001). Here, we outline the first attempt to quantify global biocultural diversity by means of a country-level index, calculated in three ways: an unadjusted richness measure, a measure of richness adjusted for land area, and a measure of richness adjusted for the size of the human population. These measures, when analyzed in concert, indicate three areas of exceptional biocultural diversity. By pinpointing these areas, the index of biocultural diversity (IBCD) will help raise awareness about the threats facing both biological and cultural diversity and could help produce more enlightened public policy for their protection.

Biocultural diversity may be thought of as the sum total of the world's differences, no matter what their origin. It includes biological diversity at all its levels, from genes to populations to species to ecosystems; cultural diversity in all its manifestations (including

---

* Corresponding author.
E-mail addresses: jonathan@livingplanet.org.uk (J. Loh), dharmon@georgewright.org (D. Harmon).

1470-160X/$ – see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ecolind.2005.02.005

linguistic diversity), ranging from individual ideas to entire cultures; and, importantly, the interactions among all of these. On a global scale, the primary importance of biocultural diversity is that it is the fundamental expression of the variety upon which all life is founded. Conceptually, biocultural diversity bridges the divide between disciplines in the social sciences that focus on human creativity and behavior, and those in the natural sciences that focus on the evolutionary fecundity of the non-human world. The result is a more integrated view of the patterns that characterize life on Earth.

A basic premise of first-generation scholarship on biocultural diversity has been that the relationships between humans and non-human species, and between them both and the landscapes they inhabit, do not run on parallel tracks. Rather, these forms of diversity are often closely linked, and sometimes may even be constitutive of each other in important ways. Much of this first-wave scholarship has aimed to establish correlations between biological and cultural/linguistic diversity in terms of geography, such as areas of overlap (Moore et al., 2002; Manne, 2003; Sutherland, 2003); theory, such as how language may be related to long-term environmental management in indigenous communities (Maffi, 2001; Harmon, 2002); and common threats to their continuation (Maffi, in press). Among the challenges for the next wave of scholars will be (1) to see if the relationships go deeper than mere correlations to something approaching actual coevolution; (2) to elucidate the complexities of how humans and non-human species interact not only with one another but also with the abiotic or geophysical diversity of the earth, including that of its landforms and geological processes, meteorology, and all other inorganic components and processes (e.g. chemical regimes) that provide the setting for life (see Gray, 2004); (3) to deepen the theoretical foundations of biocultural diversity research. In all these aims, it would be useful to have quantitative measures of biocultural diversity on a global level.

The IBCD begins to fill this gap by using a combination of five indicators to establish rankings of biocultural diversity for 238 countries and territories. We used the number of languages, religions, and ethnic groups present within each country as a proxy for its cultural diversity, and the number of bird and mammal species and the number of plant species as a measure of its biological diversity. The IBCD has three parts:

- A *biocultural diversity richness component* (BCD-RICH), which is a relative measure of a country's 'raw' biocultural diversity using unadjusted counts of the five indicators.
- An *areal component* (BCD-AREA), which adjusts the indicators for land area and therefore measures a country's biocultural diversity relative to its physical extent.
- A *population component* (BCD-POP), which adjusts the indicators for human population and therefore measures a country's biocultural diversity relative to its population size.

2. Methods

The IBCD gives equal weight to cultural and biological diversity, so a country's overall biocultural diversity score is calculated as the average of its cultural diversity score (CD) and its biological diversity score (BD).

$$\mathrm{IBCD} = \frac{\mathrm{CD} + \mathrm{BD}}{2}$$

In measuring a country's cultural diversity CD, equal weight is given to linguistic, religious and ethnic diversity. Therefore CD is calculated as the average of a country's language diversity (LD), religion diversity (RD), and ethnic group diversity (ED):

$$\mathrm{CD} = \frac{\mathrm{LD} + \mathrm{RD} + \mathrm{ED}}{3}$$

In measuring biodiversity BD, equal weight is given to animal species diversity (using birds and mammals as a proxy for all animal species marine mammals are excluded from the analysis) and plant species diversity. Therefore BD is calculated as the average of a country's bird and mammal species diversity (MD), and plant species diversity (PD):

$$\mathrm{BD} = \frac{\mathrm{MD} + \mathrm{PD}}{2}$$

Each indicator is given an equal weighting as this is the simplest way of calculating the index. As an aggregated index, the IBCD could be calculated using different weightings, to give greater or lesser impor-

Table 1
Unadjusted language diversity index (LD-RICH)

	No. of languages (L)	log L	LD-RICH (log L_i/log L_{world})
World	6800	3.83	1.000
Papua New Guinea (highest)	833	2.92	0.762
Mali (average)	45	1.65	0.431
Bermuda (lowest)	1	0.00	0.000

tance to any of the five component indicators. Alternative weightings are not analyzed here.

To derive country scores for each of the five component indicators, we compared each country's richness value with the global value. For example, for language diversity, LD is calculated as the log of the number of languages spoken in a country divided by the log of the number of languages spoken worldwide (see Table 1).

$$\text{LD} = \frac{\log L_i}{\log L_{\text{world}}}.$$

where L_i is the number of languages spoken in country i, L_{world} the number of languages spoken in the world (currently 6800).

The calculation was repeated for the other four indicators to derive BCD-RICH. Detailed discussion of the methods is included in the index's source document (Harmon and Loh, 2004). Data sources were as follows: languages (Grimes, 2000), religions (Barrett et al., 2001), ethnic groups (Barrett et al., 2001), bird/mammal species (Groombridge and Jenkins, 2002), plant species (Groombridge and Jenkins, 2002), country area (The Times, 2000; countries smaller than 1000 sq km are excluded), and country population (FAO, 2004; countries with a population of less than 10,000 are excluded).

To compensate for the fact that large countries tend to have a greater biological and cultural diversity than small ones simply because of their greater area (or greater population), we calculated two additional diversity values for each country by adjusting first for land area (BCD-AREA) and second for population size (BCD-POP). This was done by measuring how much more or less diverse a country is in comparison with an expected value based on its area or population alone. The method used is a modified version of that used by Groombridge and Jenkins (2002). As an example of the methods used, calculations for the language indicator value are shown in Tables 2 and 3. The process

Table 2
Area-adjusted language diversity index (LD-AREA)

Country or territory	Area (km^2)	log A	Total no. of languages (L)	log L	Expected log L value	Deviation from expected value	LD-AREA
World/maximum value	136605342	8.14	6800	3.83	2.33	1.50	1.000
Papua New Guinea (highest)	462840	5.67	833	2.92	1.56	1.36	0.952
Turkmenistan (average)	488100	5.69	37	1.57	1.57	0.00	0.500
Greenland (lowest)	2175600	6.34	2	0.30	1.77	−1.47	0.011
Minimum value						−1.50	0.000

Table 3
Population-adjusted language diversity index (LD-POP)

Country or territory	Population 2000 (thousand) P	log P	Total no. of languages (L)	log L	Expected log L value	Deviation from expected value	LD-POP
Maximum value	6056710	6.78	12000[a]	4.08	2.48	1.60	1.000
Papua New Guinea (highest)	4809	3.68	833	2.92	1.34	1.58	0.995
Pakistan (average)	141256	5.15	76	1.88	1.88	0.00	0.501
Korea, DPR (lowest)	22268	4.35	2	0.30	1.58	−1.28	0.099
Minimum value						−1.60	0.000

[a] Artificial number of languages chosen to create a maximum value higher than the highest-ranking country.

Table 4
IBCD-RICH: 20 highest-ranking countries

Country or territory	Total no. lang-uages (*L*)	Language diversity index, LD-RICH	Total no. of religions (*R*)	Religion diversity index, RD-RICH	No. of ethnic groups (*E*)	Ethnic group diversity index, ED-RICH	Cultural diversity index, CD-RICH	Total no. bird and mammal species (*M*)	Birds and mammal diversity index, MD-RICH	Total no. plant species (*P*)	Plants diversity index, PD-RICH	Bio-logical diversity index, BD-RICH	Index of bio-cultural diversity IBCD-RICH
World/maximum value	6800	1.000	10000	1.000	12583	1.000	1.000	14709	1.000	250876	1.000	1.000	1.000
Indonesia	736	0.748	535	0.682	744	0.700	0.710	2034	0.794	29375	0.827	0.811	0.760
Papua New Guinea	833	0.762	648	0.703	862	0.716	0.727	858	0.704	11544	0.752	0.728	0.728
Brazil	246	0.624	183	0.566	224	0.573	0.588	1886	0.786	56215	0.880	0.833	0.710
India	414	0.683	293	0.617	439	0.645	0.648	1313	0.748	18664	0.791	0.770	0.709
China	207	0.604	156	0.548	254	0.587	0.580	1494	0.762	32200	0.835	0.798	0.689
Nigeria	521	0.709	460	0.666	497	0.658	0.677	955	0.715	4715	0.680	0.698	0.688
United States	284	0.640	141	0.537	307	0.607	0.595	1078	0.728	19473	0.794	0.761	0.678
Cameroon	288	0.642	250	0.599	297	0.603	0.615	1099	0.730	8260	0.725	0.728	0.671
Congo, Dem Rep (Zaire)	221	0.612	173	0.560	260	0.589	0.587	1379	0.753	11007	0.749	0.751	0.669
Colombia	101	0.523	77	0.472	99	0.487	0.494	2054	0.795	51220	0.872	0.834	0.664
Mexico	303	0.647	36	0.389	278	0.596	0.544	1260	0.744	26071	0.818	0.781	0.663
Australia	315	0.652	83	0.480	133	0.518	0.550	901	0.709	15638	0.777	0.743	0.646
Peru	108	0.531	67	0.457	111	0.499	0.495	1998	0.792	17144	0.784	0.788	0.642
Malaysia	146	0.565	123	0.522	174	0.547	0.545	801	0.697	15500	0.776	0.736	0.640
Tanzania	141	0.561	119	0.519	163	0.540	0.540	1138	0.733	10008	0.741	0.737	0.638
Russia	119	0.542	67	0.457	169	0.543	0.514	897	0.709	11400	0.751	0.730	0.622
Myanmar	113	0.536	89	0.487	133	0.518	0.514	1167	0.736	7000	0.712	0.724	0.619
Sudan	142	0.562	119	0.519	245	0.583	0.554	947	0.714	3137	0.648	0.681	0.618
Philippines	184	0.591	152	0.545	183	0.552	0.563	349	0.610	8931	0.732	0.671	0.617
Ethiopia	88	0.507	118	0.518	145	0.527	0.518	903	0.709	6603	0.707	0.708	0.613

was repeated for the other four indicators to derive BCD-AREA and BCD-POP.

The expected diversity was calculated using the standard formula for the species–area relationship $\log S = c + z \log A$ where S = number of species, A = area, and c and z are constants derived from observation. Because the distributions of the five indicators against land area and population size are similar, we applied the same formula to indicators of cultural diversity. Hence, for BCD-AREA expected $\log N_i = c + z \log A_i$ where N_i = number of languages, religions, ethnic groups, or species in country i, and A_i = area of country i. The same formula was used for BCD-POP, except that P_i (population of country i) replaces A_i. To find the values of the constants c and z for each of the indicators, we scatter-plotted $\log N_i$ (where N_i = number of languages, religions, ethnic groups, or species in country i) against $\log A_i$ for all countries, and drew the best-fit straight line through the points. Examples for bird/mammal species and languages are in Figs. 1 and 2, respectively.

To calculate the deviation of each country from its expected value, we subtracted the expected $\log N_i$ value from the observed $\log N_i$ value. The index is calibrated such that the world, or maximum, value is set equal to 1.0, the minimum value is set equal to zero and the average or typical value is 0.5 (meaning no more or less diverse than expected given a country's area or population).

3. Results

By combining the results of BCD-RICH, BCD-AREA, and BCD-POP, we identified three 'core areas' of global biocultural diversity that include countries of various sizes and populations:

- The Amazon Basin, consisting of Brazil, Columbia and Peru, which ranked highly in BCD-RICH; Ecuador, which ranked highly in BCD-AREA; and French Guiana, Suriname and Guyana, which ranked highly in BCD-POP.

Table 5
IBCD-AREA: 20 highest-ranking countries

Country or territory	Area (km^2)	Language diversity index, LD-AREA	Religion diversity index, RD-AREA	Ethnic group diversity index, ED-AREA	Cultural diversity index, CD-AREA	Bird & mammal diversity index, MD-AREA	Plant diversity index, PD-AREA	Biodiversity index, BD-AREA	Index of biocultural diversity, IBCD-AREA
World/maximum value	136605342	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Indonesia	1919317	0.870	0.787	0.785	0.814	0.671	0.751	0.711	0.762
Papua New Guinea	462840	0.952	0.837	0.850	0.880	0.597	0.663	0.630	0.755
Colombia	1141568	0.607	0.596	0.549	0.584	0.704	0.882	0.793	0.688
Cameroon	475442	0.797	0.737	0.715	0.750	0.641	0.600	0.621	0.685
Malaysia	330442	0.715	0.671	0.660	0.682	0.605	0.736	0.671	0.676
Brunei	5765	0.602	0.552	0.515	0.557	0.767	0.798	0.782	0.669
India	3165596	0.765	0.713	0.702	0.727	0.560	0.639	0.600	0.663
Nigeria	923768	0.853	0.787	0.758	0.799	0.576	0.459	0.518	0.658
Nepal	147181	0.727	0.641	0.638	0.669	0.651	0.637	0.644	0.657
Brazil	8547404	0.645	0.643	0.586	0.625	0.567	0.782	0.675	0.650
Mexico	1958201	0.741	0.506	0.661	0.636	0.582	0.728	0.655	0.645
Peru	1285216	0.611	0.579	0.560	0.583	0.692	0.676	0.684	0.633
Ecuador	272045	0.486	0.514	0.458	0.486	0.754	0.788	0.771	0.628
Philippines	300076	0.753	0.696	0.670	0.706	0.458	0.641	0.550	0.628
Viet Nam	331041	0.656	0.621	0.591	0.623	0.592	0.665	0.629	0.626
Tanzania	942799	0.663	0.646	0.618	0.642	0.607	0.595	0.601	0.622
Laos	236800	0.656	0.628	0.598	0.627	0.589	0.641	0.615	0.621
Congo, Dem Rep	2345095	0.687	0.665	0.647	0.666	0.587	0.560	0.574	0.620
Panama	75517	0.487	0.524	0.500	0.504	0.725	0.740	0.733	0.618
Solomon Islands	28370	0.729	0.668	0.637	0.678	0.511	0.589	0.550	0.614

Table 6
IBCD-POP: 20 highest-ranking countries

Country or territory	Population 2000 (thousand)	Language diversity index, LD-POP	Religion diversity index, RD-POP	Ethnic group diversity index, ED-POP	Cultural diversity index, CD-POP	Bird & mammal diversity index, MD-POP	Plant diversity index, PD-POP	Biodiversity index, BD-POP	Index of biocultural diversity, IBCD-POP
WORLD/maximum value	6056710	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Papua New Guinea	4809	0.995	0.965	0.936	0.965	0.756	0.785	0.771	0.868
French Guiana	165	0.618	0.624	0.590	0.611	0.895	0.901	0.898	0.754
Suriname	417	0.611	0.622	0.572	0.602	0.942	0.805	0.874	0.738
Cameroon	14876	0.794	0.801	0.743	0.780	0.720	0.629	0.675	0.727
Indonesia	212092	0.789	0.807	0.756	0.784	0.641	0.682	0.662	0.723
Brunei	328	0.616	0.586	0.530	0.577	0.863	0.860	0.862	0.719
Colombia	42105	0.600	0.612	0.550	0.587	0.781	0.921	0.851	0.719
Gabon	1230	0.654	0.630	0.608	0.631	0.808	0.779	0.793	0.712
Guyana	761	0.566	0.577	0.526	0.557	0.916	0.809	0.862	0.710
Solomon Islands	447	0.786	0.762	0.705	0.751	0.628	0.706	0.667	0.709
Peru	25662	0.634	0.611	0.587	0.610	0.816	0.736	0.776	0.693
Australia	19138	0.794	0.649	0.623	0.689	0.651	0.740	0.695	0.692
Brazil	170406	0.651	0.675	0.602	0.643	0.642	0.831	0.737	0.690
Belize	226	0.593	0.542	0.545	0.560	0.878	0.741	0.809	0.685
Congo	3018	0.674	0.674	0.630	0.659	0.729	0.688	0.709	0.684
Laos	5279	0.683	0.683	0.635	0.667	0.685	0.711	0.698	0.682
Bolivia	8329	0.577	0.584	0.546	0.569	0.740	0.826	0.783	0.676
Malaysia	22218	0.682	0.695	0.654	0.677	0.610	0.727	0.668	0.673
Panama	2856	0.507	0.543	0.514	0.522	0.825	0.795	0.810	0.666
Central African Republic	3717	0.689	0.673	0.647	0.670	0.745	0.568	0.656	0.663

- Central Africa, consisting of Nigeria, Cameroon and the Democratic Republic of Congo (BCD-RICH), Tanzania (BCD-AREA) and Gabon and Congo (BCD-POP).
- Indomalaysia/Melanesia, consisting of Papua New Guinea and Indonesia (BCD-RICH), Malaysia and Brunei (BCD-AREA) and Solomon Islands (BCD-POP).

The world's four most bioculturally diverse countries – Papua New Guinea, Indonesia, Cameroon, and Colombia – rank in the top 10 for all three components of the index (see Tables 4–6 and Maps 1–3).

4. Discussion

The index of biocultural diversity has both theoretical and practical implications. For researchers of the interchanges between biological and cultural diversity, it provides a global context against which fine-grained analyses can be compared. For policy-makers and donor organizations, it is a potential framework for guiding strategic investments in biocultural diversity conservation. The three 'core areas' identified above are in that sense analogous to the results of several schemes that recently have been developed for identifying the world's most important areas for biodiversity conservation and ecoregion protection (Davis et al., 1994; Stattersfield et al., 1998; Myers et al., 2000; Olson et al., 2001). For the general public, the index serves as a reminder that no matter where a country ranks, its biocultural diversity is an important part of the global complement.

The purpose of any global index is to use simple proxies to indicate the status of complex phenomena. Our index is intended to provide a snapshot of the current distribution of the world's biocultural diversity. As more and better data become available, particularly on the numbers of individuals in each language group, religion, ethnic group, or species, it will be possible to analyze trends. Then we will be

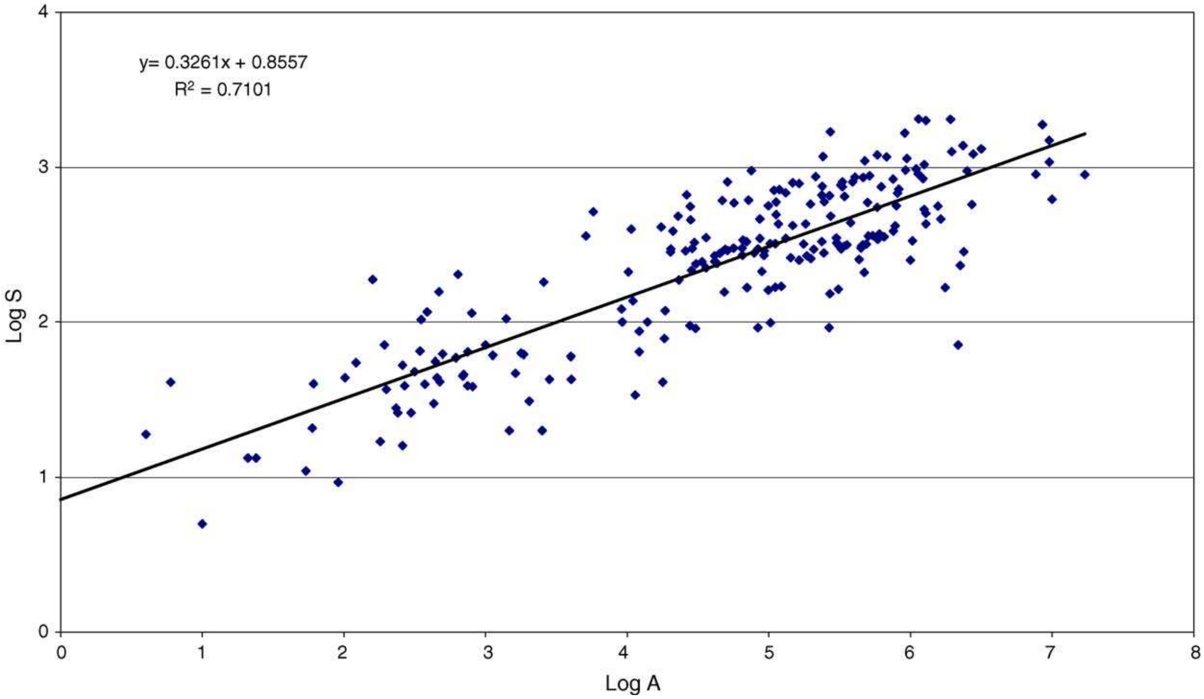


Fig. 1. Bird/mammal species—area plot.

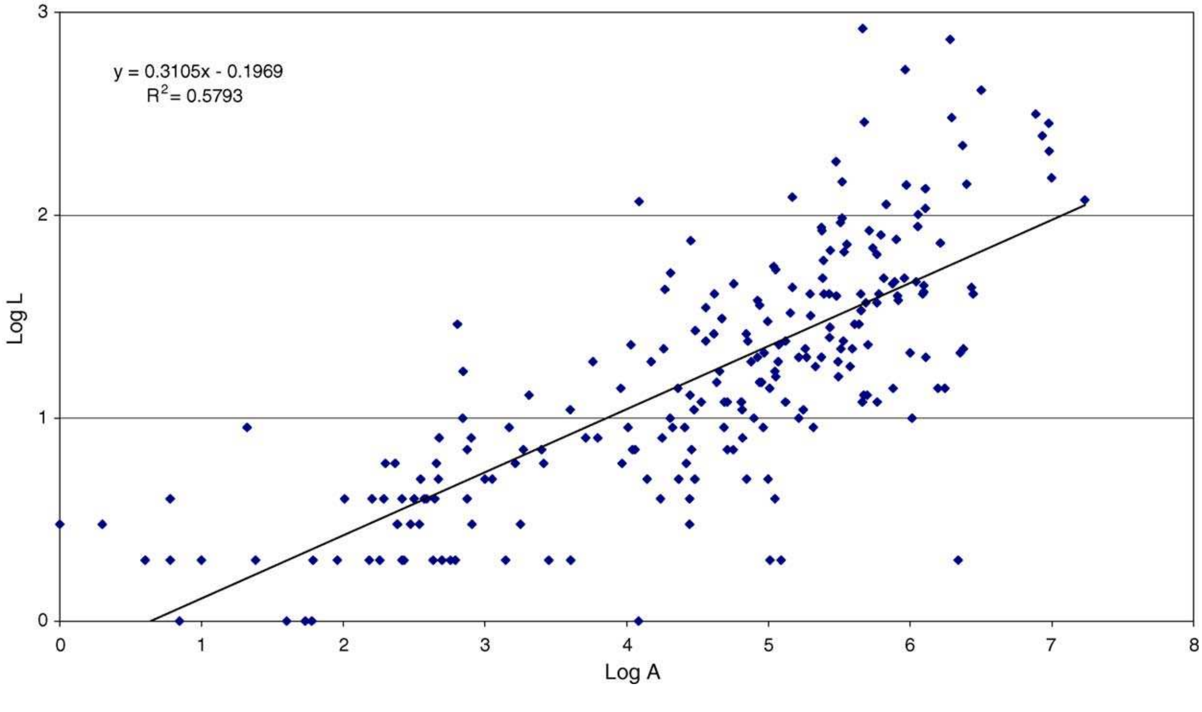
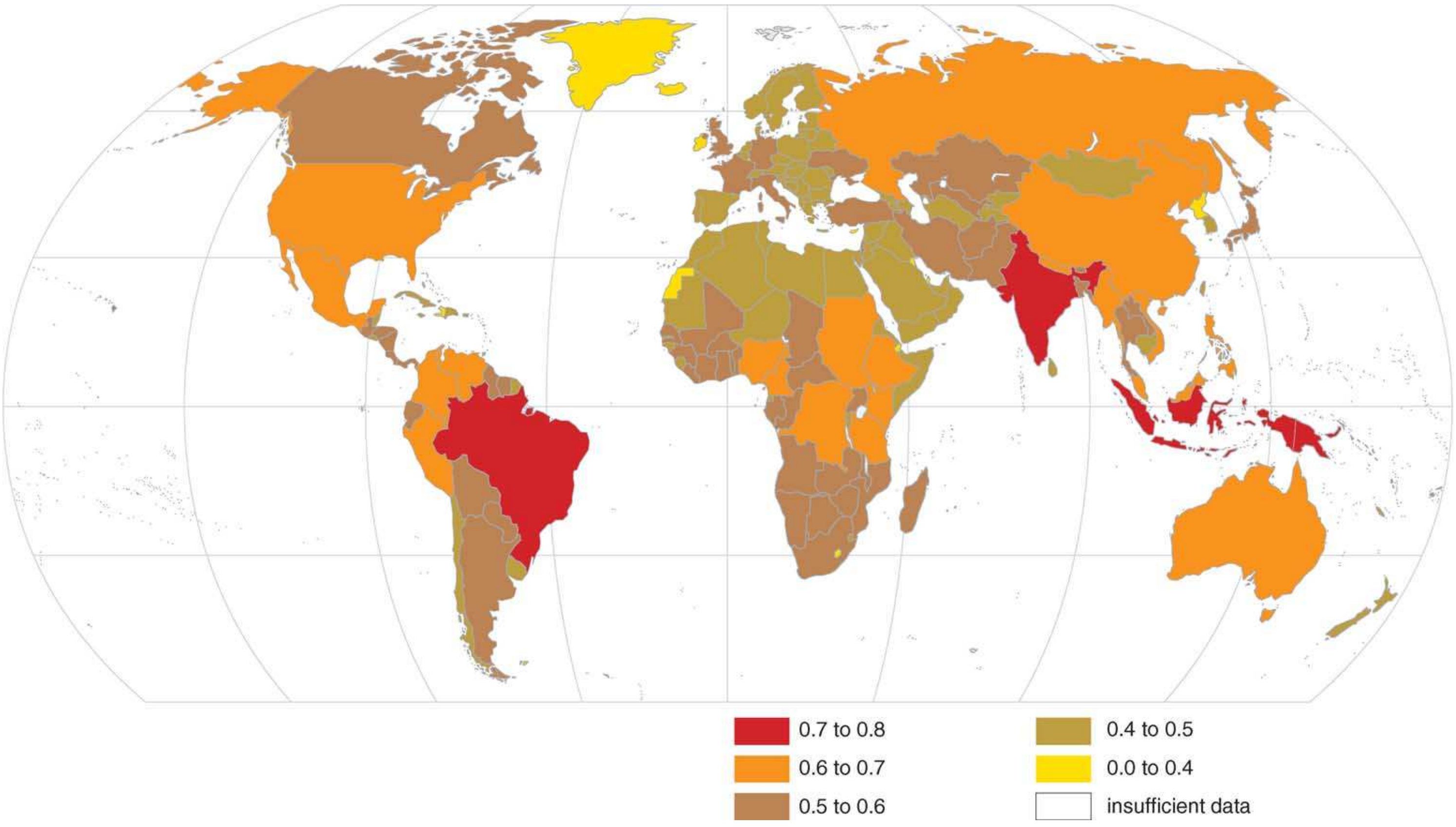
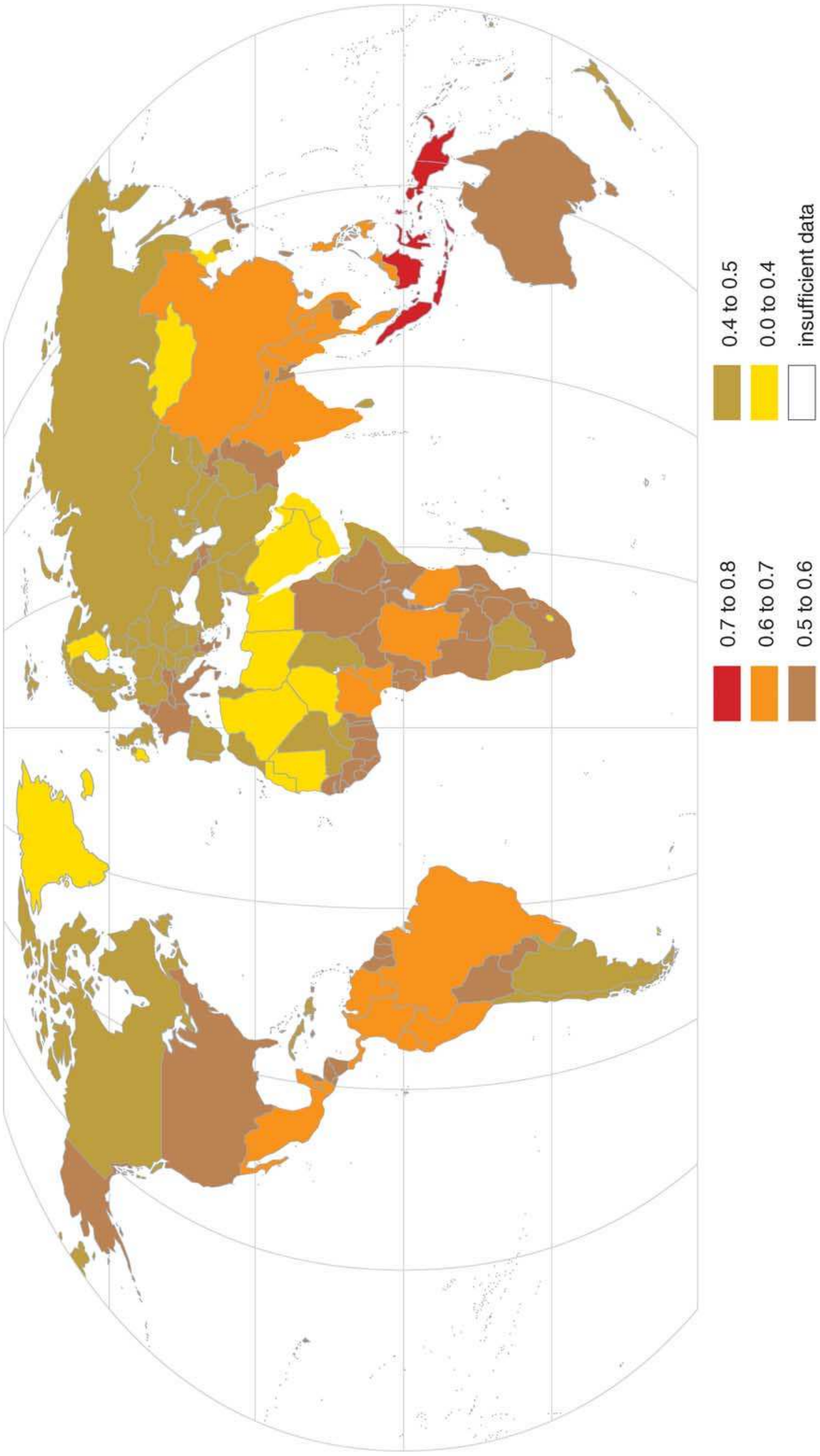


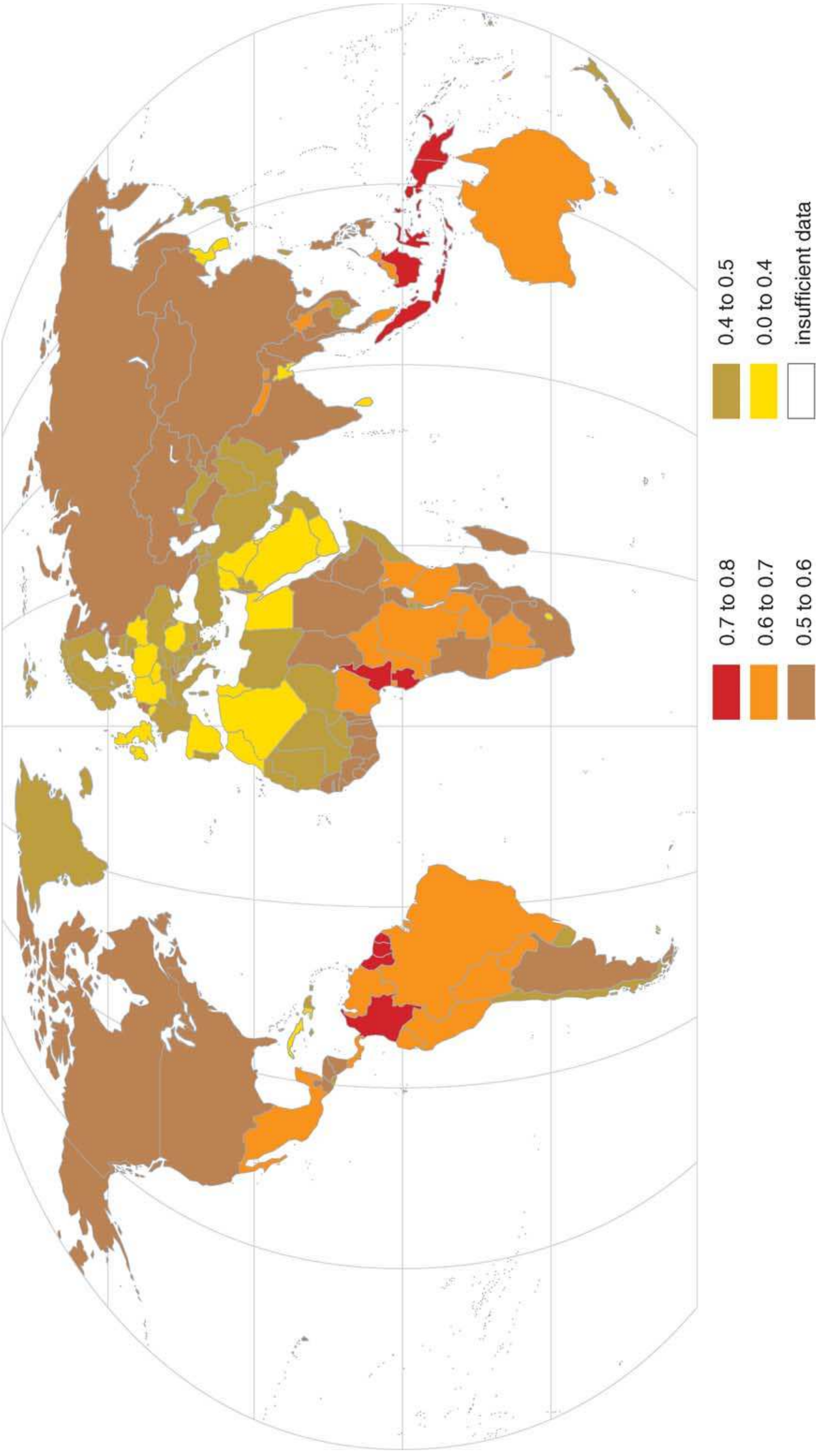
Fig. 2. Languages—area plot.



Map 1. Index of biocultural diversity IBCD-RICH.



Map 2. Index of biocultural diversity IBCD-AREA.



Map 3. Index of biocultural diversity IBCD-POP.

able to get at the critical question of the rate of change of the world's biocultural diversity.

References

Barrett, D.B., Kurian, G.T., Johnson, T.M., 2001. World Christian Encyclopedia: A Comparative Survey of Churches and Religions in the Modern World, 2nd ed. Oxford University Press, Oxford.

Davis, S.D., Heywood, V.H., Hamilton, A.C., 1994. Centres of Plant Diversity: A Guide and Strategy for their Conservation. IUCN—The World Conservation Union and World Wildlife Fund, Gland, Switzerland.

Gray, M., 2004. Geodiversity: Valuing and Conserving Abiotic Nature. John Wiley & Sons, Chichester.

FAO—Food and Agriculture Organization of the United Nations, 2004. FAOSTAT (FAO statistical databases). FAO, Rome, Italy. http://apps.fao.org (accessed February 2004).

Grimes, B.F. (Ed.), 2000. Ethnologue. 14th ed. Languages of the World, vol. 1. SIL International, Dallas.

Groombridge, B., Jenkins, M.D., 2002. World Atlas of Biodiversity: Earth's Living Resources in the 21st Century. University of California Press, Berkeley.

Harmon, D., 2002. In Light of Our Differences: How Diversity in Nature and Culture Makes Us Human. Smithsonian Institution Press, Washington, DC.

Harmon, D., J. Loh., 2004. A Global Index of Biocultural Diversity: Discussion Paper for the International Congress on Ethnobiology, University of Kent, UK, June 2004. Washington, DC: Terralingua.

Maffi, L., in press. Linguistic, cultural, and biological diversity. Annual Review of Anthropology 34.

Maffi, L. (Ed.), 2001. On Biocultural Diversity: Linking Language, Knowledge, and the Environment. Smithsonian Institution Press, Washington, DC.

Manne, L.L., 2003. Nothing has yet lasted forever: current and threatened levels of biological diversity. Evolutionary Ecol. Res. 5, 517–527.

Moore, J.L., Manne, L., Brooks, T., Burgess, N.D., Davies, R., Rahbek, C., Williams, P., Balmford, A., 2002. The distribution of cultural and biological diversity in Africa. In: Proceedings of the Royal Society of London B, vol. 269. pp. 1645–1653.

Myers, N., Mittermeier, R., Mittermeier, C.G., de Fonseca, G.A.B., Kent, J., 2000. Biodiversity hotspots for conservation priorities. Nature 403, 853–858.

Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C., D'amico, J.A., Itoua, I., Strand, H.E., Morrison, J.C., Loucks, C.J., Allnutt, T.F., Ricketts, T.H., Kura, Y., Lamoreux, J.F., Wettengel, W.W., Hedao, P., Kassem, K.R., 2001. Terrestrial ecoregions of the world: a new map of life on Earth. BioScience 51 (11), 933–938.

Stattersfield, A.J., Crosby, M.J., Long, A.J., Wege, D.C., 1998. Endemic bird areas of the World: priorities for biodiversity conservation. In: Birdlife Conservation Series No. 7unknown:-book, BirdLife International, Oxford.

Sutherland, W.J., 2003. Parallel extinction risk and global distribution of languages and species. Nature 423, 276–279.

The Times, 2000. Comprehensive Atlas of the World, 10th ed. (revised) Times Books, London.

Monitoring Change in Vertebrate Abundance: the Living Planet Index

Ben Collen, Jonathan Loh, Sarah Whitmee, Louise McRae, Rajan Amin and Jonathan E.M. Baillie

Conservation Biology (2009) **23**(2), 317-327

Conservation Biology

Contributed Paper

Monitoring Change in Vertebrate Abundance: the Living Planet Index

BEN COLLEN,*§ JONATHAN LOH,*† SARAH WHITMEE,* LOUISE McRAE,* RAJAN AMIN,‡ AND JONATHAN E. M. BAILLIE*‡

*Institute of Zoology, Zoological Society of London, Regent's Park, London NW1 4RY, United Kingdom
†WWF International, Avenue du Mont-Blanc CH-1196, Gland, Switzerland
‡Conservation Programmes, Zoological Society of London, Regent's Park, London NW1 4RY, United Kingdom

[Correction added after publication 18 November 2008: Errors in the third author's name and the fifth author's affiliation were amended.]

Abstract: *The task of measuring the decline of global biodiversity and instituting changes to halt and reverse this downturn has been taken up in response to the Convention on Biological Diversity's 2010 target. It is an undertaking made more difficult by the complex nature of biodiversity and the consequent difficulty in accurately gauging its depletion. In the Living Planet Index, aggregated population trends among vertebrate species indicate the rate of change in the status of biodiversity, and this index can be used to address the question of whether or not the 2010 target has been achieved. We investigated the use of generalized additive models in aggregating large quantities of population trend data, evaluated potential bias that results from collation of existing trends, and explored the feasibility of disaggregating the data (e.g., geographically, taxonomically, regionally, and by thematic area). Our results show strengths in length and completeness of data, little evidence of bias toward threatened species, and the possibility of disaggregation into meaningful subsets. Limitations of the data set are still apparent, in particular the dominance of bird data and gaps in tropical-species population coverage. Population-trend data complement the longer-term, but more coarse-grained, perspectives gained by evaluating species-level extinction rates. To measure progress toward the 2010 target, indicators must be adapted and strategically supplemented with existing data to generate meaningful indicators in time. Beyond 2010, it is critical a strategy be set out for the future development of indicators that will deal with existing data gaps and that is intricately tied to the goals of future biodiversity targets.*

Keywords: biodiversity decline, Convention on Biological Diversity, Living Planet Index, population time series, species abundance, threat indicators, 2010 target

Monitoreo del Cambio en la Abundancia de Vertebrados: El Índice del Planeta Viviente

Resumen: *La tarea de medir la declinación global de la biodiversidad y la institución de cambios para detener y revertir este decrecimiento se ha adoptado en respuesta al objetivo 2010 de la Convención de Diversidad Biológica. Es un cometido que se hace más difícil por la compleja naturaleza de la biodiversidad y la consecuente dificultad de evaluar su reducción con precisión. En el Índice del Planeta Viviente, las tendencias poblacionales agregadas entre especies de vertebrados indican una tasa de cambio en el estatus de la biodiversidad, y este índice puede ser utilizado para responder la pregunta sí el objetivo 2010 se ha alcanzado o no. Investigamos el uso de modelos aditivos generalizados para agregar grandes cantidades de datos de tendencias poblacionales, y exploramos la factibilidad de desagregar los datos (e.g., geográficamente, taxonómicamente, regionalmente y por área temática). Nuestros resultados muestran fortaleza en el detalle y compleción de los datos, poca evidencia de sesgo hacia especies amenazada y la posibilidad de desagregación en subconjuntos significativos. Las limitaciones de los datos aun son aparentes, en particular la dominancia*

---

§*email ben.collen@ioz.ac.uk*

Paper submitted October 8, 2007; revised manuscript accepted August 6, 2008.

Conservation Biology, Volume 23, No. 2, 317–327

DOI: 10.1111/j.1523-1739.2008.01117.x

de datos de aves y vacíos en la cobertura de poblaciones de especies tropicales. Los datos de tendencias poblacionales complementan el largo plazo, pero se obtienen perspectivas de grano más grueso mediante la evaluación de tasas de extinción a nivel de especies. Para medir el progreso hacia el objetivo 2010, los indicadores deben ser adaptados y estratégicamente suplementados con datos existentes para generar indicadores significativos a tiempo. Más allá de 2010, es crítico que se establezca una estrategia para el futuro desarrollo de indicadores que aborden los vacíos de datos y que esté intrincadamente ligada a las metas de objetivos futuros de biodiversidad.

Palabras Clave: Convención de la Diversidad Biológica, declinación de la biodiversidad, indicadores de amenaza, Índice del Planeta Viviente, objetivo 2010, series de tiempo poblacionales

Introduction

There is little prospect of effectively reducing global biodiversity loss unless trends in the state of biodiversity, and human impact on it, can be measured first. In 2002 188 nations signed on to the Convention on Biological Diversity (CBD) 2010 target of "achieving . . . a significant reduction of the current rate of biodiversity loss" (UNEP 2002). Conservation scientists must devise ways to determine whether or not the CBD 2010 target has been met. Nevertheless, the task is complex. *Biodiversity* is a multifaceted term, referring to the sum total of all biotic variation from the level of genes to ecosystems (Groombridge 1992). The sheer complexity of biological diversity means that a single measure cannot describe it or track its change. In response to the rapid rate of biological degradation seen over the past 50 years (Millennium Ecosystem Assessment 2005), the CBD has established a framework of 7 focal areas measured by 22 headline indicators to assess trends in biodiversity and ecosystem function. These indicators are intended to be rigorous, repeatable, have broad acceptance, and be easily communicated (European Academies Science Advisory Council 2005; UNEP 2006). The reporting process of these headline indicators is made more powerful by incorporating independent scientific assessment of the strategy employed (Dobson 2005). It is important that all of the CBD headline indicators follow this model. We sought to evaluate how well one of these indicators, the Living Planet Index (LPI), is achieving these aims.

The Living Planet Index

At the 8th Convention of the Parties, the LPI was adopted as one of the potential measures to address the CBD headline indicator: change in abundance of selected species (UNEP 2006). The LPI is based on what is believed to be one of the largest time-series databases on vertebrate populations, and it provides a broad range of vertebrate-population trend indicators. The LPI began life as a communications tool for a World Wildlife Fund (WWF) campaign. One of its biggest assets is that it is a simple yet powerful way of conveying information about changing trends in biodiversity to nonexperts, from policy and decision makers to the general public. Because of its developing role as a policy tool in monitoring progress toward the 2010 biodiversity target, it is becoming evermore important that the indicator is as robust, sensitive, and unbiased as possible. Ideally an indicator measuring change in population abundance would measure a randomly selected representative subset of taxa stratified across the main habitat types for which one would like information. These data do not exist. One form of recourse is to use available data on monitored populations to generate an indicator of population trends. Long-term data have much to offer current conservation efforts (Willis et al. 2007), but like many other species-based indicators, the LPI relies on compiling data collected for a range of different purposes. To ensure a robust and meaningful indicator, discrepancies in representation must be accounted for and minimized.

To assess how well this indicator informs the 2010 target, one must consider the feasibility of disaggregating the data (e.g., geographically, taxonomically, regionally, and by thematic area). This enhances the utility of the indicator and provides more targeted insight into patterns of biodiversity change. We explored the potential of augmenting population trend data to allow disaggregation and present new population-trend indices that allow finer-scale resolution of trends. There are advantages and limitations to the use of this method to track changing trends in vertebrate biodiversity. We set out a strategy for the future development of this key biodiversity indicator.

Methods

Data Collection

We collated time-series information for vertebrate species from published scientific literature, on-line databases (e.g., NERC Centre for Population Biology 1999 [Global Population Dynamics Database]; Pan-European Common Bird Monitoring Scheme 2006), and gray literature. Data were only included if a measure of population size was available for at least 2 years; information was available on how the data were collected and what the units of measurement were; the geographic location of the population was provided; the data were collected using the same method on the same population throughout the time series; and the data source was referenced and traceable.

We collated additional information on each data set so we could generate disaggregated indices. For each population data set we collected data on the system the population spent the majority of its time in and bred in (terrestrial, freshwater, marine). For terrestrial and freshwater species which biogeographic realm the population inhabited was recorded (Afrotropical, Australasian, Indo-Malayan, Nearctic, Neotropical, or Palearctic), and for marine species which ocean basin the population occupied (Arctic/Atlantic, Indian Ocean, Pacific Ocean, Southern Ocean) was recorded. We assigned terrestrial and freshwater species to either tropical or temperate regions and a habitat type (following the IUCN 2007*b* classification scheme: forest, savanna, shrubland, grassland, wetland, desert, artificial) and recorded its taxonomic class.

A quality score was generated for each time series by combining several aspects of the study: type of source (3, journal article; 2, government report or secondary source; 1, expert judgement or unpublished report; 0, unknown), type of method (3, full population count, index, density measure, or measure per unit effort; 2, estimate; 1, proxy; 0, unknown), and whether or not a measure of variation was calculated (1, yes; 0, no). We summed these scores to determine the quality score, which could range from 0 (lowest quality) to 7 (highest quality) for each time series. Time series with a score from 0 to 4 were considered poor quality, and those with scores from 5 to 7 were considered high quality.

Calculating Index Values

We used 2 different methods to generate index values: a chain method (Loh et al. 2005) and a generalized additive modeling technique (Fewster et al. 2000; Buckland et al. 2005). We followed Loh et al. (2005) to implement the chain method; however, we calculated the logarithm of the ratio of population measure for successive years (d), rather than for 5 yearly intervals:

$$d_t = \log_{10}(N_t/N_{t-1}), \tag{1}$$

where N is the population measure and t is the year. One percent of the mean population measure value for the whole time series was added to all years in time series for which N was zero in any year. Missing values were imputed with log-linear interpolation:

$$N_i = N_p(N_s/N_p)^{[(i-p)/(s-p)]}, \tag{2}$$

where i is the year for which the value is interpolated, p is the preceding year with a measured value, and s is the subsequent year with a measured value. For species with more than one time series, the mean value of d_t was calculated across all time series for that species. Species-specific values for d_t were combined:

$$\bar{d}_t = \frac{1}{n_t}\sum_{i=1}^{n_t} d_{it}. \tag{3}$$

The index value (I) was then calculated in year t as

$$I_t = I_{t-1}10^{\bar{d}_t}, \tag{4}$$

with the index value set to 1 in 1970. Insufficient data were available to continue the index beyond 2003 because of a lag in publication of data.

Time series with $n < 6$ were analyzed with the chain method. For all other time series, we implemented a generalized additive model (GAM), specified with the mgcv package framework in R (Wood 2006). For each time series, we

1. fitted a GAM on observed values with $\log_{10}(N_t)$ as the dependent variable and year (t) as the independent,
2. set the smoothing parameter to the length of the population time series divided by 2 (Wood 2006),
3. selected the smoothing-parameter value by comparing the estimated degrees of freedom when the smoothing parameter was successively incremented by 1,
4. used fitted GAM values to calculate predicted values for all years (including those with no real count data), and
5. averaged and aggregated d values from the imputed counts as described earlier.

A GAM framework might be advantageous in long-term trend analysis because it allows change in mean abundance to follow any smooth curve, not just a linear form (Fewster et al. 2000). The GAM method has greater flexibility for drawing out the long-term nonlinear trends that are generally not elicited in the discrete annual estimates of the chain method. We weighted the analysis, as described earlier, with species with more than one time series averaged across all the time series for that species. We combined specific values with geometric means at each time point and calculated the index.

To examine the effect of variation in data quality on index trajectory, the index for terrestrial species was calculated only with high-quality data, only low-quality data, and with all available time-series data. To examine the impact of equally weighting populations within species (regardless of what percentage of the global population they represented), we weighted each population by its relative share of the global population for that species for African mammals.

Weighting and Aggregation

In all indices except one, populations were weighted equally within species, and species were weighted equally within each index. Indices for terrestrial and freshwater systems were calculated as the geometric mean of tropical and temperate species $\bar{d}_t$ values. The marine index was calculated from the geometric mean of species $\bar{d}_t$ values from four ocean indices. To investigate the impact of different weighting, for the African mammal index, we estimated the proportion of the total global population that each population represents by comparing mean population size over the whole time series. For rapidly changing populations, we used the last available data point. Data were grouped into bins: <1, 1–25, 26–50, 51–75, and 76–100% of global population. It is not possible to gather these types of data for all time series in the data set; therefore, African mammals were used to investigate the effects of this different weighting strategy. Indices were calculated as earlier, but species-specific values of d_t were calculated by weighting according to the proportion of global population the time series represented.

Analyses

We used a bootstrap resampling technique to generate confidence limits around index values. To calculate a bootstrap replicate, for each interval, $t-1$ to t, a sample of n_t species-specific values of d_t was selected at random with replacement from the n_t observed values. For a given subindex within a system (tropical or temperate for terrestrial and freshwater, ocean for marine), we carried this out for each interval, and calculated $\bar{d}_t$ and I_t values as described earlier. We implemented the bootstrap procedure 10,000 times and used the bounds of the central 9,500 I values for each year to represent the 95% confidence interval for the index in that year (Loh et al. 2005). Setting the base year (1970) to unity did not mean there was no uncertainty associated with it, rather that the uncertainty was inherited by the rest of the values in the series.

Following Fewster et al. (2000), change points in the index were identified by using the bootstrap to identify time points at which the second derivative of the index differed significantly from zero. We identified years where the slope of the log-index value on year changed significantly, which identified where the annual rate of change as a proportion changed. We analyzed taxonomic representation of threatened status against coverage in the IUCN Red List of Threatened Species (IUCN 2007*a*). For each threat category, we used a binomial equality of proportions test (Crawley 2002) to compare the proportion of species in the LPI with the proportion of species on the IUCN Red List. All methods were implemented in R (version 2.5.1, R Development Core Team 2006).

Results

Trends in Vertebrate Populations

The 4218 vertebrate populations of 1411 species in the index showed a significant decrease in abundance between 1970 and 2003 (Fig. 1: 2003 index value 0.80; 95% CI 0.77–0.86). The differences between systems (Figs. 2a–c) broadly showed that abundances of terrestrial, freshwater, and marine species have been profoundly degraded over the past 3 decades and that tropical species have been more heavily affected than those in temperate regions.

We present 3 examples of disaggregation that may shed light on particular patterns of biodiversity degradation (for clarity, not all disaggregations are presented—see Supporting Information). Biogeographical disaggregations by realm (Fig. 3a) showed species in tropical realms such as the Afrotropics deteriorated faster than species in temperate realms such as the Nearctic. Disaggregated by biome (Fig. 3b), tropical forest populations declined at a greater rate than temperate forests. Taxonomic disaggregation within habitat (Fig. 3c) showed the abundance of grassland birds maintained a relatively flat trend over the past 3 decades in comparison with grassland mammals, which declined more rapidly.

Of the aggregated system-level indices (terrestrial, freshwater, marine), only the terrestrial index showed significant change points, the most recent 2 of which (2000 and 2002) showed a significant negative change in slope.

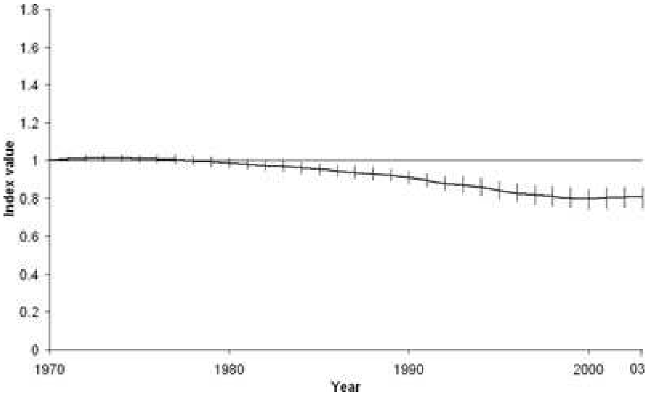


Figure 1. Average global Living Planet Index values for the 3 component indices (terrestrial, freshwater, and marine; 1411 species; 4218 populations). Confidence intervals (vertical lines) for the index are 95% values around the mean, generated with 10,000 bootstrap replicates (see Methods).

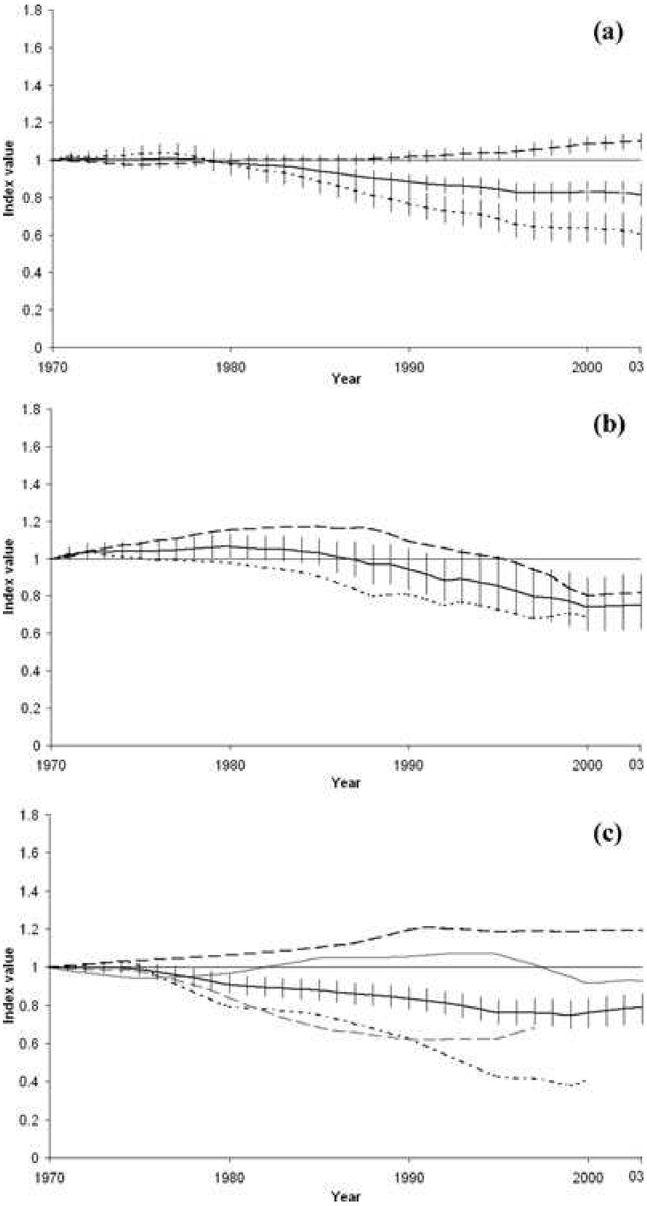


Figure 2. System-level (terrestrial, freshwater, and marine) indices for (a) temperate terrestrial (dashed), tropical terrestrial (dotted), and global terrestrial (solid) vertebrate species (739 species, 1585 populations), (b) temperate freshwater (dotted), tropical freshwater (dashed), and global freshwater (solid) vertebrate species (375 species, 1442 populations), and (c) Atlantic Ocean (heavy dashed), Pacific Ocean (light solid), Southern Ocean (light dashed), Indian Ocean (dotted), and global marine (heavy solid) vertebrate species (297 species, 1191 populations). Crosses in (a) are significant negative change points. Confidence intervals (vertical lines) are 95% values around the mean, generated with 10,000 bootstrap replicates (see Methods). For all CIs see Supporting Information.

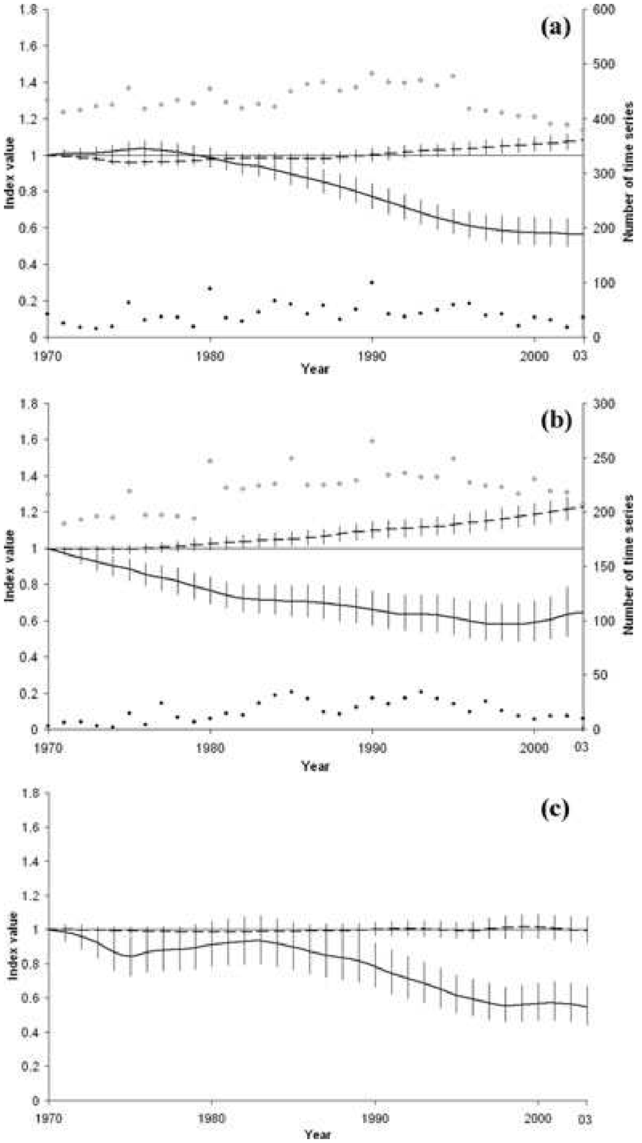


Figure 3. Living Planet indices for (a) Nearctic terrestrial (dashed) and Afrotropical terrestrial (solid) realms, (b) temperate forest (dashed) and tropical forest (solid) biomes, and (c) grassland birds (dashed) and grassland mammals (solid). Scatter plots ([a] and [b]) are the number of species contributing to the index in each year (right-hand axis) (filled circles go with the solid line and open circles with the dotted line). Confidence intervals (vertical lines) are 95% values around the mean, generated with 10,000 bootstrap replicates (see Methods).

Potential Weaknesses in Indices of Abundance

The data we used to construct the LPI were a matrix of 34 years by 4,218 populations: 143,412 potential data points. Nevertheless, not all time series started in 1970 and ended in 2003, so there were only 77,386 potential

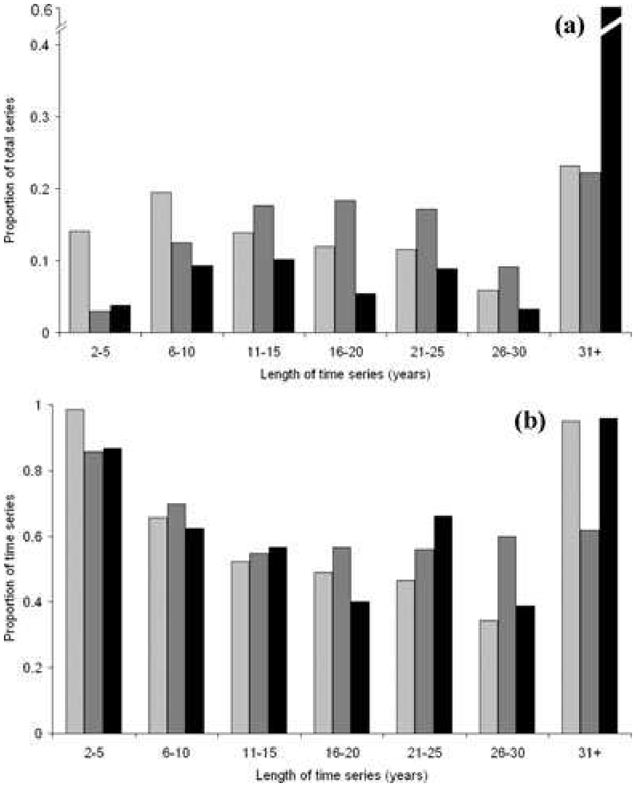


Figure 4. Distribution of length of time-series data against (a) proportion of all time-series data for each system in the Living Planet Index and (b) proportion of time-series data with noninterpolated data for populations in freshwater (light gray), marine (dark gray), and terrestrial (black) systems (4218 populations).

data points. The index contained values for 55,852 of these potential data points, a little over 72%, because not all data sets contained data for every year it covered. The missing 28% in this matrix were interpolated with the chain method or were imputed from the fitted statistical model in the GAM approach; none were extrapolated.

There are 3 key aspects of any given time series that might affect overall trends: series length (i.e., number of years from start to finish), series fullness (i.e., proportion of data points in the series that had measured [noninterpolated] values), and one as a function of the other. The majority of population time series collected covered around 20 years (Fig. 4a). The mean time-series length for marine populations was 19.88 years (SE 0.27; $n = 1193$), terrestrial populations 26.70 years (SE 0.25; $n = 1518$), and freshwater populations 17.89 years (SE 0.33; $n = 1442$). As well as being long, most time series were relatively complete. The most complete time series were also some of the longest, particularly in terrestrial and freshwater systems (Fig. 4b).

Indices accurately represented threat levels of some but not all vertebrate taxa (Table 1). Birds, amphibians, and reptiles appeared accurately represented compared with their representation in the IUCN Red List, although for the latter group, red-list coverage was highly incomplete, covering just 664 of the estimated 8240 species (IUCN 2007). Of the 289 mammals, species categorized as critically endangered and endangered were slightly overrepresented (Table 1: $\chi^2 = 9.29, p < 0.01$ and $\chi^2 = 16.57, p < 0.001$, respectively), whereas mammals categorized as least concern were slightly underrepresented ($\chi^2 = 17.12, p < 0.001$). Fish were underrepresented in all threat categories (Table 1), although coverage of the red list was highly incomplete, covering only 2899 of the estimated 29,300 species (IUCN 2007).

Relative to the species richness of the temperate and tropical regions, there were many more data from temperate regions than tropical regions in the data set. To counter this geographic bias, we gave aggregated trends for tropical and temperate species' populations equal weight in the calculation of the global index. Although coverage in temperate regions was extensive (Fig. 5) and certain tropical regions (e.g., eastern and southern Africa) had good coverage, there were some clear data gaps. This was particularly the case in South America, but it was also the case in some places where better coverage might have been expected (e.g., Australia). The distribution was similar for terrestrial and freshwater population coverage. Marine data tended to be clustered at the coast, with relatively few deep-ocean populations represented.

For terrestrial species, where data were most complete, the ratio of threatened to nonthreatened species varied little. Although there was a weak effect of year on population total (Table 2), which one would expect from a declining index, it was not significantly different between threatened and nonthreatened species (i.e., there were not more threatened species toward the end of the index [see Supporting Information]).

Data sets showing declining trends might be published more quickly and so bias the index to show a negative trend toward the present. There was no significant difference in publication date between the declining time series and nondeclining time series (Wilcoxon rank-sum test $Z = -1.33, p > 0.05$). There was also no significant difference in whether different publication types (journal article or gray or secondary literature) reported declining time series ($\chi^2 = 1.85$, df $= 1, p > 0.05$).

With the exception of poor-quality data for temperate terrestrial populations, good-quality data exhibited similar trends as poor-quality data and all data (see Supporting Information). Nevertheless, the index generated with temperate poor-quality data had a qualitatively different trajectory (increasing trend, rather than a zero trend line) than either good-quality or all temperate data. When indices for poor-quality temperate bird and

Table 1. The proportion of species in each IUCN Red List category in the Living Planet Index (LPI) and IUCN Red List of Threatened Species (IUCN 2007*a*) for 5 vertebrate classes.

Taxon	*Category*[a]	*LPI*	*IUCN*	χ^2 *value*[b]	*Representation*[c]
Mammals[d]	CR	0.07	0.03	9.29**	over
	EN	0.14	0.07	16.57***	over
	VU	0.14	0.12	0.72	—
	NT or LR/cd	0.19	0.14	5.61*	over
	LC	0.42	0.55	17.13***	under
	total no. spp.	289	4864		
Birds[d]	CR	0.02	0.02	0.02	—
	EN	0.03	0.04	0.004	—
	VU	0.06	0.07	0.4	—
	NT	0.04	0.08	14.14***	under
	LC	0.83	0.78	11.41***	over
	total no. spp	817	9934		
Reptiles	CR	0.20	0.11	2.31	—
	EN	0.15	0.15	0.00	—
	VU	0.20	0.25	0.45	—
	NT or LR/cd	0.11	0.15	0.37	—
	LC	0.13	0.20	1.06	—
	total no. spp.	46	664		
Amphibians[d]	CR	0.05	0.07	0.35	—
	EN	0.04	0.07	1.00	—
	VU	0.05	0.11	2.02	—
	NT	0.09	0.06	0.52	—
	LC	0.63	0.38	19.18***	over
	total no. spp	79	5918		
Fishes	CR	0.01	0.09	17.63***	under
	EN	0.04	0.08	5.87*	under
	VU	0.06	0.23	41.66***	under
	NT or LR/cd	0.00	0.08	20.90***	under
	LC	0.08	0.28	50.74***	under
	total no. spp.	264	2899		

[a] *Abbreviations: CR, critically endangered; EN, endangered; VU, vulnerable; NT, near threatened; LR/cd, lower risk/conservation dependent; LC, least concern.*

[b] *Chi-square calculated from a binomial test for equality of proportions (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$).*

[c] *Whether the group is over- or underrepresented for the red-list category.*

[d] *Group comprehensively assessed by IUCN.*

mammal data were plotted separately, they showed qualitatively similar results (not reported), so it is unlikely that a certain species group was driving this upward trend.

For African terrestrial mammals, weighting each population by its relative share of the global population did not significantly alter the trajectory of the index (see Supporting Information).

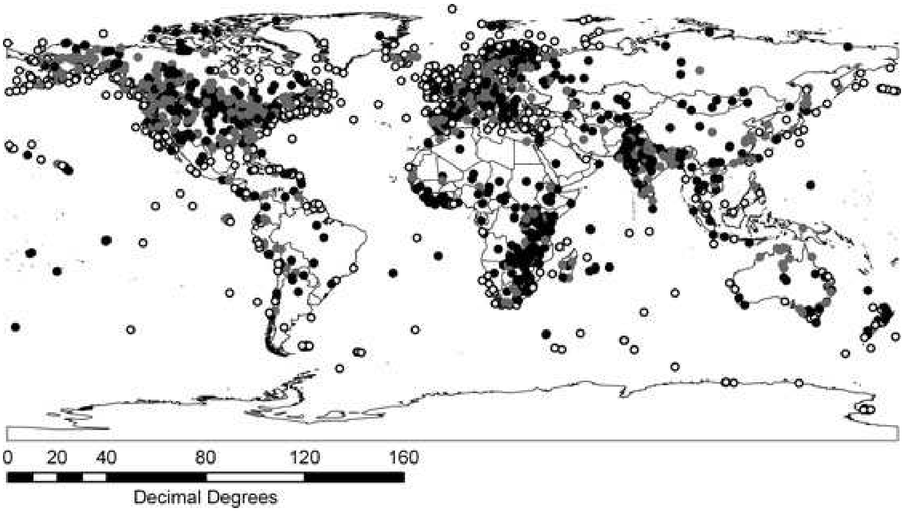


Figure 5. Geographic distribution of terrestrial (black), freshwater (gray), and marine (white) populations included in the Living Planet Index.

Table 2. Effect of year on population total for terrestrial mammals, accounting for threat status.

	df	*Sum of squares*	*Mean square*	F *ratio*
Threat status[a]	1	1.8×10^{-25}	1.8×10^{-25}	2.35×10^{-27}
Total[b]	1	1026.4	1026.4	13.40 ($p < 0.01$)
Threat status: total	1	209.5	209.5	2.74
Residuals	62	4748.1		

[a] *Threat status is binary (threatened or other; see Methods).*
[b] *Total is total number of terrestrial populations in any given year.*

Discussion

Rate of Abundance Loss

Ever increasing levels of human impact and resource consumption are driving many species toward extinction. More important, before they reach this point, local abundance of species is declining rapidly, limiting the ability of ecosystems to maintain ecological functions and services. Our study reveals an overall decline in vertebrate population abundance between 1970 and 2003. The negative change in populations in terrestrial systems since 1970 is underpinned by a more rapid rate of decline in populations of tropical species and a more stable overall trend for temperate species. Nevertheless, the data sets we examined did not extend back far enough to detect the much larger declines that likely occurred in temperate systems before 1970 (Millennium Ecosystem Assessment 2005). Similar patterns were apparent in freshwater systems, where populations of tropical species declined in abundance at a much greater rate than temperate populations. In the marine system, relatively stable trends were evident in the Pacific, Arctic, and Antarctic oceans, in comparison with the rapid declines observed in the Indian, Southeast Asian, and Southern oceans.

It is paramount that this index be interpreted in relation to the CBD's 2010 target: a decrease in rate of decline. The LPI is based on a proportional change in abundance measure. Therefore, a significant change showing the slope becoming less negative before 2010 would indicate the target has been met. Nevertheless, any downward trend, even if diminishing, would show populations in decline. Currently, there is no indication from population-abundance data that the 2010 target will be met at the global level.

Limitations of Abundance Data

It is useful to imagine a perfect LPI, which would aggregate global population time-series data for every species in the world. A representative subset of the perfect LPI would show the same trend if the selected populations remained representative of total populations. It is important to understand the areas in which the LPI falls short of these standards to provide a focus for improvement. Large quantities of data on population abundance in this study were of long and full time series that required little interpolation. Nevertheless, this reflects just one aspect of data quality.

There was little bias toward threatened species. Nevertheless, the taxonomic scope of data coverage was restricted to vertebrates; specifically, it was predominated by birds and mammals, with disproportionately fewer amphibian, reptile, and fish populations. How vertebrate trends relate to that of broader biodiversity is still largely unknown. Even so, the method is applicable to any data that meet the selection criteria, and work has begun to extend the approach to include plants, a group that currently lacks meaningful trend information, aside from small subsets (e.g., cycads and conifers; IUCN 2007*a*).

The utility of the LPI method has been questioned on the grounds of geographic coverage (Pereira & Cooper 2006). By weighting tropical and temperate species equally, geographic imbalance is addressed to some extent in the index calculation. As with all indicators of biodiversity change, the LPI relies on available data, which are inevitably more abundant in better-studied regions. Our results showed that although population trends were underrepresented in the tropics, global coverage was relatively good. A more balanced geographic sampling that enhances reliability and ensures biases are further reduced can be achieved if the network of data providers increases and if focus is placed on data-poor regions. Nonetheless, the LPI will always be restricted to those populations that happen to have been studied.

These issues highlight concerns over indicators of biodiversity change that are more wide reaching. The factor that most limits progress in understanding biodiversity change is that many of the existing data relate to species from temperate regions, whereas the majority of biodiversity is found in the tropics. Although long-term data collection in these understudied regions is becoming more widespread (e.g., Whitfield et al. 2007), most collecting efforts will not be in time for 2010. This restricts the ability to make robust statements about the global-decline dynamics of species.

Potentially the biggest limitation to producing robust indicators of change in population abundance may result from variation in the nature of underlying data. A change in type of data collected over time (e.g., from large, wide-ranging, stable populations to small populations of conservation concern) could artificially result in a declining trend toward the present. Nevertheless, the advent of conservation biology in the late 1970s and early 1980s was born out of concern over human impact on the environment. It is almost impossible to separate out a change in study focus from genuine change in population status. One possibility is to examine publication type. Publication bias can result whenever the probability that a study will be published depends on the significance of its results, the "file drawer problem" (Rosenthal 1979). The time series in the LPI were not published more quickly if they showed declining species abundance, and broadly speaking, declining populations were not disproportionately documented in peer-reviewed journals.

A criticism of the technique we used here is that all decreases in population size, regardless of whether they brought a population close to extinction, were accounted for equally (Pereira & Cooper 2006). From an ecosystem perspective, a decline in an abundant and widespread species is likely to be of greater importance than an equivalent decline in a rare endemic one of small population size. Nevertheless, a local decline in a small population of a widely distributed species is not as important as a global decline. On a related point, in our method, all populations are given equal weighting. An extreme example of possible impacts of this is that it is plausible that 2 populations, one containing 90% of the global species abundance, and the other just 10%, would get equal weighting in the index. From a subset of African mammals, for which we were able to obtain global population estimates, there was no significant difference in the indices calculated with equal and proportional weighting methods. It is our intention to extend this approach beyond this subset of species, once additional data are available.

Strengths of Abundance Data

That the LPI can be disaggregated into meaningful subsets is one of its greatest strengths. For example, grassland mammals appear to be declining at a higher rate than grassland birds. Possible mechanisms include greater hunting pressure and the effects of anthropogenic manipulation of grasslands (burning, conversion into agriculture), which affects mammals. Increasingly it appears that species groups may not all be reacting to anthropogenic pressures to the same extent (e.g., Thomas et al. 2004).

An effective indicator of biodiversity change should fit a number of both scientific and practical criteria. It should simplify information to be representative, quantitative, responsive to change, susceptible to analysis, policy relevant, and tractable (Gregory et al. 2005). Our method appears to posses many of these attributes, in particular it is sensitive to change, quantitative, policy relevant, and easy to communicate.

Measuring species' decline and extinction is not straightforward. Species extinction represents a coarse measurement of the loss of biodiversity because most species are poorly known; many species remain very rare for lengthy periods before becoming extinct; and scientists are reluctant to classify a species as extinct before being absolutely certain (Mace et al. 2003). Species extinction is also relatively insensitive to short-term change (Balmford et al. 2003). The disappearance of populations is a prelude to species-level extinctions, so perhaps measuring population reduction represents a more sensitive indicator of the loss of biodiversity (Ceballos & Ehrlich 2002). The time lag between human impact and a corresponding detectable negative trend in populations may be short enough to enable more proactive conservation action than that measured by species-level changes.

Informative Population Data

Overall population size is not always the best indicator of long-term population trend. For long-lived vertebrate species in particular, extinction debts can occur (e.g., when populations are subjected to a period of habitat degradation, but persist in a landscape for some time after degradation has made their eventual extinction inevitable) (Hanski & Ovaskainen 2002; Carroll et al. 2004). Detailed demographic information may tell us more about the fate of populations. Nevertheless, additional information on informative parameters other than population size are seldom collected, other than in certain well-developed monitoring programs (Katzner et al. 2007).

For taxa in which recruitment is more variable than survival, the relationship between population abundance and time may vary depending on the species (Alford & Richards 1999). Because of naturally high fecundity and mortality rates, certain species are likely to show noisier trends on annual timescales, but one would nevertheless expect to see longer-term trends emerge over decadal timescales given adequate monitoring. Although certain life-history traits and demographic parameters may better predict future population trends across taxa, these are currently not well understood, and the utility of any given trait is likely to vary both across taxa and under differing threatening processes, so general applicability of conclusions at a global scale might be limited. Despite the drawbacks and the difficulties associated with data collection, population abundance appears the only tangible option in the short term.

Strategy for the Future

Biodiversity is a multifaceted term, one which cannot be reduced to a simple number (Magurran 1988; Purvis & Hector 2000); therefore, following trends in abundance is just one of several important metrics that feed into a broader range of indicators. Across all monitoring programs and indicators of biodiversity loss, it is evident that a fraction of biological systems and the species and habitats integral to them are being monitored. In general, temperate terrestrial systems, birds, commercially important, and large species are monitored better than other species, simply because it is the most expedient thing to do. Regions are not evenly affected by anthropogenic pressures, and taxa do not respond in a uniform manner to these threats (Mace et al. 2003; Isaac & Cowlishaw 2004). If the changing state of biodiversity is to be measured accurately and address the 2010 target, this imbalance must be redressed.

Ultimately, population-trend data provide a complement to the longer-term, but more coarse-grained, perspectives gained by evaluating species-level extinction rates. Combined, they may offer a robust and broad view of the changing state of nature. The year 2010 is close, and its proximity invites remedial action and short-term fixes (Willis et al. 2007). Indicator development has necessarily taken an approach of adopting, adapting, and strategically supplementing existing indicators. Nevertheless, establishing long-term systematic monitoring programs and understanding the relationship between vertebrate population trends and the remainder of biodiversity in its broadest sense is integral to ensuring more useful and more policy-relevant indicators for 2010 and beyond.

We recommend 2 parallel processes of development. In the short term, the indicator will benefit from the systematic addition of time-series data, mined from the available literature to continue improving coverage. Targeted searches to fill the gaps identified here should be conducted. More representative coverage will then enable analytical development, refinement, validation, and stratification of the index.

In the medium to long term, the aim should be to implement a systematic stratified design, to address bias within the data set. After stratification of the data (e.g., by taxon within realm), recalculating the index would control for uneven coverage of strata. Data coverage of strata such as in the Neotropics and particularly for amphibian, reptile, and fish species, remains insufficient, therefore it is currently not feasible to stratify data for these data-poor areas. The gaps in the data we highlight here will serve to guide priorities and support monitoring of key species populations at unmonitored sites. These 2 concurrent strategies will improve the scope for analysis and tackle the issues raised in this study as part of the continued development of the LPI.

Acknowledgments

We acknowledge the input of J. Latham, S. Tranquilli, G. Mace, M. Ram, A. Greenwood, O. Daniel, and participants of the Global Biodiversity indicators workshop 2006 funded by Imperial College London and the Zoological Society of London. We thank A. Balmford, R. Green, and R. Gregory for their comments on this study. The Pan-European Bird Survey and Global Population Dynamics Database at Imperial College London Centre for Population Biology provided access to data. World Wildlife Fund International provided financial support.

Supporting Information

Proportion of threatened species over time (Appendix S1), disaggregations of data quality (Appendix S2), impact of population aggregation for African mammals (Appendix S3), and index values for all disaggregations (Appendix S4) are available as part of the on-line article. The authors are responsible for the content and functionality of these materials. Queries (other than absence of the material) should be directed to the corresponding author.

Literature Cited

Alford, R. A., and S. J. Richards. 1999. Global amphibian declines: a problem in applied ecology. Annual Review of Ecology and Systematics **30:**133–165.

Balmford, A., R. E. Green, and M. Jenkins. 2003. Measuring the changing state of nature. Trends in Ecology & Evolution **18:**326–330.

Buckland, S. T., A. E. Magurran, R. E. Green, and R. M. Fewster. 2005. Monitoring change in biodiversity through composite indices. Philosophical Transactions of the Royal Society of London B **360:**243–254.

Carroll, C., R. F. Noss, P. C. Paquet, and N. H. Schumaker. 2004. Extinction debt of protected areas in developing landscapes. Conservation Biology **18:**1110–1120.

Ceballos, G., and P. R. Ehrlich. 2002. Mammal population losses and the extinction crisis. Science **296:**904–907.

Crawley, M. J. 2002. Statistical computing: an introduction to data analysis using S-Plus. John Wiley & Sons, Chichester, United Kingdom.

Dobson, A. 2005. Monitoring global rates of biodiversity change: challenges that arise in meeting the Convention on Biological Diversity (CBD) 2010 goals. Philosophical Transactions of the Royal Society of London B **360:**229–241.

European Academies Science Advisory Council. 2005. A user's guide to biodiversity indicators. The Royal Society, London.

Fewster, R. M., S. T. Buckland, G. M. Siriwardena, S. R. Baillie, and J. D. Wilson. 2000. Analysis of population trends for farmland birds using generalized additive models. Ecology **81:**1970–1984.

Gregory, R. D., A. van Strien, P. Vorisek, A. W. G. Meyling, D. G. Noble, R. P. B. Foppen, and D. W. Gibbons. 2005. Developing indicators for European birds. Philosophical Transactions of the Royal Society of London B **360:**269–288.

Groombridge, B. 1992. Global biodiversity: status of the world's living resources. Chapman & Hall, London.

Hanski, I., and O. Ovaskainen. 2002. Extinction debt at extinction threshold. Conservation Biology **16:**666–673.

Isaac, N. J. B., and G. Cowlishaw. 2004. How species respond to multiple extinction threats: evidence from Primates. Proceedings of the Royal Society of London B **271:**1135–1141.

IUCN (International Union for Conservation of Nature). 2007*a*. 2007 IUCN Red List of threatened species. IUCN, Gland, Switzerland. Available from www.iucnredlist.org (accessed March 2008).

IUCN (International Union for Conservation of Nature). 2007*b*. IUCN habitat classification scheme. IUCN, Gland, Switzerland. Available from http://www.iucn.org/about/work/programmes/species/red_list/resources/technical_documents/authority_files/index.cfm (accessed March 2008).

Katzner, T., E. J. Milner-Gulland, and E. Bragin. 2007. Using modelling to improve monitoring of structured populations: are we collecting the right data? Conservation Biology **21:**241–252.

Loh, J., R. E. Green, T. Ricketts, J. F. Lamoreux, M. Jenkins, V. Kapos, and J. Randers. 2005. The Living Planet Index: using species population time series to track trends in biodiversity. Philosophical Transactions of the Royal Society of London B **360:**289–295.

Mace, G. M., J. L. Gittleman, and A. Purvis. 2003. Preserving the tree of life. Science **300:**1707–1709.

Magurran, A. E. 1988. Ecological diversity and it measurement. Croom Helm, London.

Millennium Ecosystem Assessment. 2005. Ecosystems and human wellbeing: biodiversity synthesis. World Resources Institute, Washington, D.C.

NERC Centre for Population Biology. 1999. The global population dynamics database. Imperial College London.

Pan-European Common Bird Monitoring Scheme (PECBMS). 2006. European common bird index: population trends of European common birds 2005 update. European Bird Census Council, Prague.

Pereira, H. M., and H. D. Cooper. 2006. Towards the global monitoring of biodiversity change. Trends in Ecology & Evolution **21:**123–129.

Purvis, A., and A. Hector. 2000. Getting the measure of biodiversity. Nature **405:**212–219.

R Development Core Team. 2006. R: language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.

Rosenthal, R. 1979. The “file drawer problem” and tolerance for null results. Psychological Bulletin **86:**638–641.

Thomas, J. A., M. G. Telfer, D. B. Roy, C. D. Preston, J. J. D. Greenwood, J. Asher, R. Fox, R. T. Clarke, and J. H. Lawton. 2004. Comparative losses of British butterflies, birds, and plants and the global extinction crisis. Science **303:**1879–1881.

UNEP (United Nations Environment Programme). 2002. Report on the sixth meeting of the Conference of the Parties to the Convention on Biological Diversity (UNEP/CBD/COP/20/Part 2) Strategic Plan Decision VI/26 in CBD. UNEP, Nairobi.

UNEP (United Nations Environment Programme). 2006. Report on the eighth meeting of the Conference of the Parties to the Convention on Biological Diversity, CBD. UNEP, Nairobi.

Whitfield, S. M., K. E. Bell, T. Philippi, M. Sasa, F. Bolanos, G. Chaves, J. M. Savage, and M. A. Donnelly. 2007. Amphibian and reptile declines over 35 years at La Selva, Costa Rica. Proceedings of the National Academy of Sciences of the United States of America **103:**3165–3170.

Willis, K. J., M. B. Araujo, K. D. Bennett, B. Figueroa-Rangel, C. A. Froyd, and N. Myers. 2007. How can a knowledge of the past help to conserve the future? Biodiversity conservation and the relevance of long-term ecological studies. Philosophical Transactions of the Royal Society of London B **362:**175–186.

Wood, S. N. 2006. Generalized additive models: an introduction with R. Chapman & Hall/CRC, Boca Raton, Florida.

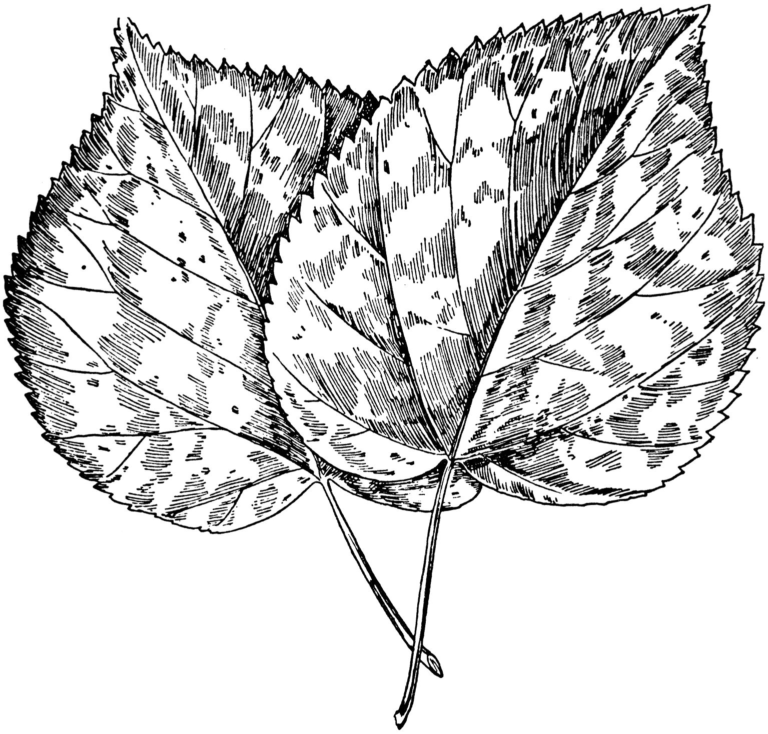


Figure S1.
Ratio of threatened to other (including DD and non-threatened) terrestrial species population totals for each year of the index. Dotted line is the mean ratio value over the whole index.

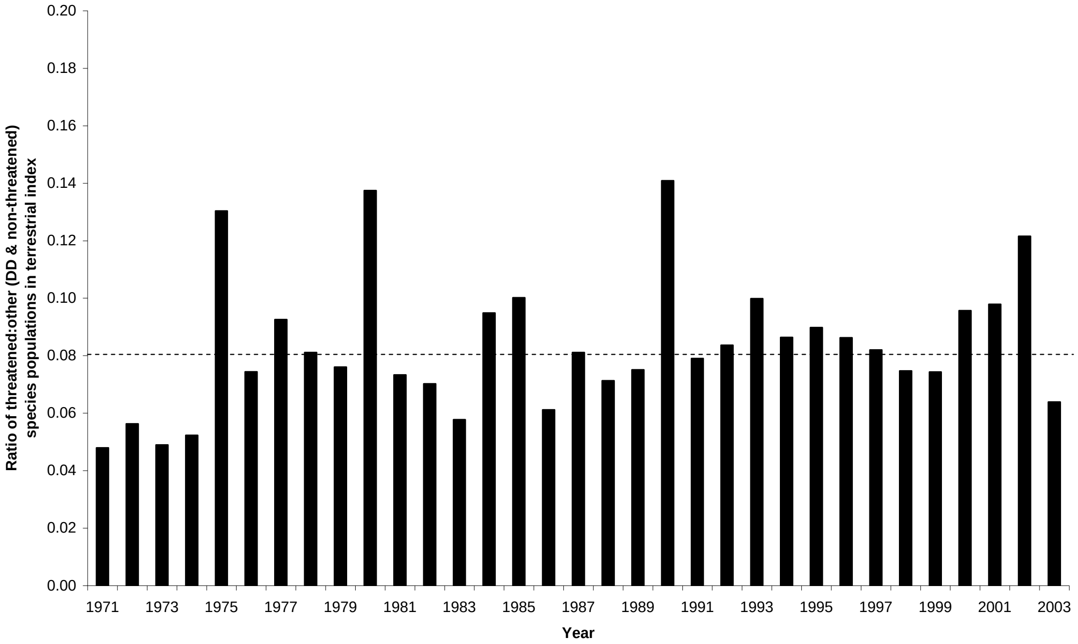
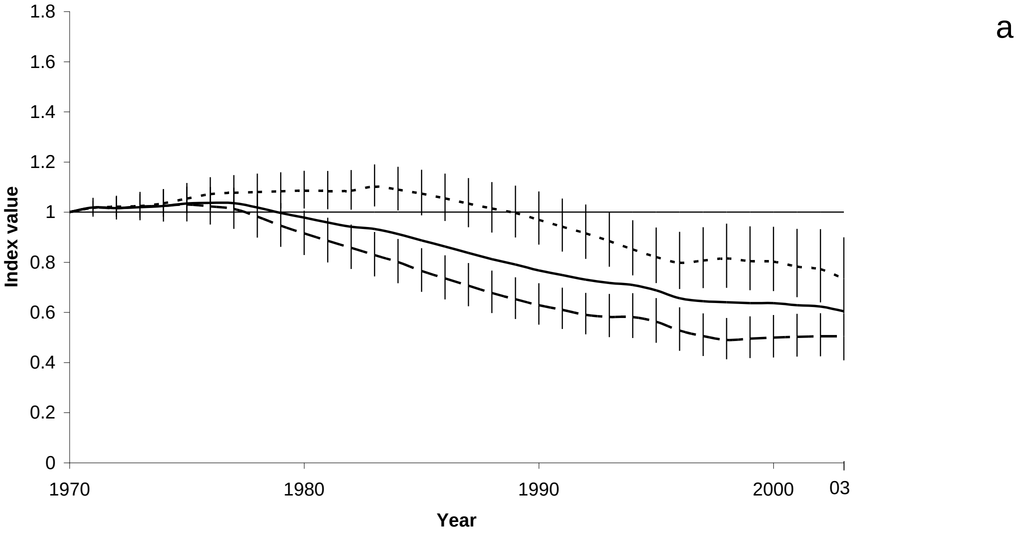
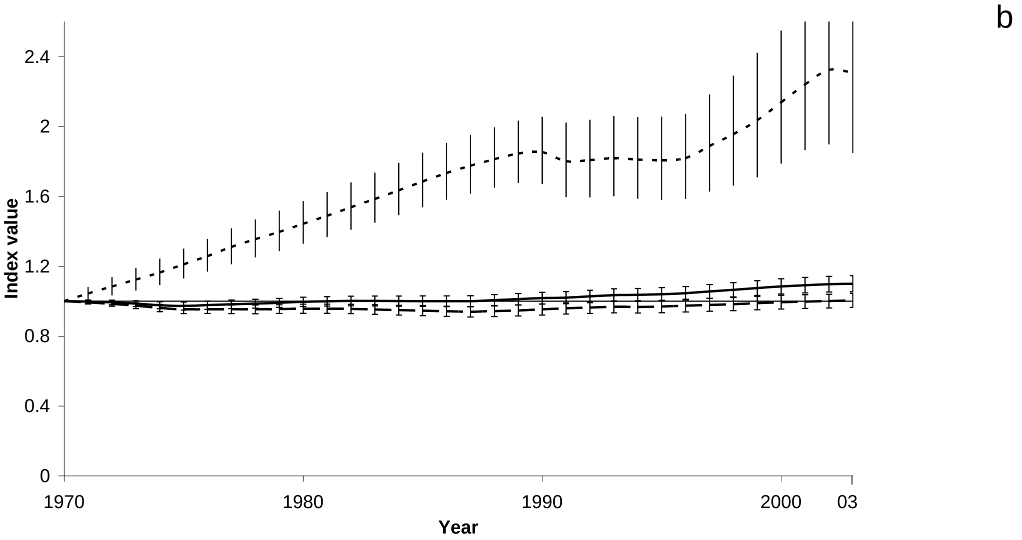


Figure S2.
Comparison of indices generated for terrestrial species in (a) tropical systems, b) temperate systems and c) aggregated into a terrestrial index. Each figure compares all data (solid line) with only good quality data (long dashed line) and only poorer quality data (short dashed line) – see Materials and Methods for detail on how quality score was generated. Confidence limits are 95% values around the mean of all data, generated using 10,000 bootstrap replicates (see Materials and Methods). To assess the impact of variation in data quality for terrestrial species, the combined index for each of the tropical terrestrial, and temperate terrestrial data sets is not an aggregation of good and poor quality data, rather all of the data given equal weight at the species level.


a
Index value
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
1970
1980
1990
2000
03
Year


b
Index value
2.4
2
1.6
1.2
0.8
0.4
0
1970
1980
1990
2000
03
Year

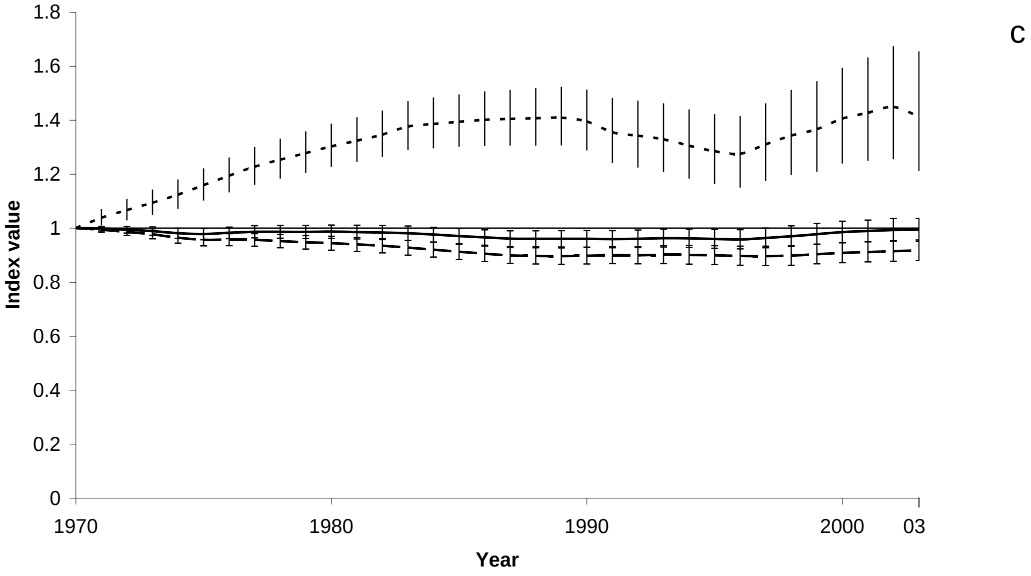

c
Index value
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
1970
1980
1990
2000
03
Year

Figure S3.
Indices of African terrestrial mammals using two different techniques for weighting populations. Equal weighting for each population = solid line, population weighted by relative contribution to global population = dashed line. N = 87 species, 288 populations. Confidence limits are 95% values around the mean, generated using 10,000 bootstrap replicates (see Materials and Methods).

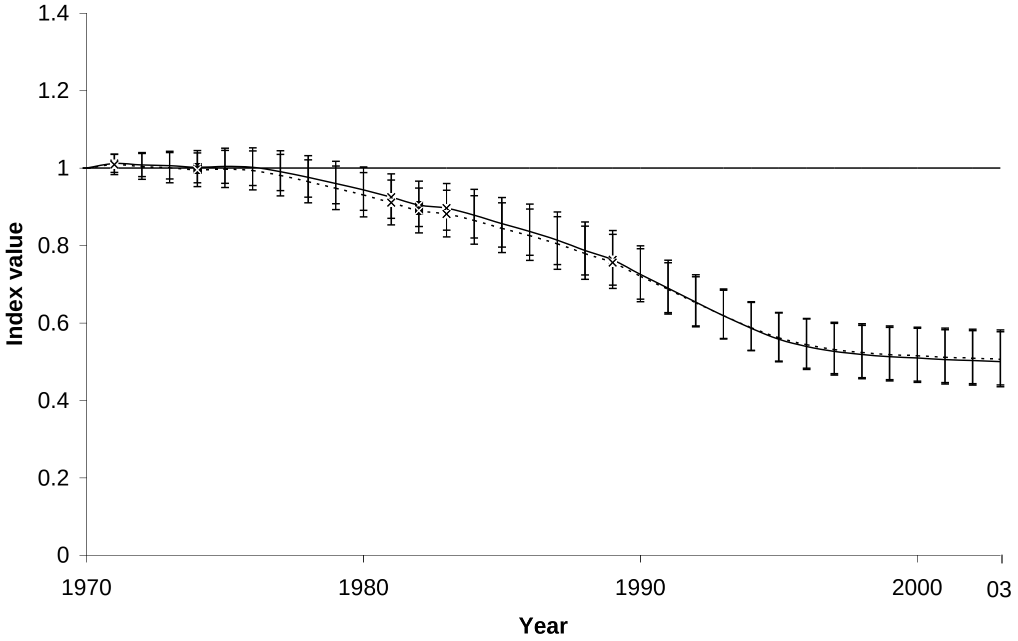


Table S4: Living Planet Index 1970-2003

Systems	**1970**	**1971**	**1972**	**1973**	**1974**	**1975**	**1976**	**1977**	**1978**	**1979**	**1980**	**1981**	**1982**	**1983**	**1984**	**1985**
Living Planet Index	1.000	1.007	1.011	1.011	1.010	1.008	1.006	1.003	0.999	0.993	0.985	0.978	0.973	0.968	0.960	0.952
lower 95% CI	1.000	0.994	0.994	0.992	0.990	0.986	0.983	0.977	0.972	0.964	0.955	0.948	0.943	0.937	0.928	0.919
upper 95% CI	1.000	1.024	1.031	1.032	1.032	1.032	1.032	1.031	1.028	1.023	1.016	1.009	1.005	1.002	0.994	0.986
Freshwater Temperate	1.000	1.010	1.038	1.055	1.071	1.081	1.097	1.107	1.127	1.145	1.155	1.158	1.165	1.170	1.168	1.174
lower 95% CI	1.000	0.995	1.014	1.027	1.041	1.048	1.060	1.065	1.080	1.096	1.103	1.102	1.106	1.109	1.106	1.108
upper 95% CI	1.000	1.024	1.071	1.089	1.107	1.120	1.139	1.153	1.178	1.198	1.211	1.219	1.229	1.236	1.236	1.245
Freshwater Tropical	1.000	1.025	1.033	1.020	1.010	1.000	0.993	0.990	0.987	0.985	0.978	0.964	0.952	0.942	0.926	0.907
lower 95% CI	1.000	0.970	0.962	0.946	0.931	0.921	0.909	0.900	0.894	0.887	0.877	0.862	0.848	0.838	0.820	0.801
upper 95% CI	1.000	1.118	1.138	1.127	1.118	1.111	1.104	1.105	1.106	1.106	1.102	1.092	1.080	1.071	1.056	1.037
Freshwater Index	1.000	1.019	1.035	1.037	1.040	1.040	1.043	1.047	1.054	1.062	1.063	1.056	1.051	1.049	1.039	1.032
lower 95% CI	1.000	0.990	0.996	0.995	0.996	0.994	0.994	0.994	0.999	1.004	1.003	0.994	0.988	0.984	0.973	0.965
upper 95% CI	1.000	1.065	1.087	1.091	1.096	1.097	1.104	1.110	1.121	1.131	1.134	1.129	1.125	1.124	1.115	1.110
Marine Atlantic	1.000	1.007	1.015	1.023	1.029	1.035	1.040	1.046	1.051	1.057	1.064	1.071	1.078	1.086	1.095	1.103
lower 95% CI	1.000	0.988	0.987	0.990	0.991	0.994	0.997	1.000	1.002	1.006	1.011	1.018	1.023	1.030	1.038	1.045
upper 95% CI	1.000	1.027	1.043	1.057	1.069	1.078	1.087	1.095	1.103	1.109	1.118	1.128	1.136	1.147	1.158	1.168
Marine Indian	1.000	1.002	1.011	1.021	1.030	1.010	0.959	0.911	0.874	0.832	0.796	0.786	0.779	0.771	0.762	0.751
lower 95% CI	1.000	0.971	0.966	0.965	0.968	0.936	0.858	0.798	0.757	0.714	0.676	0.666	0.656	0.646	0.635	0.624
upper 95% CI	1.000	1.038	1.061	1.080	1.099	1.087	1.060	1.030	0.998	0.960	0.929	0.920	0.916	0.910	0.904	0.893
Marine Pacific	1.000	0.984	0.972	0.960	0.950	0.943	0.947	0.951	0.955	0.961	0.966	0.983	1.001	1.017	1.032	1.048
lower 95% CI	1.000	0.957	0.937	0.919	0.904	0.894	0.892	0.893	0.894	0.897	0.899	0.912	0.926	0.938	0.948	0.959
upper 95% CI	1.000	1.011	1.008	1.002	0.998	0.996	1.006	1.013	1.021	1.032	1.041	1.060	1.084	1.105	1.124	1.144
Marine Southern	1.000	0.994	0.992	0.985	0.983	0.978	0.957	0.936	0.909	0.879	0.841	0.801	0.768	0.738	0.711	0.686
lower 95% CI	1.000	0.937	0.915	0.896	0.883	0.870	0.844	0.819	0.791	0.762	0.726	0.691	0.661	0.632	0.607	0.584
upper 95% CI	1.000	1.053	1.072	1.082	1.095	1.097	1.084	1.065	1.039	1.011	0.970	0.928	0.892	0.860	0.830	0.804
Marine Index	1.000	0.997	0.997	0.997	0.998	0.991	0.975	0.960	0.945	0.928	0.911	0.902	0.896	0.890	0.884	0.879
lower 95% CI	1.000	0.978	0.972	0.967	0.965	0.954	0.933	0.913	0.896	0.878	0.860	0.851	0.843	0.836	0.830	0.823
upper 95% CI	1.000	1.016	1.023	1.027	1.032	1.030	1.020	1.008	0.996	0.981	0.965	0.958	0.953	0.948	0.943	0.939
Terrestrial Index	1.000	1.008	1.005	1.003	1.001	1.003	1.007	1.009	1.002	0.994	0.987	0.979	0.971	0.966	0.955	0.942
lower 95% CI	1.000	0.995	0.987	0.982	0.978	0.977	0.979	0.979	0.970	0.961	0.953	0.944	0.935	0.929	0.918	0.903
upper 95% CI	1.000	1.021	1.023	1.024	1.025	1.029	1.035	1.038	1.034	1.028	1.022	1.015	1.008	1.004	0.995	0.982
Terrestrial Tropical	1.000	1.018	1.016	1.020	1.024	1.034	1.037	1.036	1.019	0.997	0.978	0.959	0.942	0.932	0.912	0.887
lower 95% CI	1.000	0.994	0.985	0.983	0.982	0.988	0.988	0.982	0.961	0.936	0.916	0.895	0.878	0.867	0.846	0.819
upper 95% CI	1.000	1.044	1.050	1.059	1.070	1.084	1.090	1.093	1.081	1.060	1.042	1.024	1.008	1.002	0.982	0.959
Terrestrial Temperate	1.000	0.998	0.993	0.987	0.978	0.973	0.978	0.982	0.986	0.991	0.997	0.999	1.002	1.001	1.001	1.000
lower 95% CI	1.000	0.988	0.979	0.969	0.956	0.949	0.953	0.957	0.960	0.964	0.970	0.972	0.974	0.973	0.971	0.970
upper 95% CI	1.000	1.006	1.005	1.003	0.997	0.995	1.000	1.007	1.011	1.016	1.022	1.026	1.029	1.030	1.030	1.030
WEIGHTED																
Terrestrial African Mammals Equal wt	1.000	1.012	1.008	1.006	1.002	1.005	1.002	0.991	0.976	0.960	0.944	0.925	0.905	0.897	0.878	0.857
lower 95% CI	1.000	0.989	0.978	0.972	0.962	0.961	0.955	0.942	0.925	0.908	0.891	0.870	0.849	0.840	0.820	0.796
upper 95% CI	1.000	1.037	1.040	1.043	1.045	1.051	1.052	1.045	1.032	1.018	1.003	0.985	0.966	0.960	0.945	0.924
Terrestrial African Mammals weighted	1.000	1.009	1.004	1.001	0.995	0.997	0.994	0.981	0.965	0.948	0.931	0.911	0.890	0.882	0.865	0.844
lower 95% CI	1.000	0.983	0.971	0.962	0.952	0.950	0.943	0.928	0.910	0.893	0.874	0.854	0.833	0.822	0.804	0.782
upper 95% CI	1.000	1.035	1.038	1.040	1.039	1.046	1.044	1.035	1.021	1.005	0.988	0.969	0.949	0.943	0.928	0.910
REALMS																
Terrestrial Afrotropical	1.000	1.011	1.011	1.013	1.019	1.030	1.033	1.026	1.015	1.002	0.987	0.967	0.947	0.939	0.918	0.895
lower 95% CI	1.000	0.987	0.981	0.979	0.980	0.988	0.988	0.977	0.964	0.948	0.933	0.913	0.891	0.882	0.859	0.834
upper 95% CI	1.000	1.035	1.043	1.050	1.060	1.075	1.082	1.078	1.069	1.057	1.044	1.026	1.006	1.002	0.985	0.963
Terrestrial Nearctic	1.000	0.995	0.988	0.979	0.967	0.958	0.961	0.963	0.965	0.970	0.976	0.981	0.986	0.985	0.985	0.981
lower 95% CI	1.000	0.985	0.974	0.961	0.945	0.935	0.937	0.938	0.940	0.944	0.949	0.954	0.958	0.956	0.954	0.949
upper 95% CI	1.000	1.003	1.000	0.995	0.987	0.980	0.984	0.986	0.990	0.995	1.002	1.008	1.013	1.014	1.015	1.013
QUALITY																
Terrestrial Temperate Good Quality	1.000	0.993	0.985	0.975	0.962	0.952	0.954	0.954	0.954	0.955	0.957	0.957	0.956	0.953	0.949	0.946
lower 95% CI	1.000	0.984	0.971	0.957	0.940	0.929	0.930	0.929	0.928	0.929	0.931	0.930	0.929	0.925	0.920	0.917
upper 95% CI	1.000	1.002	0.998	0.991	0.982	0.976	0.977	0.978	0.978	0.980	0.983	0.983	0.982	0.980	0.977	0.974
Terrestrial Temperate Poor Quality	1.000	1.044	1.083	1.124	1.164	1.211	1.258	1.310	1.355	1.396	1.443	1.489	1.537	1.585	1.635	1.685
lower 95% CI	1.000	1.009	1.033	1.061	1.092	1.130	1.169	1.211	1.251	1.286	1.328	1.367	1.409	1.450	1.493	1.537
upper 95% CI	1.000	1.082	1.138	1.191	1.243	1.301	1.357	1.417	1.469	1.519	1.574	1.624	1.680	1.736	1.792	1.851
Terrestrial Tropical Good Quality	1.000	1.018	1.015	1.023	1.025	1.030	1.023	1.012	0.982	0.946	0.915	0.885	0.857	0.828	0.801	0.765
lower 95% CI	1.000	0.982	0.970	0.968	0.962	0.963	0.950	0.933	0.898	0.862	0.829	0.799	0.773	0.744	0.716	0.682
upper 95% CI	1.000	1.057	1.065	1.081	1.091	1.101	1.100	1.097	1.072	1.036	1.006	0.978	0.950	0.921	0.893	0.856
Terrestrial Tropical Poor Quality	1.000	1.017	1.021	1.024	1.035	1.054	1.072	1.077	1.080	1.083	1.086	1.085	1.086	1.102	1.090	1.074
lower 95% CI	1.000	0.990	0.982	0.977	0.982	0.996	1.010	1.011	1.012	1.014	1.014	1.011	1.010	1.023	1.006	0.987
upper 95% CI	1.000	1.045	1.062	1.073	1.092	1.116	1.139	1.148	1.153	1.159	1.165	1.164	1.168	1.190	1.181	1.169
Terrestrial All Good Quality	1.000	0.994	0.986	0.977	0.965	0.957	0.958	0.957	0.952	0.948	0.944	0.940	0.935	0.928	0.920	0.912
lower 95% CI	1.000	0.985	0.973	0.960	0.944	0.934	0.935	0.933	0.927	0.922	0.918	0.913	0.908	0.900	0.893	0.884
upper 95% CI	1.000	1.002	0.998	0.992	0.984	0.978	0.980	0.979	0.976	0.972	0.970	0.965	0.961	0.954	0.948	0.941
Terrestrial All Poor Quality	1.000	1.038	1.067	1.094	1.123	1.159	1.194	1.227	1.254	1.278	1.303	1.324	1.346	1.376	1.386	1.394
lower 95% CI	1.000	1.011	1.028	1.049	1.071	1.102	1.133	1.161	1.184	1.205	1.227	1.245	1.264	1.289	1.296	1.301
upper 95% CI	1.000	1.070	1.109	1.143	1.180	1.222	1.262	1.301	1.331	1.359	1.387	1.411	1.436	1.471	1.484	1.496
Terrestrial All Quality	1.000	0.999	0.994	0.989	0.981	0.978	0.983	0.987	0.987	0.986	0.987	0.986	0.984	0.981	0.976	0.971
lower 95% CI	1.000	0.989	0.982	0.973	0.962	0.957	0.961	0.964	0.963	0.962	0.962	0.960	0.958	0.955	0.949	0.943
upper 95% CI	1.000	1.007	1.006	1.004	1.000	0.999	1.004	1.009	1.010	1.010	1.012	1.011	1.010	1.008	1.003	0.999
HABITATS																
Terrestrial Grassland Birds	1.000	1.000	0.998	0.995	0.993	0.992	0.990	0.988	0.987	0.986	0.986	0.987	0.987	0.988	0.990	0.991
lower 95% CI	1.000	0.989	0.981	0.976	0.971	0.967	0.963	0.959	0.956	0.954	0.952	0.952	0.951	0.952	0.952	0.953
upper 95% CI	1.000	1.013	1.015	1.017	1.018	1.020	1.020	1.020	1.021	1.022	1.023	1.025	1.026	1.028	1.031	1.033
Terrestrial Grassland Mammals	1.000	0.987	0.962	0.924	0.870	0.843	0.870	0.880	0.883	0.896	0.912	0.923	0.931	0.934	0.919	0.899
lower 95% CI	1.000	0.927	0.882	0.828	0.758	0.724	0.748	0.756	0.755	0.765	0.780	0.788	0.793	0.796	0.782	0.762
upper 95% CI	1.000	1.034	1.034	1.016	0.982	0.966	0.996	1.012	1.016	1.033	1.053	1.067	1.078	1.084	1.067	1.044
Terrestrial Temperate Forest	1.000	0.997	0.995	0.995	0.994	0.995	1.000	1.005	1.011	1.018	1.025	1.031	1.036	1.042	1.049	1.051
lower 95% CI	1.000	0.982	0.975	0.970	0.967	0.965	0.969	0.972	0.977	0.983	0.989	0.994	0.998	1.003	1.008	1.009
upper 95% CI	1.000	1.009	1.014	1.017	1.019	1.023	1.030	1.037	1.044	1.053	1.062	1.069	1.075	1.082	1.090	1.093
Terrestrial Tropical Forest	1.000	0.975	0.948	0.924	0.904	0.884	0.857	0.836	0.819	0.791	0.767	0.742	0.722	0.717	0.714	0.706
lower 95% CI	1.000	0.958	0.914	0.873	0.846	0.821	0.790	0.765	0.743	0.717	0.692	0.668	0.647	0.638	0.632	0.621
upper 95% CI	1.000	0.989	0.976	0.969	0.958	0.944	0.924	0.907	0.895	0.867	0.844	0.821	0.801	0.800	0.802	0.795

Table S4: Living Planet Index 19

Systems	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
Living Planet Index	0.943	0.936	0.929	0.920	0.908	0.893	0.877	0.871	0.857	0.840	0.827	0.816	0.811	0.801	0.797	0.801	0.806	0.806
lower 95% CI	0.910	0.903	0.895	0.885	0.872	0.857	0.838	0.824	0.810	0.793	0.779	0.767	0.761	0.749	0.741	0.744	0.747	0.745
upper 95% CI	0.978	0.971	0.965	0.956	0.945	0.931	0.918	0.921	0.909	0.892	0.879	0.870	0.866	0.857	0.852	0.856	0.861	0.862
Freshwater Temperate	1.163	1.164	1.159	1.131	1.096	1.077	1.058	1.036	1.022	1.007	0.981	0.942	0.911	0.843	0.804	0.809	0.815	0.816
lower 95% CI	1.095	1.093	1.085	1.053	1.011	0.992	0.972	0.950	0.934	0.919	0.889	0.850	0.818	0.747	0.710	0.714	0.717	0.716
upper 95% CI	1.236	1.242	1.241	1.218	1.191	1.171	1.152	1.130	1.119	1.105	1.083	1.046	1.016	0.948	0.910	0.917	0.925	0.929
Freshwater Tropical	0.872	0.839	0.803	0.805	0.807	0.781	0.750	0.770	0.751	0.727	0.707	0.684	0.691	0.708	0.687			
lower 95% CI	0.741	0.690	0.648	0.648	0.646	0.623	0.585	0.571	0.555	0.535	0.516	0.496	0.497	0.508	0.482			
upper 95% CI	1.016	1.000	0.974	0.978	0.988	0.961	0.945	1.050	1.027	0.995	0.973	0.943	0.957	0.985	0.974			
Freshwater Index	1.008	0.995	0.967	0.968	0.942	0.916	0.883	0.892	0.871	0.854	0.829	0.799	0.790	0.770	0.744	0.747	0.751	0.754
lower 95% CI	0.926	0.901	0.862	0.861	0.833	0.807	0.770	0.759	0.739	0.721	0.698	0.671	0.660	0.640	0.614	0.614	0.617	0.618
upper 95% CI	1.096	1.095	1.075	1.079	1.059	1.032	1.003	1.056	1.034	1.017	0.992	0.956	0.947	0.931	0.903	0.907	0.915	0.922
Marine Atlantic	1.116	1.128	1.149	1.172	1.195	1.206	1.205	1.200	1.193	1.186	1.189	1.190	1.188	1.186	1.191	1.193	1.192	1.192
lower 95% CI	1.055	1.065	1.084	1.104	1.123	1.130	1.128	1.121	1.113	1.105	1.107	1.106	1.103	1.099	1.101	1.100	1.096	1.093
upper 95% CI	1.183	1.196	1.221	1.247	1.275	1.289	1.290	1.285	1.279	1.273	1.278	1.281	1.280	1.280	1.289	1.296	1.299	1.304
Marine Indian	0.730	0.705	0.679	0.653	0.626	0.584	0.543	0.505	0.465	0.428	0.415	0.415	0.400	0.380	0.408			
lower 95% CI	0.605	0.582	0.557	0.532	0.508	0.471	0.436	0.402	0.367	0.336	0.320	0.301	0.277	-0.845	-1.008			
upper 95% CI	0.872	0.846	0.821	0.795	0.764	0.717	0.671	0.631	0.586	0.543	0.531	0.545	0.531	0.517	0.558			
Marine Pacific	1.049	1.050	1.049	1.052	1.056	1.064	1.070	1.072	1.070	1.067	1.039	1.011	0.981	0.949	0.918	0.925	0.931	0.929
lower 95% CI	0.959	0.956	0.954	0.954	0.956	0.961	0.964	0.964	0.962	0.959	0.932	0.904	0.875	0.845	0.816	0.821	0.824	0.822
upper 95% CI	1.148	1.151	1.153	1.160	1.165	1.176	1.183	1.188	1.187	1.187	1.158	1.129	1.100	1.067	1.034	1.043	1.050	1.048
Marine Southern	0.668	0.659	0.645	0.633	0.622	0.620	0.620	0.622	0.622	0.624	0.657	0.680						
lower 95% CI	0.567	0.556	0.543	0.532	0.520	0.515	0.512	0.510	0.507	0.504	0.528	0.543						
upper 95% CI	0.785	0.778	0.765	0.753	0.742	0.743	0.748	0.755	0.761	0.770	0.813	0.842						
Marine Index	0.869	0.861	0.853	0.845	0.837	0.826	0.812	0.797	0.779	0.763	0.762	0.764	0.757	0.748	0.762	0.772	0.783	0.792
lower 95% CI	0.813	0.804	0.795	0.786	0.777	0.765	0.750	0.734	0.716	0.699	0.696	0.695	0.687	0.673	0.680	0.687	0.694	0.698
upper 95% CI	0.929	0.922	0.915	0.907	0.900	0.890	0.878	0.864	0.847	0.831	0.833	0.838	0.834	0.827	0.839	0.850	0.859	0.866
Terrestrial Index	0.929	0.915	0.904	0.895	0.884	0.874	0.866	0.862	0.857	0.846	0.829	0.825	0.826	0.827	0.831	0.828	0.827	0.815
lower 95% CI	0.889	0.875	0.863	0.852	0.841	0.831	0.821	0.816	0.809	0.797	0.779	0.774	0.773	0.775	0.778	0.773	0.767	0.753
upper 95% CI	0.970	0.957	0.947	0.938	0.928	0.919	0.913	0.911	0.910	0.898	0.883	0.880	0.882	0.884	0.889	0.889	0.891	0.880
Terrestrial Tropical	0.863	0.837	0.812	0.791	0.767	0.749	0.730	0.718	0.709	0.688	0.657	0.645	0.641	0.637	0.637	0.629	0.624	0.604
lower 95% CI	0.794	0.769	0.744	0.723	0.699	0.681	0.662	0.648	0.636	0.614	0.583	0.571	0.565	0.562	0.561	0.549	0.541	0.518
upper 95% CI	0.936	0.911	0.886	0.866	0.842	0.825	0.809	0.797	0.792	0.771	0.739	0.728	0.727	0.724	0.724	0.719	0.720	0.703
Terrestrial Temperate	1.000	1.000	1.006	1.012	1.017	1.021	1.027	1.035	1.036	1.039	1.046	1.055	1.065	1.075	1.085	1.091	1.097	1.100
lower 95% CI	0.969	0.968	0.974	0.979	0.984	0.986	0.992	0.998	0.999	1.002	1.007	1.016	1.024	1.033	1.042	1.047	1.052	1.054
upper 95% CI	1.031	1.031	1.038	1.044	1.051	1.055	1.063	1.071	1.073	1.078	1.084	1.096	1.106	1.118	1.129	1.136	1.142	1.146
WEIGHTED																		
Terrestrial African Mammals Equal wt	0.837	0.814	0.787	0.763	0.726	0.690	0.654	0.619	0.586	0.558	0.539	0.526	0.518	0.513	0.510	0.505	0.503	0.500
lower 95% CI	0.775	0.751	0.724	0.698	0.661	0.626	0.592	0.560	0.528	0.499	0.480	0.465	0.456	0.450	0.447	0.442	0.439	0.435
upper 95% CI	0.907	0.887	0.861	0.839	0.799	0.762	0.724	0.688	0.655	0.626	0.610	0.599	0.594	0.589	0.586	0.582	0.580	0.577
Terrestrial African Mammals weighted	0.826	0.804	0.779	0.756	0.720	0.686	0.652	0.619	0.588	0.562	0.544	0.531	0.524	0.518	0.515	0.512	0.509	0.506
lower 95% CI	0.762	0.738	0.713	0.689	0.655	0.623	0.590	0.558	0.529	0.502	0.483	0.469	0.459	0.453	0.450	0.446	0.443	0.440
upper 95% CI	0.894	0.875	0.850	0.829	0.792	0.756	0.719	0.684	0.653	0.627	0.611	0.602	0.598	0.592	0.589	0.587	0.584	0.582
REALMS																		
Terrestrial Afrotropical	0.874	0.852	0.826	0.803	0.772	0.743	0.713	0.683	0.657	0.633	0.612	0.597	0.585	0.580	0.577	0.572	0.570	0.566
lower 95% CI	0.811	0.787	0.759	0.735	0.705	0.676	0.646	0.618	0.591	0.566	0.545	0.527	0.512	0.506	0.504	0.499	0.496	0.493
upper 95% CI	0.945	0.924	0.900	0.880	0.849	0.819	0.788	0.757	0.731	0.708	0.691	0.678	0.669	0.664	0.662	0.657	0.658	0.654
Terrestrial Nearctic	0.981	0.981	0.989	0.996	1.003	1.011	1.018	1.025	1.026	1.030	1.036	1.043	1.048	1.055	1.061	1.066	1.073	1.082
lower 95% CI	0.947	0.948	0.955	0.961	0.967	0.974	0.980	0.988	0.988	0.990	0.996	1.002	1.006	1.012	1.018	1.021	1.029	1.036
upper 95% CI	1.014	1.015	1.024	1.032	1.040	1.048	1.055	1.064	1.065	1.070	1.078	1.085	1.091	1.098	1.106	1.111	1.120	1.130
QUALITY																		
Terrestrial Temperate Good Quality	0.942	0.940	0.943	0.947	0.953	0.959	0.963	0.968	0.967	0.970	0.974	0.979	0.983	0.989	0.994	0.997	1.000	1.004
lower 95% CI	0.913	0.910	0.912	0.915	0.920	0.926	0.930	0.933	0.932	0.934	0.938	0.942	0.946	0.951	0.955	0.958	0.961	0.964
upper 95% CI	0.971	0.969	0.973	0.978	0.984	0.991	0.996	1.002	1.002	1.005	1.011	1.016	1.021	1.028	1.033	1.037	1.040	1.044
Terrestrial Temperate Poor Quality	1.733	1.775	1.812	1.845	1.853	1.801	1.808	1.819	1.811	1.808	1.819	1.888	1.956	2.035	2.138	2.241	2.325	2.308
lower 95% CI	1.580	1.615	1.650	1.676	1.670	1.596	1.594	1.601	1.587	1.579	1.586	1.627	1.662	1.709	1.788	1.865	1.898	1.848
upper 95% CI	1.906	1.952	1.996	2.033	2.055	2.023	2.039	2.060	2.054	2.057	2.073	2.184	2.291	2.421	2.550	2.705	2.862	2.894
Terrestrial Tropical Good Quality	0.735	0.706	0.677	0.652	0.629	0.611	0.590	0.582	0.582	0.562	0.528	0.505	0.490	0.495	0.499	0.503	0.505	0.505
lower 95% CI	0.652	0.625	0.597	0.573	0.551	0.534	0.512	0.501	0.498	0.479	0.447	0.426	0.413	0.418	0.420	0.423	0.425	0.409
upper 95% CI	0.827	0.797	0.766	0.740	0.716	0.698	0.678	0.674	0.677	0.657	0.621	0.596	0.577	0.584	0.589	0.594	0.597	0.596
Terrestrial Tropical Poor Quality	1.054	1.034	1.014	0.996	0.969	0.941	0.915	0.884	0.851	0.821	0.798	0.807	0.815	0.805	0.802	0.783	0.772	0.734
lower 95% CI	0.964	0.940	0.919	0.899	0.871	0.843	0.814	0.782	0.748	0.717	0.693	0.697	0.698	0.688	0.685	0.660	0.640	0.596
upper 95% CI	1.153	1.136	1.120	1.106	1.082	1.054	1.030	0.999	0.967	0.938	0.921	0.940	0.954	0.943	0.941	0.933	0.932	0.899
Terrestrial All Good Quality	0.905	0.899	0.897	0.896	0.898	0.900	0.900	0.901	0.901	0.899	0.897	0.897	0.898	0.904	0.909	0.912	0.915	0.918
lower 95% CI	0.876	0.870	0.867	0.866	0.867	0.869	0.868	0.868	0.867	0.865	0.862	0.862	0.863	0.868	0.872	0.875	0.878	0.880
upper 95% CI	0.934	0.928	0.927	0.927	0.929	0.931	0.932	0.935	0.935	0.935	0.933	0.933	0.935	0.941	0.946	0.950	0.953	0.957
Terrestrial All Poor Quality	1.401	1.405	1.407	1.409	1.395	1.355	1.342	1.329	1.305	1.285	1.275	1.310	1.342	1.367	1.405	1.427	1.449	1.413
lower 95% CI	1.304	1.306	1.306	1.306	1.288	1.241	1.225	1.208	1.184	1.163	1.150	1.174	1.196	1.209	1.239	1.249	1.255	1.212
upper 95% CI	1.507	1.512	1.519	1.524	1.514	1.483	1.472	1.462	1.440	1.422	1.415	1.463	1.512	1.545	1.594	1.632	1.674	1.655
Terrestrial All Quality	0.966	0.961	0.961	0.961	0.961	0.959	0.960	0.963	0.962	0.960	0.958	0.963	0.970	0.978	0.985	0.990	0.993	0.994
lower 95% CI	0.937	0.932	0.931	0.931	0.930	0.927	0.928	0.930	0.928	0.926	0.923	0.927	0.933	0.940	0.946	0.949	0.952	0.952
upper 95% CI	0.994	0.990	0.990	0.991	0.991	0.991	0.993	0.997	0.997	0.996	0.994	1.000	1.008	1.018	1.026	1.030	1.036	1.036
HABITATS																		
Terrestrial Grassland Birds	0.993	0.994	0.995	0.996	1.001	1.005	1.006	1.005	1.001	0.997	0.994	1.005	1.012	1.014	1.016	1.009	1.001	0.992
lower 95% CI	0.953	0.954	0.954	0.955	0.959	0.961	0.960	0.958	0.953	0.946	0.941	0.945	0.946	0.945	0.946	0.937	0.928	0.917
upper 95% CI	1.035	1.037	1.039	1.041	1.047	1.053	1.055	1.055	1.053	1.050	1.049	1.071	1.087	1.092	1.095	1.088	1.082	1.076
Terrestrial Grassland Mammals	0.873	0.852	0.835	0.817	0.786	0.746	0.714	0.685	0.651	0.617	0.594	0.573	0.555	0.560	0.569	0.573	0.564	0.547
lower 95% CI	0.738	0.719	0.704	0.689	0.659	0.621	0.593	0.569	0.540	0.510	0.490	0.472	0.457	0.460	0.465	0.467	0.458	0.442
upper 95% CI	1.019	0.994	0.979	0.961	0.925	0.884	0.849	0.814	0.776	0.738	0.714	0.690	0.669	0.677	0.691	0.700	0.689	0.672
Terrestrial Temperate Forest	1.058	1.065	1.079	1.090	1.100	1.105	1.109	1.116	1.122	1.132	1.141	1.151	1.163	1.175	1.187	1.199	1.215	1.230
lower 95% CI	1.016	1.022	1.033	1.042	1.051	1.055	1.058	1.064	1.068	1.076	1.085	1.093	1.103	1.115	1.126	1.136	1.149	1.162
upper 95% CI	1.101	1.109	1.126	1.138	1.149	1.155	1.161	1.170	1.177	1.189	1.200	1.213	1.226	1.240	1.254	1.268	1.287	1.303
Terrestrial Tropical Forest	0.706	0.698	0.687	0.676	0.660	0.648	0.637	0.635	0.632	0.617	0.600	0.585	0.582	0.582	0.590	0.606	0.636	0.647
lower 95% CI	0.617	0.606	0.594	0.583	0.568	0.554	0.543	0.537	0.532	0.517	0.501	0.487	0.484	0.483	0.485	0.496	0.512	0.521
upper 95% CI	0.802	0.799	0.788	0.781	0.766	0.754	0.745	0.746	0.747	0.731	0.714	0.700	0.696	0.698	0.712	0.736	0.790	0.806

The Index of Linguistic Diversity: A New Quantitative Measure of Trends in the Status of the World's Languages

David Harmon and Jonathan Loh

Language Documentation and Conservation
(2010) 4, 97-151

LANGUAGE DOCUMENTATION & CONSERVATION

Vol. 4 (2010), pp. 97-151
http://nflrc.hawaii.edu/ldc/
http://hdl.handle.net/10125/4474

The Index of Linguistic Diversity: A New Quantitative Measure of Trends in the Status of the World's Languages

David Harmon
George Wright Society/Terralingua

Jonathan Loh
Zoological Society of London/Terralingua

The Index of Linguistic Diversity (ILD) is a new quantitative measure of trends in linguistic diversity. To derive the ILD we created a database of time-series data on language demographics, which we believe to be the world's largest. So far, the database contains information from nine editions of *Ethnologue* and five other compendia of speaker numbers. The initial version of the ILD, which draws solely on the *Ethnologue* subset of these data, is based on a representative random sample of 1,500 of the world's 7,299 languages (as listed in the 2005 edition). At the global level, the ILD measures how far, on average, the world's languages deviate from a hypothetical situation of stability in which each language is neither increasing nor decreasing its share of the total population of the grouping. The ILD can also be used to assess trends at various subglobal groupings. Key findings:

- Globally, linguistic diversity declined 20% over the period 1970–2005.
- The diversity of the world's indigenous languages declined 21%.
- Regionally, indigenous linguistic diversity declined over 60% in the Americas, 30% in the Pacific (including Australia), and almost 20% in Africa.

1. INTRODUCTION.[1] Concern about the future of the world's languages has been building for the better part of two decades. A large amount of qualitative evidence points to an impending mass extinction[2] of languages. The quality of this evidence ranges from merely

---

[1] We are grateful to The Christensen Fund for underwriting this work as part of a larger project on Global Indicators of the Status and Trends of Linguistic Diversity and Traditional Knowledge, which is being carried out by the NGO Terralingua. Luisa Maffi of Terralingua provided valuable comments throughout the project. We owe a large debt of thanks to M. Paul Lewis, editor of *Ethnologue*, for providing copies of the earliest editions; we also thank him for answering questions about the book's publishing history and reviewing the technical report upon which this paper was based. We are indebted to Margaret Florey for her comments on an earlier draft of the manuscript, to Ashbindu Singh for reviewing the technical report, and to Kenneth L. Rehg and two anonymous referees for their comments on the final draft.

[2] Outside of specialist discussions, the issue of language endangerment is almost always couched in terms of "extinction." Applying the extinction concept to language is fraught with theoretical difficulties—and, even more troublingly, can be used by unsympathetic authorities to thwart the interests of language communities. Still, the metaphor is firmly ensconced in the both the popular and professional literature, and the alternatives (such as "sleeping" or "silent" languages) also have problems. For a good discussion of the difficulties in determining the precise moment when a language goes extinct, see Evans 2001.

E-ISSN 1934-5275

anecdotal to very accurate narrative accounts based on firsthand knowledge of the language demographics of individual speech communities. It is a highly valuable body of evidence, leaving no room to doubt that the entirety of the world's languages—not just their number, but also the linguistic and cultural diversity they represent—is being severely diminished. For a host of complex reasons, people are abandoning their mother tongues and switching to other languages, almost always ones with larger numbers of speakers; thereby, more and more people are being concentrated into fewer and fewer languages.

However, there is much less quantitative evidence of a global linguistic diversity crisis. To help fill this gap we have created the Index of Linguistic Diversity (ILD), which we believe to be the first-ever quantitative index of trends in linguistic diversity based on time-series data on numbers of mother-tongue speakers. The ILD assesses trends in linguistic diversity by comparing changes in the relative distribution of mother-tongue speakers against a benchmark of the situation prevailing in 1970, the earliest year we could set the index based on the data available. The ILD measures how far, on average, the languages in a given geographical grouping deviate from a hypothetical situation of stability in which each language is neither increasing nor decreasing its share of the total population of the grouping. For example, ILD Global, an index of the world's overall linguistic diversity, measures the average deviation of the world's languages from a hypothetical situation in which each language is neither increasing nor decreasing its share of the global population. The index does this by measuring changes in the number of mother-tongue speakers from a globally representative sample of 1,500 languages over the period 1970–2005. (See Appendix A for a discussion of the ILD database.) The ILD can be calculated at different geographic scales and for different groupings of languages; each of these versions of the index uses the same methods.

The main finding of this research is that linguistic diversity has seriously declined since 1970. The overall linguistic diversity of the world, as measured by ILD Global, declined by 20% over the 35–year period (Figure 1). We also assessed the diversity of the world's indigenous languages—which make up 80–85% of the total number—on both global and regional levels. We did this because the status of the world's indigenous languages is important to global initiatives such as the Convention on Biological Diversity, as well as to indigenous communities themselves. ILD Global Indigenous, which measures the diversity of the world's indigenous languages, declined by 21% (Figure 2). The diversity of indigenous languages declined in all regions as well.

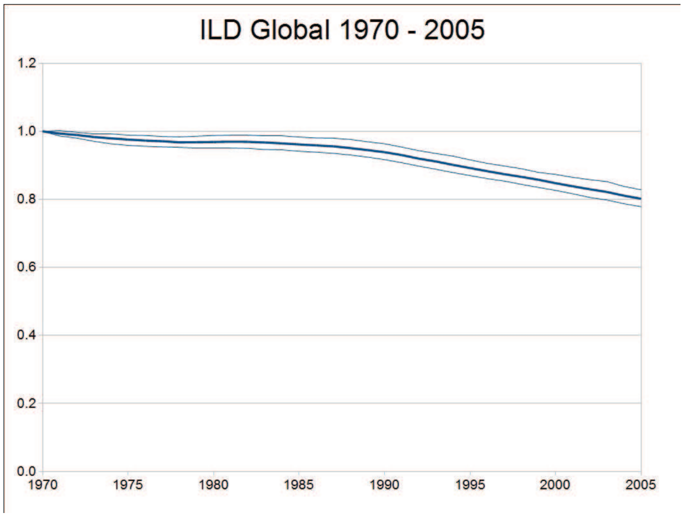


FIGURE 1: ILD Global, 1970–2005.
In Figures 1–7, The upper and lower confidence limits (CLs), showing the boundaries of the 95% confidence interval, are depicted as small lines above and below the main trendline.

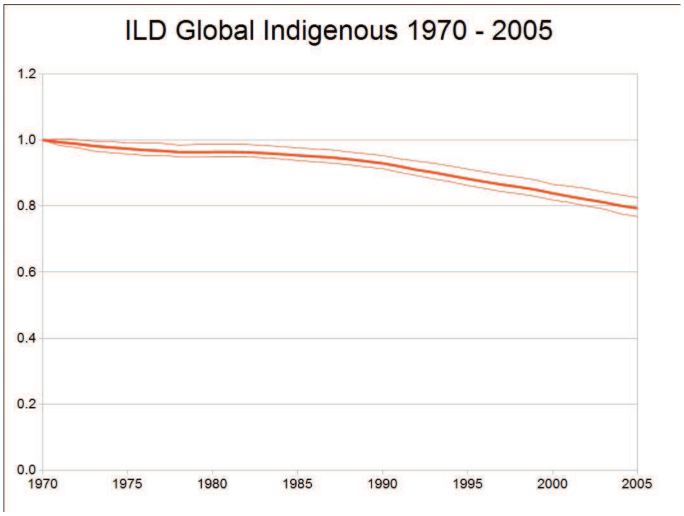


FIGURE 2: ILD Global Indigenous, 1970–2005.

2. WHAT IS LINGUISTIC DIVERSITY? Linguistic diversity is often viewed from three related (but not necessarily correlated) perspectives; this is the approach taken, for instance, by Daniel Nettle (1999). The first is what he calls *language diversity,* and we will call *language richness,* "the number of different languages in a given geographical area" (Nettle 1999:10). The term "language richness" encapsulates two points: first, that speech forms can be and routinely are classified as discrete languages, despite the well-known difficulties of distinguishing languages from dialects; and second, that these discrete languages are countable.

Another perspective on linguistic diversity is that of *phylogenetic diversity,* or variation above the level of languages, such as "the number of different lineages of languages found in an area." Nettle notes that phylogenetic groupings can be identified on many levels—language families, for example (Nettle 1999:10, 115). An area where many closely related languages are spoken therefore has greater language richness but less phylogenetic diversity than one with fewer languages belonging to several different families. The third perspective often used is *structural diversity,* which is the variation found among structures within languages, such as morphology, word order, phonology, and so on (Nettle 1999:130–148).

For the purposes of developing a quantitative measure such as the ILD, we depart slightly from the definitions of linguistic diversity outlined above, and borrow some related concepts from the field of ecology. Language richness can be thought of as being analogous to species richness, the number of species found in a given area. In addition to richness, a second component in species diversity is evenness, or the distribution of individual organisms among species. In the case of linguistic diversity, evenness is the distribution of individual speakers among languages. For example, two regions in both of which ten languages are spoken each have the same richness, but the region in which each language is spoken by 10% of the population has greater evenness, and therefore higher linguistic diversity, than one in which 91% of the population speaks one language and only 1% of the population speaks each of the other nine. We think that this concept is critical in measuring changes in linguistic diversity over comparatively short time scales. Relatively few of the world's languages have become extinct as mother tongues in the last few decades, so language richness in most areas of the world has declined only slightly. And yet, we would argue, diversity has declined much more than this because the distribution of mother-tongue speakers among extant languages has become more uneven: more speakers are becoming concentrated in fewer languages. While phylogenetic and structural diversity are important, these concepts are not currently incorporated into the index. In summary, for the purposes of the ILD, we define linguistic diversity as the number of languages and the evenness of distribution of mother-tongue speakers among languages in a given area.

3. THE NEED FOR A LINGUISTIC DIVERSITY INDEX. If there are already projections of the future magnitude of language extinctions, why is there a need for an index like the ILD? First, published estimates of the percentage of languages likely to die out during this century are, to date, little more than informed conjecture. Categorical statements of the rate of extinction—"X number of languages are dying every year"—are widely quoted but almost never referenced to a rigorous estimate.[3]

Second, even if better estimates were available, merely tracking when particular languages go extinct does not account for the loss of linguistic diversity occurring during the course of pre-extinction language shift. A great deal of linguistic diversity is lost well before a declining language finally goes extinct, as speakers shift to other (usually larger) languages, intergenerational transmission declines, and usage becomes restricted to fewer speakers, domains, and functions. Quantifying changing distributions of mother-tongue speakers prior to extinction is therefore important.

Moreover, focusing on language extinction rates places undue emphasis on what is perceived to be the terminal state of linguistic diversity decline. If "language extinction" is to have any useful meaning, it must be specified that the term actually refers to the condition of a language no longer being spoken as a mother tongue. While there are several possible definitions of "mother tongue,"[4] what we mean by the term is that language which an individual would speak first (though not necessarily exclusively) if given free rein to choose. The term "first language" as used in *Ethnologue* (see Lewis 2009:13) captures the essence of what we mean. For the purposes of constructing the ILD, we assume that even multilingual people can have only one mother tongue.

Moreover, many languages, extinct as mother tongues in the sense just defined, continue to be spoken in everyday use as second languages or in one or more select domains (e.g., at home, as part of ceremonies, etc.). A language that is extinct as a mother tongue may live on as a *language of heritage* and, in some cases, might one day be revived as a mother tongue. Popular accounts usually gloss over or omit the fact that, with reference to language, the extinction metaphor does not necessarily imply absolute irreversibility.[5]

---

[3] It appears that many estimates originate in speculations made by Michael Krauss in his seminal 1992 *Language* paper, in which he said it is conceivable that as many as 90% of the world's languages could become extinct or irreversibly moribund by the end of the 21st century, and speculated that 50% of the world's 6,000 languages (his consensus figure of that time) were already moribund (Krauss 1992:6–7). Crystal (2000:19) notes this, and then proceeds to work through the math, deriving an estimate of 26 extinctions per year by extrapolating Krauss' estimate (6,000 languages, 50% loss over the next 100 years). This or a similar calculation appears to be the basis for statements such as that of the Living Tongues Institute for Endangered Languages, which says on its web site: "Every two weeks the last fluent speaker of a language passes on and with him/her goes literally hundreds of generations of traditional knowledge encoded in these ancestral tongues" (http://www.livingtongues.org/index.html; accessed June 2009). Similar statements can be found in journalistic accounts. Summarizing languages according to an endangerment typology (e.g., Krauss 2006; UNESCO 2009a) holds promise for a more accurate projection of likely extinctions; for an example see Table 10.2 in Evans 2010, and the accompanying discussion (pp. 211–216).

[4] See the discussion in Skutnabb-Kangas 2000:105–115 and the commentary thereon in Harmon 2002:56–58; see also Gunnemark and Kenrick 1985:242.

[5] For a somewhat different definition of "heritage language," see Golla 2007:8–9. In addition, a small number of languages, which we refer to as "auxiliary languages," were never spoken as mother tongues but instead always restricted to a particular domain. For these auxiliary languages, the term "extinction" simply refers to their no longer being spoken at all.

So, while obtaining accurate projections of mother-tongue language extinctions is important, they need to be augmented by a quantitative measure of current global trends in linguistic diversity. Clearly, the claims of those who tout the loss of linguistic diversity as a major problem for the world would be strengthened if there were quantitative evidence to support their arguments. Government officials, other decision-makers, and the general public will likely take the decline of linguistic diversity more seriously if there is a readily understandable global metric that captures the current magnitude of the problem. That is what the ILD is designed to provide.

4. WHAT THE ILD MEASURES. As stated earlier, the ILD uses language evenness in conjunction with language richness as a proxy for linguistic diversity. Because the goal of the index is to measure trends in linguistic diversity, it must account for changes in evenness and richness: that is, changes in the relative distribution of mother-tongue speakers among discrete languages within the total population, as measured from the starting point of the index to its ending point. The ILD indicates the rate of change in linguistic diversity by measuring how far, on average, the languages in a given grouping deviate from a hypothetical situation in which each language is neither increasing nor decreasing its share of the total population of that grouping.

To illustrate this, let us look at ILD Global, with measures the world's overall linguistic diversity. ILD Global tracks the trend in the world's linguistic diversity since 1970, the earliest year for which sufficient data are available to calculate the index. The index value is set equal to 1 in the baseline year, and in each subsequent year shows the trend in the share of the world population represented by the average[6] of all the languages in the sample relative to the baseline year. If the average is declining, it means that the distribution is becoming less even (i.e., more skewed), with a few large languages increasing their global share at the expense of many smaller languages. If the average is increasing, it means that the distribution is becoming more even, with many languages increasing their share at the expense of a few large languages. If somehow each language could maintain its initial proportion, the ILD Global trendline would be flat. Any increases and decreases in the index can also be thought of as changes in the relative abundance of the world's languages: a rising trendline means more people are shifting away from dominant languages to minority languages, while a falling trendline means more people are shifting to majority languages and away from minority ones.

To calculate an ILD for languages spoken in a given population, we track the proportion of the total population speaking each language in each year, and then take the average. The index measures how that average changes over time. Thus the ILD can be said to measure the concentration or distribution of mother-tongue speakers among the world's languages. What does it mean to say that ILD Global declined 20% over the period 1970–2005? It means that, for all languages spoken worldwide in 1970, their average share of the world's population declined by 20% over 35 years. (Appendix B contains a technical discussion of how the ILD tracks changes in the average share, and also provides some

---

[6] The average is calculated using the geometric mean rather than the arithmetic mean. See Appendix B for a more detailed explanation.

simple comparative scenarios with accompanying graphics that make clear how the ILD changes under different conditions of language shift.)

Although we have used ILD Global as an example, the same methods and reasoning apply to subglobal ILDs. For instance, ILD Americas measures the trend in the share of the population of the Americas represented by the average of all the region's languages in the sample relative to the baseline year.

The ILD is entirely retrospective, indicating changes that have taken place in the past, and is not designed to predict future changes. It is not a measure of the future viability of any one language or group of languages. Rather, it provides a snapshot of the trends in the distribution of speakers among the world's languages between the starting year of the index (1970) and the final year (2005 in the current version).

In the ILD each language carries equal weight, regardless of its relative size. While it is possible to produce a weighted index that would impart more importance to phylogenetic diversity—say, by giving extra weight to isolates—such weightings are always more or less arbitrary. Making the ILD be unweighted means that the phylogenetic uniqueness of any particular language does not differentially affect the calculation of the index. Neither does the mode of shift that any particular language may be undergoing, so that, for example, language attrition caused by rural-to-urban migration is, in terms of its effect on the ILD, no different than attrition caused by intergenerational transmission failure within a geographically homogeneous speech community, or language loss caused by a catastrophic decline of a community of speakers. This means the ILD is not useful for illuminating the sociolinguistic bases of language shift.

Finally, it is worth noting again that the ILD is not a measure of language extinction: a 10% decline in the index does not mean that 10% of languages went extinct over the period being measured. For example, it is possible that most of the world's languages could decline until only a few speakers of each are left, while a few languages become dominant with many millions of speakers: the ILD would show a marked decline and yet the total number of extant languages would remain constant. In that case the number of extinctions would remain zero, yet the ILD would indicate that almost all linguistic diversity had been lost.

5. THE ILD DATABASE. The ILD is based on a sample of 1,500 languages selected at random from the 7,299 languages listed in the 15th edition of *Ethnologue* (Gordon 2005). (The 16th edition, Lewis 2009, appeared too late for us to include in this study.) This sample size—representing just over 20% of the world's languages—is higher than is needed to constitute a statistically representative global sample. Having a sample size much larger than required for global analysis allows statistically valid analysis of subglobal samples. A larger-than-needed sample size also provides a cushion against sample attrition.

Our long-term aim is to base the ILD on a variety of data sources, not just *Ethnologue*. However, we decided to restrict the first version of the ILD to *Ethnologue* data to minimize potential inconsistencies in language-status assessment that could come from incorporating multiple sources of data into a single time series. Thanks to the assistance of M. Paul Lewis of SIL International, we were able to obtain copies of some extremely rare early editions, which allowed us to complete a collection of all 15 editions available at the time of analysis. This enabled us to move the ILD's starting date further back than we initially anticipated. After reviewing all the editions, we selected the following nine on which to base the initial version of the ILD: 1st (WBT 1951), 5th (Canonge and Pittman n.d. [ca. 1958]), 9th (B.F. Grimes 1978), 10th (B.F. Grimes 1984), 11th (B.F. Grimes 1988); 12th (B.F. Grimes 1992a), 13th (B.F. Grimes 1996a), 14th (B.F. Grimes 2000a), and 15th (Gordon 2005). We chose these editions because (at the time) they spanned the entire history of *Ethnologue* while giving priority to later editions whose contents are much more comprehensive.

Our first step was to enter benchmark demographic information from the 15th edition of *Ethnologue* (Gordon 2005) into a Base Data Entry Form. Next, we reviewed the nine editions listed above looking for data on the number of mother-tongue speakers for our sample languages.[7] Within the time available to us we were able to examine six of the editions (1st, 5th, 9th, 12th, 14th, and 15th) for data on the full sample of 1,500 languages. For the remaining three editions (10th, 11th, and 13th), we were able to search for the 751 languages in our sample from Africa and the Americas only. Thus, we performed a total of 11,253 data searches. After eliminating duplicates, we were left with 2,703 unique datapoints; these form the basis for the first iteration of the ILD.

Our protocol was to enter results for each of these data searches into the database using a Mother-Tongue Speaker Trend Data Form, even though the vast majority of them did not produce unique datapoints. Doing this ensures that there is no ambiguity about whether a particular data source has been consulted with regard to any given language.

Before creating the ILD we analyzed the data for representativeness, eliminated duplicate datapoints and entries having no data, assessed and adjusted for data trend anomalies, removed discrepant datapoints, and dealt with apparent extinctions within time series. Technical detail about these steps, along with samples of the Base Data Entry Form and the Mother-Tongue Speaker Trend Data Form (and explanations of the fields in these forms) and a discussion of other points related to the creation of the database and data analysis, are in Appendix A. Readers who wish to review the data can do so at http://www.terralingua.org/projects/ild/ild.htm.

6. CALCULATING THE ILD. The following account describes a simplified method for calculating the ILD that requires data on each language for each year included in the index.

---

[7] As noted above, there are a small number of auxiliary languages that have never had any mother-tongue speakers. Five of these happened to fall within our sample (The 3-letter codes in square brackets are the languages' ISO 639-3 codes; for more, see Appendix A.): Amerax [aex], reputed to be spoken only as a second language by neo-Muslims in American prisons (and which, incidentally, is no longer listed in the 16th edition of *Ethnologue*); To [toz], an ancient secret male initiation language of the Gbaya people of Cameroon; Lucumi [luq], a secret language used for ritual by the Santeria religion; Yinglish [yib], a blend of Yiddish [yid] and English [eng] that is used as a second language only; and Europanto [eur], an artificial language mixing elements from major European languages, which is spoken in the European Union buildings in Belgium. Because these auxiliary languages do not currently have any mother-tongue speakers we excluded them from the calculation of the ILD. They could, conceivably, gain mother-tongue speakers in the future, as has the best-known intentionally constructed auxiliary language, Esperanto [esp], which now has 200–2,000 mother-tongue speakers (Gordon 2005).

For a detailed explanation of the method needed when there are gaps in the data, as with numbers of speakers, refer to Appendix B. The method has three steps that remain the same whether the global level or a regional grouping is being analyzed:

The fraction F of the total population (global or regional) represented by each datapoint (N speakers of language l in year y) was calculated.

$$F_{ly} = N_{ly}/P_y$$

where

N_{ly} is the number of speakers of language l in year y, and
P_y is the total population in year y.

The total populations from 1950 to 2005 of the world and five regions—Africa, Asia, Pacific, Europe, and the Americas—were taken from UN Population Division (2006 revision), downloaded from http://esa.un.org/unpp/index.asp.

The geometric mean of the F values in each year was calculated:

$$M = (F_1.F_2.F_3...F_n)^{1/n}$$

where

n = total number of languages.

Finally, the geometric means in each year were chained together to form an index, such that:

$$I_y = I_{y-1}(M_y/M_{y-1})$$

where

I_y = the Index of Linguistic Diversity in year y
M_y = the geometric mean F value in year y, and
M_{y-1} = the geometric mean F value the previous year

and the index value in 1970 was set to unity

$$I_{1970} = 1.0$$

In this way, the ILD shows the trend in the fraction of the total population that speaks a language that is average or typical of all languages in the sample.

7. RESULTS.

Global Linguistic Diversity. ILD Global (Figure 1), which covers all the languages in the sample, both indigenous and non-indigenous, shows a slow decline from 1.0 to 0.95 between 1970 and 1988, but a steeper decline from 0.95 to 0.80 between 1988 and 2005. The upper and lower confidence limits (CLs) show the boundaries of the 95% confidence interval, and are depicted in this and the other graphs as small lines above and below the main trendline.[8]

Global Indigenous Linguistic Diversity. ILD Global Indigenous (Figure 2), which covers only the indigenous languages in the sample, declined from 1.0 to 0.94 between 1970 and 1988, and from 0.94 to 0.79 between 1988 and 2005. It shows a marginally greater decline than the global ILD, but the two trends are largely similar as most of the languages in the global dataset are indigenous languages (see Appendix A for discussion).

Regional Indigenous Linguistic Diversity. Changes in indigenous linguistic diversity differ among regions. ILD Africa Indigenous increased from 1.00 to 1.07 between 1970 and 1985, and then declined rapidly from 1.07 to 0.83 in 2005 (Figure 3). The increase in the 1970s and early 1980s suggests that African indigenous languages were becoming more equally distributed in terms of speaker numbers during that period, but from the mid-1980s on the distribution became increasingly skewed, with many languages' share of the total African population declining.

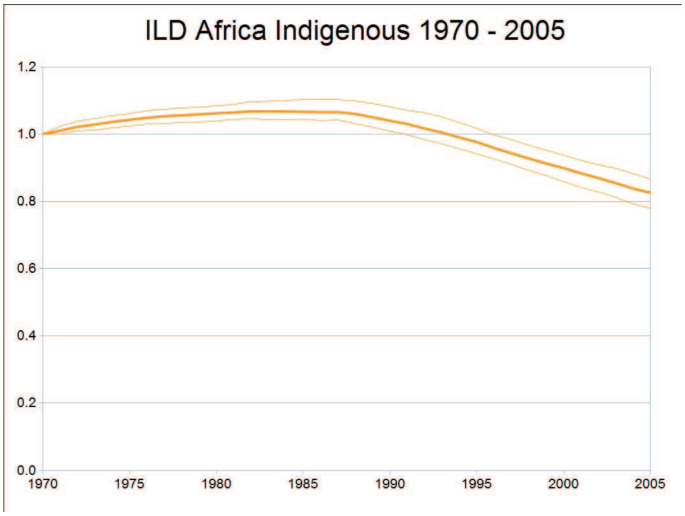


FIGURE 3: ILD Africa Indigenous, 1970–2005.

---

[8] Confidence limits were calculated by bootstrapping with 1,000 bootstraps.

ILD Americas Indigenous shows the steepest decline of any region, falling from 1.00 to 0.71 between 1970 and 1980, and from 0.71 to 0.36 between 1980 and 2005 (Figure 4).

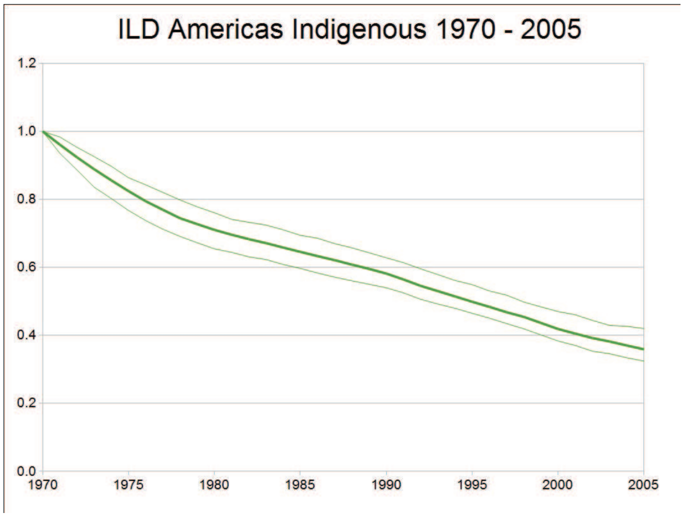


Figure 4: ILD Americas Indigenous, 1970–2005.

ILD Eurasia Indigenous, like its African counterpart, showed an initial increase from 1.00 to 1.10 between 1970 and 1981, suggesting that there was a slight gain in the proportion of the total population speaking an indigenous language. It flattened out for about a decade between 1981 and 1991, and then declined very slightly to 1.07 in 2005 (Figure 5). Overall the index shows little change in linguistic diversity in Eurasia.

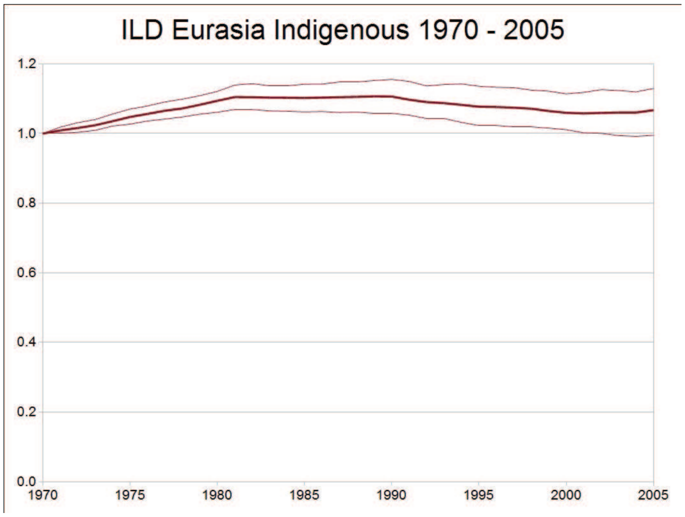


Figure 5: ILD Eurasia Indigenous, 1970–2005.

ILD Pacific Indigenous (which includes Australia) shows the second steepest decline after the Americas. The index fell steadily from 1.0 to 0.82 in 1999, then dropped steeply from 0.82 to 0.70 between 1999 and 2005 (Figure 6). The widening confidence intervals in the last few years of the index suggest a higher degree of uncertainty in the trend after 1999, which would be reduced with additional data.

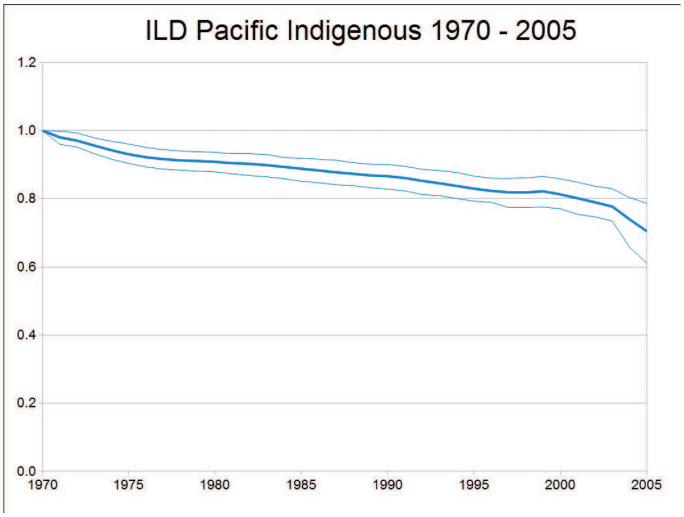


Figure 6: ILD Pacific Indigenous, 1970–2005.

Because the linguistic situation in Australia is distinctive within in the Pacific region, ILD Australia Indigenous (Figure 7) shows a national ILD for Australian Aboriginal/Torres Strait Islander languages alone. The graph includes a second, non-*Ethnologue* data source for comparison: the Australian Bureau of Statistics (ABS). Data from the two sources are shown both separately and combined. ABS data from on numbers of speakers of aboriginal languages from 1996 to 2006 were used to compare trends derived from these data (blue line) with those derived from *Ethnologue* data (red line). The *Ethnologue* data show a decline from 1.0 to 0.7 between 1970 and 1991, then a faster decline from 0.70 to 0.38 between 1991 and 2005. The ABS data show a decline from 1.0 in 1996 to 0.87 in 2006, which is similar to the rate of the *Ethnologue*-based ILD from 1970 to 1991. Combining data from both sources results in an index that declines from 1.0 to 0.53 between 1970 and 2006. Whichever data source[9] is used, Australia shows a more rapid loss of linguistic diversity than the rest of the Pacific region. The rate of loss is comparable to that of the Americas.

---

[9] As acknowledged by the compilers of *Ethnologue*, the data for Australia in editions immediately preceding the current one (i.e., the 16th, published in 2009) were quite out-of-date (M. Paul Lewis, pers. comm., 25 May 2009), and some linguists have raised concerns about the accuracy of ABS data. Nonetheless, the overall point—that Aboriginal languages are in sharp decline—does not seem to be dispute.

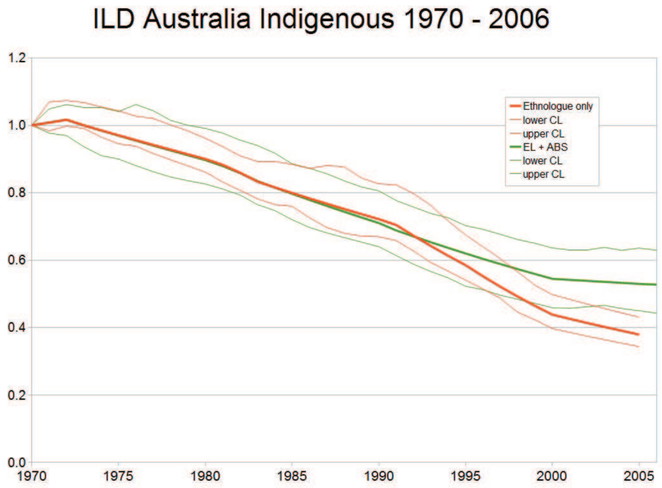


FIGURE 7: ILD Australia Indigenous, 1970–2005, and ABS Data, 1996–2006.

Figure 8 shows the four regional ILDs in one chart for ease of comparison.

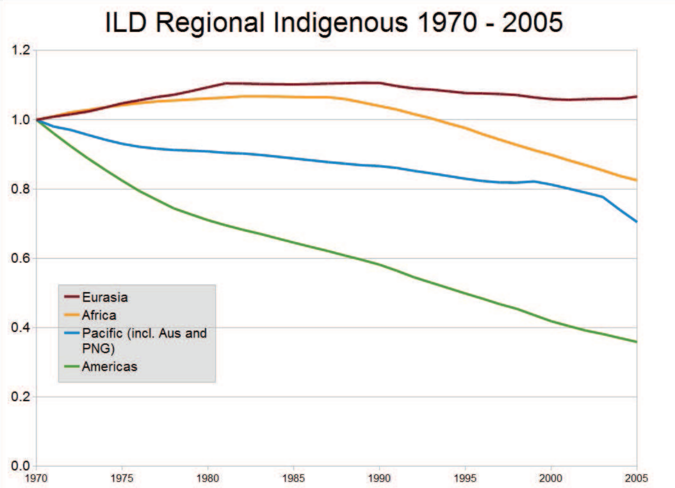


FIGURE 8: Regional Indigenous ILDs, 1970–2005.

8. DISCUSSION.

Decline in Global Linguistic Diversity. Figure 1 shows the global trendline for the ILD. ILD Global shows a slow decline from 1.0 to 0.95 between 1970 and 1988, but a steeper decline from 0.95 to 0.80[10] between 1988 and 2005. The overall decline of 20% in the space of 35 years shows that linguistic diversity is being lost at a significant rate, but even more importantly, the rate of loss has increased from about – 0.3% per year in the 1970s and 1980s to more than –1.0% per year in the 1990s and 2000s. This is a stark indication of the scale of the recent loss of global linguistic diversity. The rapid disappearance of one-fifth of the linguistic diversity that existed in the world in 1970 is a quantitative depiction of the continuing widespread shift from smaller languages to larger languages. The more the ILD Global declines, the more the world's mother-tongue speakers are concentrated into fewer languages.

Decline in Global Indigineous Linguistic Diversity. Figure 2 shows that the decline in the diversity of the world's indigenous languages has been similar, which is unsurprising in that most of the languages in the world (by our estimate, 80–85%) are indigenous languages. (See Appendix A for discussion.) ILD Global Indigenous declined from 1.0 to 0.79 between 1970 and 2005—a 21% decrease. The average annual rate of decline in indigenous linguistic diversity was slightly faster than the global average in the 1970s and 1980s, but only by a fraction of a percent per year.

Making judgments about whether particular languages are to be considered "indigenous" can be difficult, and this problem is discussed further in Appendix A. Suffice it to say here that it is indeed important to make such judgments. Indigenous communities themselves certainly want to know the status of indigenous languages; see, for example, the documents associated with the International Expert Group Meeting on Indigenous Languages (UNPFII 2008).

Moreover, the Convention on Biological Diversity has identified stemming the rate of loss of linguistic diversity and in the number of speakers of indigenous languages as one its indicators for assessing progress toward meeting its 2010 Biodiversity Target. The acceleration in the loss of linguistic diversity indicated by the ILD Global Indigenous implies that this particular CBD target will not be met.

Declines in Regional Indigenous Linguistic Diversity. A comparison of the various regional indigenous ILDs (Figure 8) shows some interesting results. Some regions are declining more rapidly than others, particularly the Americas, which declined by 64% over the period (Figure 4). The fact that the Americas showed the greatest overall decline should not necessarily be interpreted as meaning that linguistic diversity is, consequently, lower there than in other regions. It simply means that the Americas underwent the most rapid decline of all four regions between 1970 and 2005. It may well have been the case that the

---

[10] It has been suggested to us that this pattern may indicate, at a "macro" level, the phenomenon of what Nancy Dorian has called "abrupt transmission failure" or "tip," in contrast to "gradual shift." Dorian wrote: "In terms of possible routes toward language death, it would seem that a language which has been demographically highly stable for several centuries may experience a sudden 'tip,' after which the demographic tide flows strongly in favor of some other language" (1981:51).

Americas were much more linguistically diverse in 1970 compared with other regions, such as Europe for example, in which the majority of linguistic diversity was lost prior to 1970.

The Pacific region (Figure 6) shows the second greatest rate of decline, 30% over 35 years, while ILD Africa Indigenous (Figure 3) declined by nearly 20%. This suggests that indigenous languages are in very rapid decline in comparison to total population growth in the region as a whole in the Americas, and in rapid decline in Africa and the Pacific.

Eurasia was the only region to show an increase in its indigenous ILD (Figure 5). There, indigenous languages are growing at the same rate as the overall population.[11]

In addition to the regional analyses, we calculated a national ILD for indigenous Australian languages (Figure 7). We did this in two ways: first based on the data for the 20 Australian languages in the ILD database, and then using additional data from the Australian Bureau of Statistics. The ABS data are based on censuses conducted in 1996 and 2006, and show an average decline of 13% over ten years across 45 languages for which there were two datapoints. This gives a yardstick of trends in a relatively well-monitored group of indigenous languages with which to compare the trends reported in *Ethnologue* (over the last decade of the index at least). ILD Australia Indigenous, based on *Ethnologue* data only, showed a decline of over 60%, but with the addition of the ABS data from 1996 to 2006, this decline was reduced to less than 50% (Figure 7). Nevertheless, this reflects a severe and rapid loss of linguistic diversity in Australia since the 1970s.

There are aspects of these results which may change with further analysis. Africa, Asia, and Europe show increases in diversity in the 1970s and 1980s. These increases are possibly an artifact of some *Ethnologue* data which do not reflect genuine changes in diversity. Some of these data anomalies may be discovered with additional scrutiny of the dataset.

Starting Point of the Index. Another consideration of an index based on numbers of speakers is when to fix the initial starting point. A flat trendline describes a state in which richness is being maintained (i.e., not being lost) in relative chronological terms; that is, relative to the starting year of the analysis. It is important to understand that the initial starting point (in this case 1970) does not describe a maximal state of linguistic diversity in absolute terms. For any set of languages spoken in a given region, maximum diversity is reached when each language has an equal number of speakers. The starting point of the ILD for any given region is highly unlikely to be maximal. Qualitative estimates point toward global linguistic richness having reached its peak thousands of years ago, long before there were any quantitative data by which to measure it.

Ideally, the ILD's starting point should be as early as possible. *Ethnologue* has sufficient quantitative data to set the starting point at 1970, but prior to 1970 there are not enough datapoints from which to derive global numerical trends. Yet as just noted the

---

[11] We combined Europe and Asia into a single Eurasian region because the sample size in Europe is so small. In Europe, indigenous languages in our sample are in moderate decline in comparison to Europe's total population, but the sample size is not large enough to enable us to draw significant conclusions. Therefore, the ILD Eurasian Indigenous trendline can mostly be attributed to indigenous languages in Asia growing at the same rate as the overall population in Asia.

global peak in linguistic diversity was reached centuries before 1970. That context—the knowledge that most of the world's linguistic diversity was lost before we were even able to start measuring it—must always be kept in mind when interpreting the ILD.

Data Quality: Is the ILD Valid? *Ethnologue* is widely recognized as the most authoritative source of information on the number of speakers of the world's languages. In 1992, at the dawn of concern over language endangerment, linguist Michael Krauss called it "by the far the best single source available" on the number of languages and their speakers globally (Krauss 1992:4, n1). That assessment has not changed: in 2007, the editor of the *Encyclopedia of the World's Endangered Languages* referred to it as the "most comprehensive compendium of the world's languages that has yet been produced..." (Moseley 2007b:x).

Nonetheless, these experts—and *Ethnologue's* compilers themselves—also acknowledge that the quality of its data is uneven. *Ethnologue* draws its speaker data from a wide variety of sources, "everything from popular reference books to missionary field reports to specialist monographs by professional linguists" (Harmon 1995:12). Even data taken from government censuses, which might be taken as reliable on their face, are in fact often inaccurate when it comes to reporting language statistics (see, for example, Voegelin and Voegelin 1977:8; Garza Cuarón and Lastra 1991:94, 96; and esp. Skutnabb-Kangas 2000:30–32, and the cites thereunder). The problem is underlined by the current *Ethnologue* editor, M. Paul Lewis, who writes that calculating the number of speakers "is probably the most difficult component of the language information for us to stay on top of" (Lewis, pers. comm., 25 May 2009).[12]

Given these difficulties, it is reasonable to ask whether the underlying data are so inaccurate as to make the ILD (or any time-series linguistic diversity index) invalid. We think that the answer is no, for several reasons:

There is no reason to think that there is a systematic bias towards either overcounting or undercounting the number of mother-tongue speakers within the *Ethnologue* dataset upon which the ILD is based. There are many reasons why a particular datapoint may be an overcount (e.g., the enumerator simply reported, without investigation, the entire ethnic group as mother-tongue speakers) or an undercount (e.g., the enumerator was unaware of the existence of additional mother-tongue speakers elsewhere). We are not aware of any evidence that shows one of these types of error being more prevalent than the other within *Ethnologue* (or any other data source that we have consulted to date). Indeed, one might instead argue that precisely because there has been no systematic means for counting mother-tongue speakers, there is no reason to think that the results are systematically biased one way or the other. If it could be shown that enumerator errors consistently tended (or are likely to tend) toward either overcounting or undercounting, then the ILD trendline would indeed be invalid (absent statistical adjustment to correct for the bias). If there is no systematic bias one way or the other, then—given a large enough sample size, which we believe ours is—it is reasonable to assume that instances of overcounting and undercounting would, on average, cancel each other out.

---

[12] For further discussion of *Ethnologue* data quality and its ramifications for trend analyses, see Harmon 1995.

This caveat about sample size is important. The ILD methodology is designed to measure average trends in large groups of languages. This means that inaccuracies in the time series for any one language cannot unduly affect the overall trendline, for the reasons just given. There is no inherent reason why the ILD methodology could not be applied to small groups of languages, but the results would be valid only if it could assured that the data were gathered in a consistent way.

The ILD Global trendline aligns with the large and convincing body of qualitative evidence pointing to a decline in linguistic diversity. A decline of 20% over the period 1970–2005 is an entirely plausible outcome in view of this evidence. Had the ILD shown, say, an increase over the period, or a precipitous global decline, then that would be *prima facie* evidence calling into question the accuracy of the underlying data.

The ILD is premised on there being a one-to-one equivalence between the cumulative number of mother-tongue speakers of the world's languages and the global population; in other words, on the assumption that each person can have only one mother tongue. This premise depends for its validity on a precisely specified definition of "mother tongue"; as discussed earlier, our definition is "that language an individual would speak most often if given free rein to choose." Under this definition, even multilingual people can have only one mother tongue. With this in mind, let us imagine what a perfect global census of the number of mother-tongue speakers would look like:

- The census would be a true temporal snapshot, taking place worldwide over a very short time; say, a single day.
- It would query every single person in the world.
- Each census-taker would have exactly the same understanding of our definition of "mother tongue," and would have the ability to explain it with such fidelity that every respondent would have a 100% identical understanding of our meaning of the term.
- Every multilingual respondent would be able (and willing) to prioritize among his/her languages as to which one is his/her single mother tongue according to our "first preference" definition.
- Every respondent would feel free to answer truthfully, without fear of political, social, economic, or other repercussions, and would indeed answer truthfully.
- The census would be replicated at regular intervals to produce accurate time-series data.

The fact that our imaginary census is obviously unattainable does not mean that the less-than-perfect numbers available to us have no value. For example, an analysis of the 1992 *Ethnologue* data found a reasonably close correspondence between the cumulative number of mother-tongue speakers and the global population at that time (Harmon 1995:12–13). This strongly suggests an underlying plausibility in the *Ethnologue* speaker-totals data.

In the final analysis, it is always possible to dismiss quantitative assessments of complex global phenomena by claiming that the data aren't good enough—or to categorically rule out any quantitative representation of such phenomena on ideological or philosophical grounds. Accepting such criticisms, however, leaves us in the position of likely never being able to say anything very precise about the global status of linguistic diversity.

Other Caveats and Limitations. In the course of developing the ILD, we had the opportunity to present it as a work-in-progress at two international meetings and in a variety of informal discussions with colleagues. In those exchanges, several points emerged repeatedly that are worth sharing here:

- Many people are skeptical of the validity and usefulness of global indices, often because they don't understand their purpose. The technical basis—and inherent limitations—of such indices not only must be carefully explained, but potential political misuses must be acknowledged and warned against.
- Key concepts that underlie the ILD, such as "language extinction" and "mother-tongue speaker," are nuanced and must be carefully qualified.
- Quantitative indicators are a complement to, not a substitute for, in-depth qualitative knowledge of linguistic and cultural diversity.
- Virtually all indigenous peoples who care about the continuation of their traditional culture believe that maintaining their native language is the linchpin (cf. UNDESA 2009:57–59).

While we expect the ILD to prove a useful tool to communities, analysts and academics, policymakers, and the general public, any index is only as good as the underlying data available at the time. *Ethnologue* is the best single source for data on the numbers of speakers of languages around the world, and information from its various editions is an indispensable part of any analysis of recent trends in language demographics. Nonetheless, as we discussed above, *Ethnologue* data come from a variety of primary and secondary sources and are, inevitably, uneven. We believe that *Ethnologue* time-series data are valid, but without question language demographic data in general can be improved. It should be borne in mind when using the initial version of the ILD that better data will, in the future, produce even more accurate trendlines.

It is also important to acknowledge that global indices such as the ILD should be used to provide broad contextual background for policy frameworks, rather than as guidance for on-the-ground policy decisions. No large-scale language index can hope to fully represent the complexities that must be accounted for in any policy affecting individual language communities. Nor can a global or regional index do more than outline the state of linguistic diversity at these levels; much more fine-grained analyses are required to get a complete picture.

As suggested above, quantitative analyses such as the ILD must be supplemented by knowledge derived through other methods. This is especially relevant with respect to languages because most linguistic diversity is tied to traditional knowledge systems of indigenous people. These systems primarily rely on non-quantitative observational science and narrative, often transmitted orally rather than in writing. Therefore, any global numerical index, including the ILD, runs the risk of being irrelevant (or, worse, antithetical) to the needs of indigenous communities if it is not properly qualified as noted above—and, in addition, supplemented by other information that is generated by the communities themselves.

In short, the ILD and similar global indices that deal with potentially controversial phenomena, such as language policy, must carefully be placed in context whenever they are used as an educational or policy-orientation tool, and should never be used as a sole source of information.

Future Development of the ILD. As part of future work, we plan to add data from the 16th edition of *Ethnologue* (Lewis 2009) for our 1,500 sample languages; in fact, we would like to expand the database to achieve complete coverage of all the world's languages. We also hope to be able to enter into the ILD database all available speaker-numbers data from other global compendia of language statistics, such as Voegelin and Voegelin 1977, the series of monographs produced under the editorship of T. Sebeok in the 1960s and 1970s, and the recent *Encyclopedia of the World's Endangered Languages* (Moseley 2007a, and citations thereunder listed in the References), as well as information from UNESCO's *Atlas of the World's Languages in Danger* (UNESCO 2009b) and other UNESCO-led data-gathering efforts. All of these will provide data with which to compare, or add to, those from *Ethnologue*.

But the full potential of the ILD methodology won't be realized until we are able to expand it to include language demographic data in addition to counts of mother-tongue speakers. To fully understand the status of and trends in the world's linguistic diversity, we need to go beyond using language richness (the number of discrete languages) and language distribution as a proxy—although that is where we, of necessity, have had to begin our work with the ILD. For example, it may be possible to create versions of the ILD that address phylogenetic diversity by using data on language family affiliations that are already included in *Ethnologue*. The methodology could also be applied to certain special language categories, thus producing versions such as ILD Creoles or ILD Isolates. There may be scope for incorporating structural diversity into the ILD by drawing on data from the World Atlas of Language Structures (Haspelmath et al. 2005; http://wals.info). Even better understanding will come when we are able to augment speaker-numbers data with deeper knowledge about all the factors that determine language demographics and drive trends in linguistic diversity.

APPENDIX A. TECHNICAL DISCUSSION OF THE ILD DATABASE

ORIGINS. The ILD database has its origins in work done in the mid-1990s in which a shadow database of the 12th edition of *Ethnologue* (B.F. Grimes 1992a) was created and analyzed for demographic trends. The work involved entering into a FileMaker Pro database a variety of demographic information relevant to speaker trends and language viability on all 6,760 languages reported in the 12th edition. Each record represented a discrete language distinguished by a unique three-letter code assigned by *Ethnologue*. The information was used to produce a basic analysis of the demographic structure of the world's languages (Harmon 1995). The ILD database expands on the Harmon 1995 database and is organized on the same principle. It too is keyed to discrete languages as reported in *Ethnologue*: in this case, the 15th (Gordon 2005). It was in this edition that the unique three-letter language identifier codes assigned by the International Standards Organisation (ISO) first came into use.

STRUCTURE. The ILD database is structured around these ISO codes, which follow the ISO 639-3 standard. In terms of quality control, the ISO code is the most critical piece of information in the ILD database because it signifies a discrete language. A number of languages share the same name, many have variant names, and still others have self-names that are different from those that have become established in English (e.g., Magyar = Hungarian). ISO codes avoid confusion by assigning a unique three-letter code to each language that is considered discrete. In the 16th edition of *Ethnologue*, which appeared too late for use in calculating the initial ILD, ISO codes are also assigned to "macrolanguages," defined by ISO as "multiple, closely related individual languages that are deemed in some usage contexts to be a single language" (quoted in Lewis 2009:9). The 16th edition lists 55 such macrolanguages. Arabic [ara], Chinese [zho], Serbo-Croatian [hbs], and Kurdish [kur] are some prominent examples.

The purpose of ISO codes is the same as that of Linnean binomials for biological species: they serve to uniquely identify separate entities no matter what their vernacular names are in different languages. The ISO code is written in lowercase; when we refer to them in this paper, they appear in square brackets.

The ISO codes derive, in part, from *Ethnologue's* earlier proprietary three-letter codes, first published in the 10th edition (B.F. Grimes 1984), which were written in uppercase. As noted by the current *Ethnologue* editor, M. Paul Lewis, the "adoption of the ISO 639-3 standard both 'took over' the previously existing [*Ethnologue*] codes but also involved an alignment of those codes with the already existing [interim] ISO 639-2 code set." This resulted in some confusing re-assignments of the old *Ethnologue* codes to new languages under the ISO code set, "but the ongoing principle of the [ISO 639-3] standard that no code is ever re-assigned (henceforth) or re-used provides the immensely valuable benefit that we can now not only uniquely identify languages but also be able to trace their identification history (what they were split from, merged with, etc.) since all of the ISO codes retain their original denotations" (Lewis, pers. comm., 25 May 2009). *Ethnologue* now follows a prescribed process in which all changes to the language roster—whether additions, deletions, mergers, splits, or name changes—are recorded with ISO and published on the *Ethnologue* website.

The adoption of ISO 639-3 will go a long way toward ending confusion over language names. Moreover, henceforth it should be simple for anyone to see trace reclassifications in linguistic status (e.g., a dialect being elevated to consideration as a discrete language, or vice versa) made by *Ethnologue's* editors from edition to edition. These problems remain, of course, for those like us who wish to retrospectively analyze data from editions of *Ethnologue* prior to the 15th.

BUILDING THE DATABASE

The ILD database was built in a series of steps:

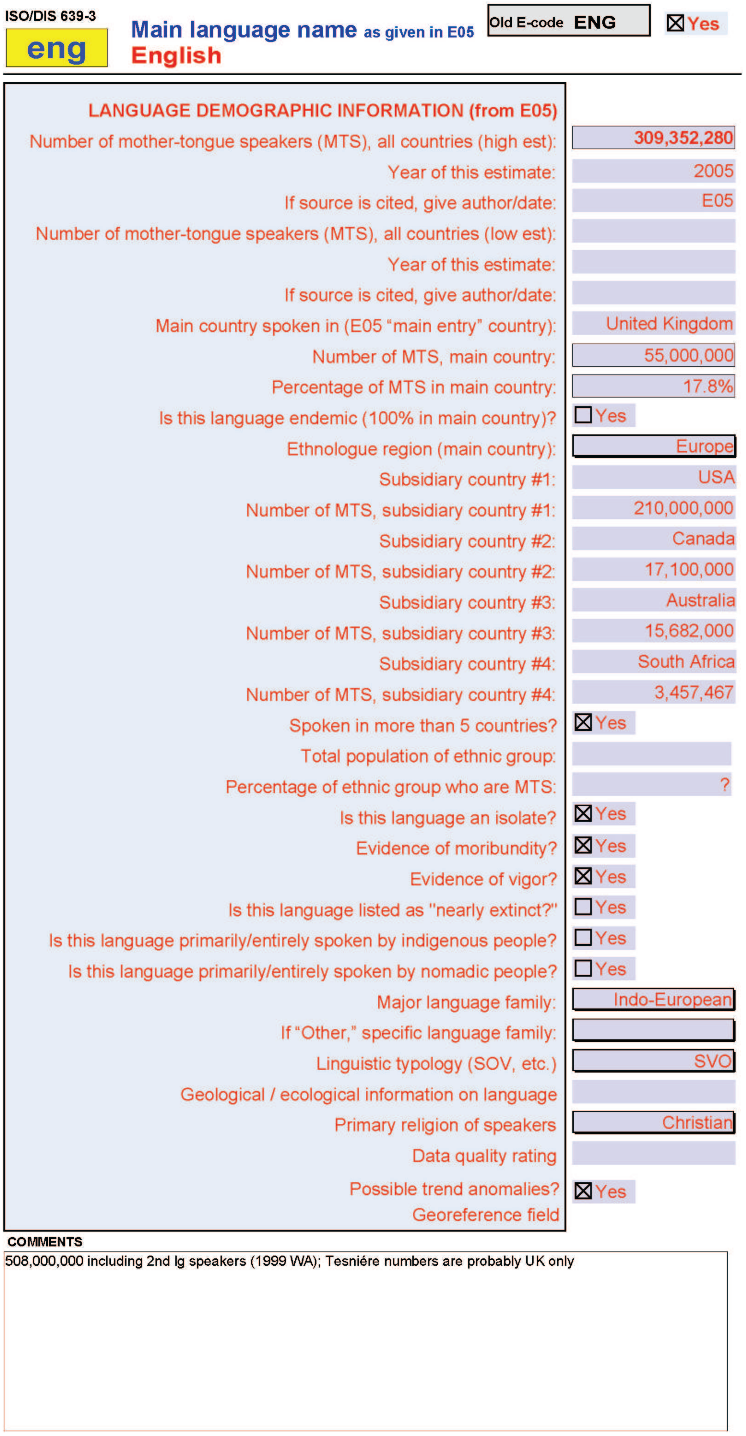
1. Select random sample;
2. Enter base demographic information from the 15th edition of *Ethnologue*;
3. Enter mother-tongue speaker numbers from earlier editions of *Ethnologue*;
4. Analyze sample for representativeness;
5. Eliminate duplicate datapoints;
6. Assess and adjust for possible data trend anomalies;
7. Remove discrepant datapoints; and
8. Deal with apparent extinctions within time series.

1. SELECTION OF RANDOM SAMPLE. The ILD is based on a random sample of 1,500 of the world's 7,299 languages. This sample size was chosen because we determined it to be the largest we could reasonably deal with over the period of project funding. A sample size of 1,500 is far higher than is needed to constitute a statistically representative global sample, but is also allows statistically valid analysis of subglobal samples and provides a cushion against sample attrition (more on this below).

To create the sample, we used the statistics program "R" to generate 1,500 random numbers between 1 and 7,299. An alphabetical list of the 7,299 ISO codes—[aaa] (Ghutuo) through [zyp] (Zyphe)—was imported from FileMaker Pro into an Excel spreadsheet, numbered consecutively, and then the random numbers matched to the ISO codes. The result was a random sample of 1,500 languages.

2. ENTRY OF BASE DEMOGRAPHIC INFORMATION. We extracted all demographic information from the 15th edition of *Ethnologue* for the 1,500 languages in our sample. Figure A-1 shows the form used to record the base demographic information; Table A-1 explains the fields in the form.

FIGURE A-1



ISO/DIS 639-3

eng

Main language name as given in E05

English

Old E-code **ENG**

☒ Yes

LANGUAGE DEMOGRAPHIC INFORMATION (from E05)

Number of mother-tongue speakers (MTS), all countries (high est):	309,352,280
Year of this estimate:	2005
If source is cited, give author/date:	E05
Number of mother-tongue speakers (MTS), all countries (low est):	
Year of this estimate:	
If source is cited, give author/date:	
Main country spoken in (E05 "main entry" country):	United Kingdom
Number of MTS, main country:	55,000,000
Percentage of MTS in main country:	17.8%
Is this language endemic (100% in main country)?	☐ Yes
Ethnologue region (main country):	Europe
Subsidiary country #1:	USA
Number of MTS, subsidiary country #1:	210,000,000
Subsidiary country #2:	Canada
Number of MTS, subsidiary country #2:	17,100,000
Subsidiary country #3:	Australia
Number of MTS, subsidiary country #3:	15,682,000
Subsidiary country #4:	South Africa
Number of MTS, subsidiary country #4:	3,457,467
Spoken in more than 5 countries?	☒ Yes
Total population of ethnic group:	
Percentage of ethnic group who are MTS:	?
Is this language an isolate?	☒ Yes
Evidence of moribundity?	☒ Yes
Evidence of vigor?	☒ Yes
Is this language listed as "nearly extinct?"	☐ Yes
Is this language primarily/entirely spoken by indigenous people?	☐ Yes
Is this language primarily/entirely spoken by nomadic people?	☐ Yes
Major language family:	Indo-European
If "Other," specific language family:	
Linguistic typology (SOV, etc.)	SVO
Geological / ecological information on language	
Primary religion of speakers	Christian
Data quality rating	
Possible trend anomalies?	☒ Yes
Georeference field	

COMMENTS

508,000,000 including 2nd lg speakers (1999 WA); Tesniére numbers are probably UK only

TABLE A-1

Field label as shown on Base Data Entry Form	**Type of field**	**Explanation**
ISO-DIS 639/3	text	The unique three-letter ISO code that identifies each discrete language.
Main language name as given in E[thnologue 20]05	text	The primary name for the language as given in *Ethnologue*. For languages spoken in more than one country, *Ethnologue* generally provides separate entries for each country, with cross-references back to a main entry, which is usually under the country where the language originated. In such instances, we took the primary language name as given in the main entry, and also took all the demographic information from the main entry.
Old E[thnologue] code	text	The unique three-letter code assigned in previous editions of *Ethnologue*. These have been superseded by the ISO codes.
(unlabeled check field, upper righthand corner)	check field	The sorting check field was used to demarcate languages that are part of the random sample of 1,500.
Number of mother-tongue speakers (MTS), all countries (high est):	number	The number of mother-tongue speakers reported for the language. If a range is given, this number is the high estimate.
Year of this estimate:	text	The year the above figure was estimated. If no year was given, “2005” was entered.
If source is cited, give author/ date:	text	If given, a source of the estimate and date for the source. If no source is given, “E05” is entered.
Number of mother-tongue speakers (MTS), all countries (low est):	number	If a range is given, the low estimate of the number of mother-tongues speakers. Otherwise left blank.
Year of this estimate:	text	The year the above figure was estimated. If no year was given, “2005” was entered.
If source is cited, give author/ date:	text	If given, a source of the estimate and date for the source. If no source is given, “E05” is entered.
Main country spoken in (E05 “main entry” country):	text	The country under which the main entry for the language is to be found. This is not always the country in which the most speakers live. For example, for English [eng] the main entry country is not that with the largest number of speakers (USA), but is instead the UK, the language’s country of origin.
Number of MTS, main country:	number	If spoken in more than one country, the number of mother-tongue speakers given under the main entry.

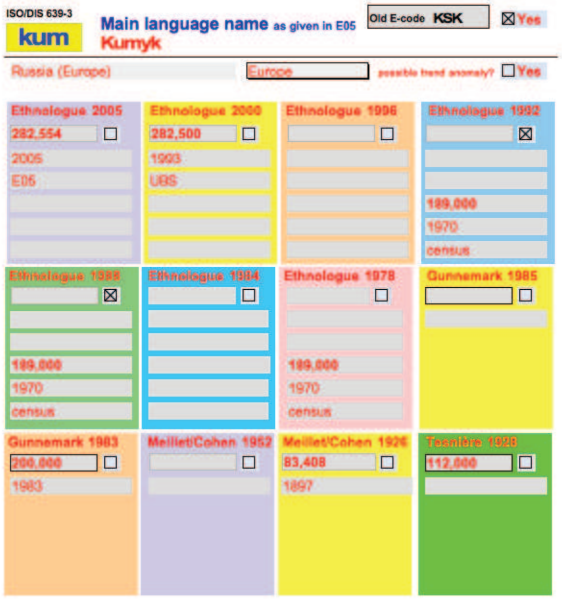
Percentage of MTS in main country:	automatic calculation	[Number of MTS, main country]/[Number of mother-tongue speakers (MTS), all countries (high est)]
Is this language endemic (100% in main country)?	check field	If the above calculation is 100%, the language is considered endemic and this field is checked.
Ethnologue region (main country):	drop-down text menu	*Ethnologue* is organized according to five regions: Africa, Americas, Asia, Europe, and Pacific; for the ILD, we separated out Australia from the Pacific.
Subsidiary country #1:	text	If spoken in more than one country, the name of the first subsidiary country listed under the main entry.
Number of MTS, subsidiary country #1:	number	Number of mother-tongue speakers in first subsidiary country.
Subsidiary country #2:	text	The name of the second subsidiary country listed under the main entry.
Number of MTS, subsidiary country #2:	number	Number of mother-tongue speakers in second subsidiary country.
Subsidiary country #3:	text	The name of the third subsidiary country listed under the main entry.
Number of MTS, subsidiary country #3:	number	Number of mother-tongue speakers in third subsidiary country.
Subsidiary country #4:	text	The name of the fourth subsidiary country listed under the main entry.
Number of MTS, subsidiary country #4:	number	Number of mother-tongue speakers in fourth subsidiary country.
Spoken in more than 5 countries?	check field	Checked if “yes.”
Total population of ethnic group:	number	If given by *Ethnologue*, the total number in the ethnic group.
Percentage of ethnic group who are MTS:	automatic calculation	[Percentage of ethnic group who are MTS]/[Number of mother-tongue speakers (MTS), all countries (high est)]
Is this language an isolate?	check field	Checked “yes” if language is considered an isolate (unrelated to any other language).
Evidence of moribundity?	check field	Checked “yes” if *Ethnologue’s* description of the language shows any indications of moribundity. For further explanation, see text.
Evidence of vigor?	check field	Checked “yes” if *Ethnologue’s* description of the language shows any indications of vigor. For further explanation, see text.
Is this language listed as “nearly extinct?”	check field	Checked “yes” if *Ethnologue* lists the language as “nearly extinct.”

Is this language primarily/entirely spoken by indigenous people?	check field	Checked if our analysis determined that the language is spoken by an indigenous people. For further explanation, see text.
Is this language primarily/ entirely spoken by nomadic people?	check field	Checked "yes" if *Ethnologue* description indicates that the speakers are nomads/mobile peoples.
Major language family:	drop-down text menu	*Ethnologue* assigns languages to one of the following categories: Afro-Asiatic, Austronesian, Indo-European, Language isolate, Niger-Congo, Sino-Tibetan, Trans-New Guinea, or Other.
If "Other," specific language family:	drop-down text menu	If the language falls into the "Other" category, it is assigned to one of the following subcategories: Alacalufan, Algic, Altaic, Amto-Musan, Andamanese, Araun, Araucanian, Arawakan, Artificial Language, Arutani-Sape, Australian, Austro-Asiatic, Aymaran, Barbacoan, Basque, Bayono-Awbono, Caddoan, Cahuapanan, Carib, Chapacura-Wanham, Chibchan, Chimakuan, Choco, Chon, Chukotko-Kamchatkan, Chumash, Creole, Deaf Sign Language, Dravidian, East Bird's Head, East Papuan, Eskimo-Aleut, Geelvink Bay, Guahiban, Harakmbet, Hmong-Mien, Hokan, Huavean, Iroquoian, Japanese, Jivaroan, Kartvelian, Katukinan, Keres, Khoisan, Kiowa Tanoan, Kwomtari-Baibai, Language Isolate, Left May, Lower Mamberamo, Lule-Vilela, Macro-Ge, Maku, Mascoian, Mataco-Guaicuru, Mayan, Misumalpan, Mixe-Zoque, Mixed Language, Mura, Muskogean, Na-Dene, Nambiquaran, Nilo-Saharan, North Caucasian, Oto-Manguean, Panoan, Peba-Yaguan, Penutian, Pidgin, Quechuan, Salishan, Salivan, Sepik-Ramu, Siouan, Sko, Subtiaba-Tlapanec, Tacanan, Tai-Kadai, Tarascan, Torricelli, Totonacan, Tucanoan, Tupi, Unclassified, Uralic, Uru-Chipaya, Uto-Aztecan, Wakashan, West Papuan, Witotoan, Yanomam, Yeniseian, Yukaghir, Zamucoan, Zaparoan
Linguistic typology (SOV, etc.)	drop-down text menu	For some languages, *Ethnologue* indicates the linguistic typology: OSV, OVS, SOV, SVO, VOS, or VSO.
Geological/ecological information on language	text	For some languages, *Ethnologue* indicates the general geological/ecological conditions of the main area inhabited by its speakers.

Primary religion of speakers	drop-down text menu	For some languages, *Ethnologue* indicates the primary religion of speakers, using the following list: Buddhist (unspecified), Buddhist (Lamaist), Christian, Confucianism, Daoist, Hindu, Jewish, Mandaism, Muslim (unspecified), Muslim (Alevi), Muslim (Shi'a), Muslim (Sunni), Polytheist, Shamanist, Syncretism, Traditional Religion, Zoroastrianism
Data quality rating	(undetermined)	Country entries in some pre-2005 editions of *Ethnologue* contained a simple A–D data quality rating. This was discontinued in the 2005 edition, but we are reserving such a field for possible future use in the database.
Possible trend anomalies?	check field	For further explanation, see text.
Georeference field	(undetermined)	We are reserving a field that would be tied to current work on mapping biocultural diversity being undertaken by Terralingua, which includes *Ethnologue's* GIS coordinates for the world's languages.

3. ENTRY OF MOTHER-TONGUE SPEAKER INFORMATION. We searched the nine editions of *Ethnologue* used as the basis of the initial ILD for the number of mother-tongue speakers of the 1,500 languages in our sample. Figure A-2 shows the form used to record the base demographic information; Table A-2 explains the fields in the form.

FIGURE A-2



ISO/DIS 639-3
kum
Main language name as given in E05
Kumyk
Old E-code KSK
☒ Yes
Russia (Europe)
Europe
possible trend anomaly? ☐ Yes

Ethnologue 2005
282,554 ☐
2005
E05

Ethnologue 2000
282,500 ☐
1993
UBS

Ethnologue 1996
☐

Ethnologue 1992
☒
189,000
1970
census

Ethnologue 1988
☒
189,000
1970
census

Ethnologue 1984
☐

Ethnologue 1978
☐
189,000
1970
census

Gunnemark 1985
☐

Gunnemark 1983
200,000 ☐
1983

Meillet/Cohen 1952
☐

Meillet/Cohen 1926
83,408 ☐
1897

Tesnière 1928
112,000 ☐

Table A-2

Field label as shown on form	**Type of field**	**Explanation**
ISO-DIS 639/3	text	The unique three-letter ISO code that identifies each discrete language. For further explanation, see text.
Main language name as given in E[thnologue 20]05	text	The primary name for the language as given in *Ethnologue*. For languages spoken in more than one country, *Ethnologue* generally provides separate entries for each country, with cross-references back to a main entry, which is usually under the country where the language originated. In such instances, we took the primary language name as given in the main entry, and also took all the demographic information from the main entry.
Old E[thnologue] code	text	The unique three-letter code assigned in previous editions of *Ethnologue*. These have been superseded by the ISO codes.
(unlabeled field, left side, just beneath horizontal line) Main country spoken in (E05 “main entry” country):	text	The country under which the main entry for the language is to be found. This is not always the country in which the most speakers live. For example, for English [eng] the main entry country is not that with the largest number of speakers (USA), but is instead the UK, the language’s country of origin.
(unlabeled field, center, just beneath horizontal line) Ethnologue region (main country):	drop-down text menu	*Ethnologue* is organized according to five regions: Africa, Americas, Asia, Europe, and Pacific; for the ILD, we separated out Australia from the Pacific.
Possible trend anomaly?	check field	For further explanation, see text.
Description of fields in *Ethnologue* edition source blocks (Ethnologue 2005, Ethnologue 2000 ... Ethnologue 1951), from top to bottom		
Number of mother-tongue speakers (MTS), all countries (high est):	number	The number of mother-tongue speakers reported for the language. If a range is given, this number is the high estimate.
Year of this estimate:	text	The year the above figure was estimated. If no year was given, the year of the edition is entered.
Source of estimate:	text	If given, a source of the estimate.

Number of mother-tongue speakers (MTS), all countries (low est):	number	If a range is given, this is the low estimate of the number of mother-tongue speakers. If a number is given here but no high estimate is given above, it means that this number represents a minimum estimate. If neither of foregoing two conditions applies, the field is left blank.
Year of this estimate:	text	The year the above figure was estimated. If no year was given, the year of the edition is entered.
Source of estimate:	text	If given, a source of the estimate.
(unlabeled check box) Duplicate datapoint control check field:	check field	If the figures reported in this block are identical (in both number of speakers reported and in terms of source citation), then this datapoint is a duplicate and is omitted from the trend analysis.

4. ANALYSIS OF SAMPLE REPRESENTATIVENESS. The 15th edition of *Ethnologue* (Gordon 2005) provides global statistics for three language demographic variables that we used to assess our sample's representativeness: language size, language family, and main region of the language. We compared our sample to the global total for the three variables, and found that it is closely representative of the global distribution for all three variables (Table A-3).

TABLE A-3.1: Representativeness by language size

Representativeness by language size (extinct languages excluded)									
Number of mother-tongue speakers per language	**0**	**1–100**	**101–1,000**	**1,001–10,000**	**10,001–100,000**	**100,001–1,000,000**	**>1,000,000**	**No data**	**Total**
Number of languages, ILD sample	73	129	216	430	363	171	65	53	1,500
% of ILD sample	4.9	8.6	14.4	28.7	24.2	11.4	4.3	3.5	100.0
Number of languages, global total	387	548	1,071	1,967	1,779	892	347	308	7,299
% of global total	5,3	7.5	14.7	26.9	24.4	12.2	4.8	4.2	100.0

Table A-3.2: Representativeness by language family

Representativeness by language family (extinct languages excluded)				
	Ethnologue 2005	**% of global total**	**ILD sample**	**% of sample**
Major lg families				
Afro-Asiatic	353	5.11	67	4.70
Austronesian	1,246	18.03	242	16.96
Indo-European	430	6.22	102	7.15
Niger-Congo	1,495	21.63	295	20.67
Sino-Tibetan	399	5.77	93	6.52
Trans-New Guinea	561	8.12	112	7.85
	4,484	64.87	911	63.84
Other lg families & classifications				
Alacalufan	1	0.01	0	0.00
Algic	31	0.45	6	0.42
Altaic	64	0.93	17	1.19
Amto-Musan	2	0.03	0	0.00
Andamanese	4	0.06	0	0.00
Araun	7	0.10	1	0.07
Araucanian	2	0.03	1	0.07
Arawakan	49	0.71	14	0.98
Artificial Language	1	0.01	1	0.07
Arutani-Sape	2	0.03	0	0.00
Australian	224	3.24	46	3.22
Austro-Asiatic	169	2.45	36	2.52
Aymaran	3	0.04	1	0.07
Barbacoan	5	0.07	1	0.07
Basque	3	0.04	0	0.00
Bayono-Awbono	2	0.03	0	0.00
Caddoan	4	0.06	1	0.07
Cahuapanan	2	0.03	1	0.07
Carib	29	0.42	5	0.35
Chapacura-Wanham	4	0.06	0	0.00
Chibchan	21	0.30	6	0.42
Chimakuan	1	0.01	0	0.00
Choco	7	0.10	1	0.07
Chon	2	0.03	0	0.00

Chukotko-Kamchatkan	5	0.07	1	0.07
Creole	82	1.19	20	1.40
Deaf Sign Language	119	1.72	25	1.75
Dravidian	73	1.06	19	1.33
East Bird's Head	3	0.04	1	0.07
East Papuan	33	0.48	6	0.42
Eskimo-Aleut	10	0.14	1	0.07
Geelvink Bay	33	0.48	7	0.49
Guahiban	5	0.07	2	0.14
Harakmbet	2	0.03	0	0.00
Hmong-Mien	35	0.51	5	0.35
Hokan	19	0.27	4	0.28
Huavean	4	0.06	0	0.00
Iroquoian	7	0.10	1	0.07
Japanese	12	0.17	1	0.07
Jivaroan	4	0.06	0	0.00
Kartvelian	5	0.07	0	0.00
Katukinan	3	0.04	1	0.07
Keres	2	0.03	0	0.00
Khoisan	22	0.32	3	0.21
Kiowa Tanoan	5	0.07	2	0.14
Kwomtari-Baibai	6	0.09	0	0.00
Language Isolate	36	0.52	8	0.56
Left May	6	0.09	1	0.07
Lower Mamberamo	2	0.03	0	0.00
Lule-Vilela	1	0.01	1	0.07
Macro-Ge	24	0.35	3	0.21
Maku	6	0.09	1	0.07
Mascoian	4	0.06	0	0.00
Mataco-Guaicuru	11	0.16	1	0.07
Mayan	68	0.98	15	1.05
Misumalpan	2	0.03	0	0.00
Mixe-Zoque	17	0.25	8	0.56
Mixed Language	19	0.27	4	0.28
Mura	1	0.01	0	0.00
Muskogean	6	0.09	2	0.14

Na-Dene	41	0.59	9	0.63
Nambiquaran	3	0.04	0	0.00
Nilo-Saharan	197	2.85	40	2.80
North Caucasian	33	0.48	7	0.49
Oto-Manguean	172	2.49	35	2.45
Panoan	19	0.27	4	0.28
Peba-Yaguan	1	0.01	0	0.00
Penutian	23	0.33	7	0.49
Pidgin	5	0.07	3	0.21
Quechuan	45	0.65	9	0.63
Salishan	19	0.27	5	0.35
Salivan	3	0.04	0	0.00
Sepik-Ramu	100	1.45	24	1.68
Siouan	12	0.17	6	0.42
Sko	7	0.10	2	0.14
Subtiaba-Tlapanec	4	0.06	1	0.07
Tacanan	6	0.09	2	0.14
Tai-Kadai	74	1.07	14	0.98
Tarascan	2	0.03	0	0.00
Torricelli	53	0.77	12	0.84
Totonacan	11	0.16	1	0.07
Tucanoan	20	0.29	5	0.35
Tupi	60	0.87	13	0.91
Unclassified	43	0.62	11	0.77
Uralic	36	0.52	5	0.35
Uru-Chipaya	2	0.03	0	0.00
Uto-Aztecan	56	0.81	8	0.56
Wakashan	4	0.06	1	0.07
West Papuan	26	0.38	6	0.42
Witotoan	6	0.09	3	0.21
Yanomam	4	0.06	0	0.00
Yeniseian	2	0.03	0	0.00
Yukaghir	2	0.03	1	0.07
Zamucoan	2	0.03	0	0.00
Zaparoan	4	0.06	1	0.07
	2,428	35.13	516	36.16

TABLE A-3.3: Representativeness by region

Representativeness by region (extinct languages excluded)						
	Africa	**Americas**	**Asia**	**Europe**	**Pacific**	**Total**
Ethnologue 2005	2,092	1,002	2,269	239	1,310	6,912
% of global total	30.3	14.5	32.8	3.5	19.0	100.0
ILD sample	408	226	476	44	273	1,427
% of sample	28.6	15.8	33.4	3.1	19.1	100.0

5. ELIMINATION OF DUPLICATE DATAPOINTS. As we entered mother-tongue speaker information, we analyzed each datapoint to see if it was unique (i.e., represented new data) or a duplicate of an earlier datapoint. In terms of developing time-series data on speaker numbers, it would have been ideal if each of our 11,253 data searches had produced a unique datapoint. The reality is far from the ideal, however, and one reason to have a sample size much larger than the minimum required for statistical validity is to account for attrition: in our case, languages having to be excluded from the ILD because they have fewer than two unique datapoints from which to construct a trend. Because of the paucity of speaker statistics for many languages, this is not an uncommon occurrence. It is standard practice for *Ethnologue* to carry over estimates from earlier editions if a newer (and therefore presumably more current estimate) is unavailable. Out of our sample of 1,500 languages, 391 had to be excluded from the ILD calculation because they had either no speaker totals listed in any of the editions of *Ethnologue* we consulted, or else had only one unique datapoint. We can expect this situation to improve with future editions of *Ethnologue*, for there is now a sustained effort by the editors to report speaker totals from as many languages as possible. (A major gap is deaf languages, for which speaker totals are rarely reported.) After all steps of the data analysis were completed, we were left with 2,703 unique datapoints from which we created the initial version of the ILD.

6. ASSESSMENT AND TREATMENT OF POSSIBLE TREND ANOMALIES. It is not uncommon for successive estimates of speaker numbers for a particular language to vary widely. To account for this, we assessed the time series for all 1,500 languages in our sample for possible trend anomalies: large or rapid changes in the reported numbers of mother-tongue speakers within the chronological sequence of estimates for that language.

There are many reasons why a particular datapoint in a time series could possibly be anomalous. It might reflect some kind of major difference in the way the speakers were counted, or in interpretation of what constitutes the language itself, or in who qualifies as a mother-tongue speaker. Perhaps the datapoint in question could simply reflect an incomplete count of speakers despite the researcher having canvassed all known locations where speakers live. Or it could reflect an incomplete count of speakers because the researcher failed to canvass all known locations. It could even be (though it is more unlikely) that the seemingly discrepant datapoint is, in fact, accurate and all the others are wrong—for example, maybe all the other datapoints included second-language speakers.

Not all languages will show linear trends in their speaker numbers, and this in itself is not a reason to suspect a possible trend anomaly. It may be that a particular language's numbers truly are fluctuating. Similarly, some languages may show an unbroken upward or downward trend, but within that trend there will be huge jumps or declines that might lead one to question the accuracy of the numbers. In all these situations, the controlling questions are, what is the magnitude of the reported change, how quickly is it happening, and how plausible is it relative to the size of the language?

In terms of plausibility, a cardinal principle is that smaller percentage changes are more plausible across the board, no matter if the language has 100 speakers to start with or 1,000,000. We can easily imagine a small language going from 100 to 99 speakers over a 20-year period, and just as easily imagine a language with 1,000,000 going to 990,000 over the same period. However, as the percentage changes become larger, the plausibility of those changes depends on how large the language is initially. For a language starting out at 100 speakers, it is plausible to imagine a situation in which it declines 90% over the 20-year period, going from 100 to 10 speakers. Perhaps most or all of the speakers were old to begin with (a not uncommon occurrence in such cases) and they died over the period while at the same time no children were being brought up using the language as their mother tongue. Or perhaps there was a catastrophe that struck the village where all the speakers lived, causing most of the speakers to die. These are plausible scenarios. But it is far less plausible to see how a language could go from 1,000,000 speakers to 100,000 in just 20 years. So a corollary point is that the plausibility of changes in speaker numbers declines as (a) the percentage of change increases, (b) the language size increases, and (c) the time period over which the change is said to occur decreases. That is, large percentage changes in the size of large languages over short periods of time are the least plausible.

We identified possible trend anomalies by calculating the rate of change in number of speakers between one datapoint and the next. Three degrees of possible anomaly were identified: differences between successive datapoints that represented a rate of change equivalent to a doubling or halving in number over a period of (a) ten years, (b) five years, or (c) three years. Languages where numbers of speakers were below 1,000 were excluded because, as just noted, very small populations are liable to undergo rapid fluctuations.

For our assessment of possible trend anomalies within the ILD's 1,500-language sample, we analyzed all instances flagged by the 3-year doubling/halving filter. There were 157 languages having such instances (Table A-4). Our analysis consisted of:

- Identifying the flagged datapoint(s).
- Assessing the likelihood of that datapoint being anomalous. This involved a number of considerations, including size of the language, keeping in mind that, in general, the smaller the language, the easier it is to accurately count its number of speakers (Voegelin and Voegelin 1977:8); special qualities of certain data sources used by *Ethnologue*, based on our experience in working with the database (e.g., long-time observation has shown estimates from certain data sources cited by *Ethnologue* tend to run higher or lower than others cited in other editions of *Ethnologue*); and whether or not the trendline for the language contains one possible anomalous datapoint or several.
- Excluding the anomalous datapoint, if necessary.

For datapoints that we assessed as being "definitely anomalous" or "probably anomalous," the datapoint was excluded from the ILD. This has the effect of smoothing the trendline. For datapoints assessed as "possibly anomalous," or "may not be anomalous," the decision to exclude or not varied depending on our judgment, using the considerations outlined above: some were excluded, while some were left unchanged. When in doubt, our policy was to leave the data unchanged. See Table A-4.

TABLE A-4

ISO	**Language name**	**Is data trend anomalous?**	**Reason**	**Data treatment**
afb	Arabic, Gulf Spoken	probably anomalous	decline from E92 to later estimates too steep to be plausible in a language with millions of speakers	disregard E92 as outlier
ald	Alladian	probably anomalous	Taber's estimates generally run low	disregard E69 as outlier; all other estimates show steady rise in numbers
amx	Anmatyerre	possibly anomalous	Wurm and Hatori estimate very low in comparison to Black's	disregard Wurm and Hatori estimate (1981) because subsequent editions endorse Black's 1983 estimate
apl	Apache, Lipan	may not be anomalous	extremely low speaker numbers might explain pecentage decline	leave data unchanged
apm	Apache, Mescalero-Chiricahua	probably anomalous	unlikely that [apm] gained 800 speakers between 1969 and 1978	disregard E69 because subsequent editions endorse E78 estimate
arg	Aragonese	probably anomalous	unlikely that [arg] declined by 19,000 speakers between 1989 and 1993; prior to E92 had been lumped in with Spanish	disregard E92 because two subsequent editions give identical estimates from 2 different sources
asb	Assiniboine	probably anomalous	unlikely that [asb] declined from 1000-2000 in 1969 to 100 in 1977; also possible that it did not actually increase from 100 in 1977 to 150-200 in 1986	disregard E69 as outlier because subsequent editions give much more comparable estimates
ask	Ashkun	may not be anomalous	*Ethnologue* estimate of 7000 held from E78 through E2000; since no external data sources given, it appears that the E78 esimate was simply carried over to subsequent editions	disregard E84, E88, E92, E96, and E00; use 2-datapoint trendline: E78 and E05

aue	=Kx’au||´ein	probably anomalous	unlikely that [aue] went from 4890 in 1977 to 3000 in 1991 and then back up to 5000 in 1993	disregard E92 as outlier
bae	Baré	possibly anomalous	debatable that [bae] went from 263 speakers in 1988 to 0 in 2005, but not inconceivable	leave data unchanged
bis	Bislama	possibly anomalous	debatable that [bis] went from 1200 in 2000 to 6200 in 2005, but not inconceivable	leave data unchanged
bjl	Bulu (Papua New Guinea)	probably anomalous	unlikely that [bjl] went from 200 speakers in 1978 to 566 in 1982	disregard E78 as outlier
bjz	Baruga	probably anomalous	unlikely that [bjz] went from 4000-6000 in 1969 to 1051 in 1971	disregard E69 as outlier
bpp	Kaure	probably anomalous	unlikely that [bpp] spiked at 4000 in 1991 when other estimates are less than 1000	disregard E92 as outlier
bra	Braj Bhasha	probably anomalous	unlikely that [bra] declined from 11+million in 1977 to 44500 in 1997	disregard E78 in favor of more recent datapoint (which is repeated in E05)
brn	Boruca	probably anomalous	unlikely that [brn] went from 500 in in 1978 to 5 in 1986	disregard E78 as outlier
bsr	Bassa-Kontagora	may not be anomalous	datapoints in E00 and E05 revert to a count of 10 in 1987, and E05 says [bsr] is extinct, or nearly extinct; less conservatively, E92 give a count of 0	disregard E78, use 2-datapoint trendline: E00 (10 in 1987) and E92 (0 in 1992)
bwt	Bafaw-Balong	definitely anomalous	E78 and E69 estimates are for Balong only	disregard E78 and E69
bzp	Kemberano	probably anomalous	unlikely that [bzp] went from 150 in 1978 to 1500 in 1987	disregard E78 because subsequent editions endorse E92
caz	Canichana	may not be anomalous	E92 estimate of 25 is dated to 1958; E69 estimate of 25 is dated to 1968; E78 and E88 estimates of 25 are undated, so are listed in the database as being from 1978 and 1988, but this is likely misleading	disregard E78 and E88; use 3-datapoint trendline: E92, E69, and E00
cbb	Cabiyarí	may not be anomalous	fluctuations of this magnitude are conceivable with a small lg like [cbb]	leave data unchanged

cbn	Nyahkur	probably anomalous	discrepancies between E92 and E78, and between the high/low average of E92 and E00 and E05, appear to be of too great a magnitude to be true demographics	disregard E92 and E78
cbr	Cashibo-Cacataibo	probably anomalous	E00 estimate is undated and therefore attributed to 1998, but probably dates from earlier	disregard E00; use E69, E78, and E05 as trendline
cch	Atsam	probably anomalous	unlikely that [cch] increased from 8500 in 1969 to 35000 in 1972	disregard E69 as outlier
cku	Koasati	possibly anomalous	debatable increase to, and then decline from 1996 datapoint	leave data unchanged
cle	Chinantec, Lealao	possibly anomalous	debatable that [cle] went from 3500/5000 in 1978 t0 800/900 in 1982	disregard E78 as outlier
cod	Cocama-Cocamilla	probably anomalous	no discernable pattern: estimates vary widely and cannot be reconciled	disregard all datapoints
cun	K'iché, Cunén	may not be anomalous	E05 notes "significant monolingualism" and only a slight move toward Spanish; this suggests E78 estimate (repeated in E88) is an undercount	disregard E78 and E88
dal	Dahalo	probably anomalous	unlikely decrease from 1987 to 1992 datapoints	disregard E92
djm	Dogon, Jamsay	probably anomalous	unlikely decrease from 1995 to 1998 datapoints	disregard E00
dng	Dungan	probably anomalous	unlikely increase from 1969 to 1970 datapoints	disregard E69 as outlier
dor	Dori'o	probably anomalous	unlikely increase from 1998 to 1999 datapoints	disregard E00 as outlier
dyi	Senoufo, Djimini	probably anomalous	unlikely increase from 1991 to 1993 datapoints	disregard E92
eee	E	probably anomalous	unlikely increase from 1990 to 1992 datapoints	disregard E92
eot	Beti (Cote d'Ivoire)	probably anomalous	unlikely decrease from 1966 to 1977 datapoints	disregard E69
faf	Fagani	probably anomalous	unlikely increase from 1998 to 1999 datapoints	disregard E00
fgr	Fongoro	definitely anomalous	thought extinct in E92; more recent editions reference 1983 estimate	disregard E92
fip	Fipa	probably anomalous	unlikely increase from 1990 to 1992 datapoints	disregard E92

fri	Frisian, Western	probably anomalous	unlikely decrease from 1976 to 1978 datapoints	disregard E78; construct 2-datapoint trendline using E05 (1976) and E88 (1988)
gid	Gidar	probably anomalous	unlikely increase from 1966 to 1967 datapoints	disregard E69
glk	Gilaki	probably anomalous	unlikely increase from 1991 to 1993 datapoints	disregard E92
goa	Guro	probably anomalous	unlikely decrease from 1966 to 1967 datapoints	disregard E78
gvf	Golin	probably anomalous	unlikely increase from 1978 to 1981 datapoints	disregard E78
gvl	Gulay	probably anomalous	unlikely increase from 1990 to 1993 datapoints	disregard E92
gyi	Gyele	probably anomalous	unlikely increase from 2000 to 2005 datapoints	disregard E00
hae	Oromo, Eastern	possibly anomalous	debatable increase from 1978 to 1998 datapoints	leave data unchanged
hbn	Heiban	probably anomalous	unlikely increase from 1966 to 1972 datapoints	disregard E78 as outlier
hio	Tsoa	probably anomalous	unlikely increases between 1976 and 1977 datapoints and between 2000 and 2004 datapoints	disregard E88 and E00; construct 2-datapoint trendline using E78 and E05
hmd	Hmong, Northeastern	probably anomalous	unlikely decrease from 1982 to 1987 datapoints	disregard E00
hsb	Sorbian, Upper	possibly anomalous	debatable decrease from 1991 to 1196 datapoints	leave data unchanged
huc	ǀHua	possibly anomalous	debatable increase from 1966 to 1978 datapoints	disregard E00 as a duplicate datapoint; otherwise leave data unchanged
huu	Huitoto, Murui	probably anomalous	unlikely increase from 1969 to 1976 datapoints and from 1976 and 1982 datapoints	disregard E69 and E78 as probable underestimations
igl	Igala	probably anomalous	unlikely increase from 1969 to 1973 datapoints	disregard E69
ilb	Ila	probably anomalous	unlikely decrease to, and then increase from, 1973 datapoint	disregard E78 as outlier
iru	Irula	probably anomalous	trendline highly variable; E05 vastly out of line with all previous estimates	disregard E78 and E05
itv	Itawit	probably anomalous	unlikely decrease from 1969 to 1973 datapoints	disregard E78 as outlier
izi	Izi-Ezaa-Ikwo-Mgbo	probably anomalous	unlikely increase from 1969 to 1973 datapoints	disregard E69

jai	Jakalteco, Western	probably anomalous	unlikely increases from 1988 to 1992 and again from 1992 to 2000	disregard E88 and E92
jeg	Jeng	probably anomalous	unlikely increase from 1978 to 1981 datapoints	disregard E78; construct 3-datapoint trendline using E69, E92, and E05
kav	Katukina	probably anomalous	unlikely decrease from 1969 to 1976 datapoints	disregard E69
kca	Khanty	possibly anomalous	debatable decrease from 1969 to 1970 datapoints	leave data unchanged
kdr	Karaim	possibly anomalous	debatable decrease from 1969 to 1977 datapoints	leave data unchanged
khg	Tibetan, Khams	probably anomalous	unlikely increase from 1977 to 1987 datapoints	disregard E78 as outlier
khy	Kele (Democratic Repulic of Congo)	probably anomalous	unlikely increase from 1971 to 1980 datapoints	disregard E78
kia	Kim	probably anomalous	unlikely increase from 1991 to 1993 datapoints	disregard E92
kll	Kalagan, Kagan	possibly anomalous	debatable decrease from 1969 to 1977 datapoints	disregard E78
kln	Kalenjin	probably anomalous	E92, E78, and E69 all appear to be underestimations	disregard E69, E78, and E92
kng	Koongo	probably anomalous	E78 and E88 appear to be overestimations (and E05 discards E88 estimate dating from 1987 in favor of a lower estimate dating from 1986)	disregard E78 and E88
kou	Koke	probably anomalous	unlikely decrease from 1969 to 1971 datapoints	disregard E69
krk	Kerek	possibly anomalous	debatable decrease from 1975 to 1991 datapoints	leave data unchanged
ksi	Krisa	probably anomalous	unlikely increase from 1969 to 1972 datapoints	disregard E69 as outlier
kum	Kumyk	probably anomalous	unlikely increase from 1969 to 1970 datapoints	disregard E69 as outlier
kwz	Kwadi	probably anomalous	E78 datapoint conflicts with note in E05 that says [kwr] had 3 speakers in 1971 and was considered by J.C. Winter in 1981 to have been extinct by then; since E00 listed as extinct	disregard E78; perhaps we could construct trendline with 3 speakers in 1971 and 0 in 1981?
kyr	Kuruáya	definitely anomalous	E78 datapoint obviously erroneous	disregard E78
kzw	Karirí-Xocó	possibly anomalous	debatable decrease to 0 (1978; E78) from 163 (1969; E69), but conceivable	leave data unchanged

lbo	Laven	probably anomalous	unlikely increase from 1978 to 1981 datapoints	disregard E78
leb	Lala-Bisa	probably anomalous	unlikely increase from 1969 to 1973 datapoints	disregard E69
lez	Lezghian	probably anomalous	unlikely increase from 1969 to 1970 datapoints	disregard E69
lga	Lungga	may not be anomalous	trendline looks plausible	leave data unchanged
lif	Limbu	probably anomalous	unlikely increase from 1967 to 1971 datapoints; debatable increase from 1998 to 2005 datapoints	disregard E69; leaves others unchanged
lma	Limba, East	probably anomalous	unlikely increase from 1991 to 1993 datapoints	disregard E92
lpa	Lelepa	possibly anomalous	unlikely decrease from 1983 to 1989 datapoints; but not inconceivable	leave data unchanged
mbb	Manobo, Western Bukidnon	probably anomalous	unlikely decrease from 1969 to 1977 datapoints	disregard E69
mco	Mixe, Coatlán	probably anomalous	trendline highly variable	disregard all datapoints??
mdt	Mbere	probably anomalous	unlikely increase from 1990 to 2005 datapoints	disregard E92
mei	Midob	probably anomalous	unlikely increase from 1977 to 1983 datapoints and from 1983 to 1993 datapoints	disregard E78 and E92?
mez	Menominee	possibly anomalous	debatable decrease from 1969 to 1977 datapoints	disregard E69
mit	Mixtec, Southern Puebla	possibly anomalous	unlikely decrease from 1977 to 1982 datapoints; but not inconceivable	leave data unchanged
mjj	Mawak	possibly anomalous	unlikely decrease from 1969 to 1981 datapoints; but not inconceivable	leave data unchanged
mjp	Malapandaram	probably anomalous	it appears E78 and E92 were simply repeating the 1961 census figures	disregard E78 and E92
mju	Manna-Dora	probably anomalous	it appears E78 and E92 were simply repeating the 1961 census figures	disregard E78 and E92
mnc	Manchu	probably anomalous	unlikely decrease from E78 high estimate to subsequent estimates	disregard E78
msm	Manobo, Agusan	probably anomalous	unlikely increase from 1977 to 1981 datapoints	disregard E78

mug	Musgu	probably anomalous	unlikely increase from 1969 to 1972 datapoints	disregard E69
mvk	Mekmek			
myx	Masaba	probably anomalous	unlikely increases from 1966 to 1978 datapoints and again from 1990 to 1991 datapoints	disregard E69 and E92
ngu	Náhuatl, Guerrero	definitely anomalous	unlikely increase from 1969 to 1977 datapoints; E78 and E88 contradict each other	disregard E69, E78, and E88
niv	Gilyak	probably anomalous	E78 high / low estimates too far out of line, as is E92 high estimate	disregard E78 and E92 high estimate; construct 3-datapoint trendline out of E92 low estimate and E00 and E05 estimates
nkf	Naga, Inpui	possibly anomalous	debatable increase to, and then decrease from, E92 and E00 estimates	disregard E92 and E00 as probable overestimations
nog	Nogai	probably anomalous	unlikely increase from 1969 to 1970 datapoints	disregard E69
now	Nyambo	probably anomalous	unlikely increase from 1987 to 2005 datapoints	disregard E05 (or else disregard all except E05, if we want to go with the latest estimate as being the most accurate)
nut	Nung (Viet Nam)	probably anomalous	unlikely increase from 1973 to 1981 datapoints	disregard E92 as outlier
nza	Mbembe, Tigon	may not be anomalous	trendline debatable but not implausible	leave data unchanged
nzm	Naga, Zeme	probably anomalous	unlikely increase from 1961 to 1971 datapoints and from 1990 to 1994 datapoints	disregard E69 and E92
ojw	Ojibwa, Western	may not be anomalous	trendline looks plausible	leave data unchanged (keeping in mind E00 is a duplicate datapoint)
ots	Otomi, Estado fr Mexico	probably anomalous	unlikely increase to, and then decrease from, 1978 datapoint	disregard E78 as outlier
paf	Paranawat	possibly anomalous	debatable decrease from 1969 to 1978 datapoints	disregard E69
pcg	Paniya	probably anomalous	unlikely decrease to, and then increase from, 1971 datapoint	disregard E78 as outlier
pch	Pardhan	probably anomalous	unlikely decrease to, and then increase from, 1971 datapoint	disregard E78 as outlier

pis	Pijin	may not be anomalous	debatable increase from 1975 to 1997 datapoints	leave data unchanged
pmu	Panjabi, Mirpur	probably anomalous	unlikely increase from 2000 to 2005 datapoints	disregard E00
pou	Poqomam, Southern	probably anomalous	unlikely increase from 1982 to 1991 datapoints	disregard E88
quu	K'iché, Eastern	probably anomalous	unlikely increase from 1978 to 1982 datapoints; E05 notes that it is spoken by "all ages"	disregard E88 as outlier
rej	Rejang	probably anomalous	unlikely increase from 1977 to 1981 datapoints	disregard E78
she	Sheko	probably anomalous	unlikely increase from 1966 to 1972 datapoints	disregard E69
sih	Zire	definitely anomalous	E00 datapoint erroneous	disregard E00; also, I suspect E69 datapoint (which didn't give a year and so is listed as 1969) is the same as the E78 datapoint, which is sourced to 1939
smn	Inari Sami	possibly anomalous	debatable decrease from 1978 to 1983 datapoints, but conceivable	leave data unchanged
soe	Songomeno	possibly anomalous	unlikely increase from 1971 to 1972 datapoints	disregard E78
syc	Syriac	probably anomalous	unlikely decrease from 1978 to 2000 datapoints	disregard E78
tab	Tabassaran	probably anomalous	unlikely increase from 1969 to 1970 datapoints	disregard E69
tan	Tangale	probably anomalous	unlikely increase from 1969 to 1973 datapoints	disregard E69
tbe	Tanimbili	possibly anomalous	unlikely decrease from 1998 to 1999 datapoints, but not inconceivable	leave data unchanged
tbx	Kapin	probably anomalous	unlikely decrease from 1979 to 1980 datapoints	disregard E92
tcc	Datooga	probably anomalous	trendline highly variable; E05 notes lg use vigorous, so overall upward trendline seems most likely possibility	disregard E92 and E00 as probable overestimations
tdg	Tamang, Western	probably anomalous	unlikely increase from 1989 to 1991 datapoints	disregard E92

thr	Tharu, Rana	probably anomalous	unlikely increase from 1985 to 2000 datapoints; debatable increase from 2000 to 2005 datapoints	disregard E92
thu	Thuri	probably anomalous	unlikely increase from 1966 to 1971 datapoints	disregard E78 as outlier
thv	Tamahaq, Tahaggart	probably anomalous	unlikely decrease from 1976 to 1987 datapoints; 1987 datapoint is low-end estimate (probably for Algeria only) and out of line with subsequent estimates	disregard E78 and E88; use 2-datapoint trendline: E96 and E00
tic	Tira	probably anomalous	unlikely increase from 1977 to 1982 datapoints; 1977 appears to be repetition (from a different source) of 1966 datapoint)	disregard E78
tou	Tho	probably anomalous	unlikely increase from 1996 to 1999 datapoints	disregard E00
tpi	Tok Pisin	probably anomalous	unlikely increase from 1977 to 1982 datapoints	disregard E78
tqu	Touo	probably anomalous	unlikely increase from 1976 to 1981 datapoints; debatable increase from 1998 to 1999 datapoints	disregard E78
trr	Taushiro	may not be anomalous	trendline looks plausible (very small language)	leave data unchanged
tsg	Tausug	may not be anomalous	increases are large, but not totally implausible	leave data unchanged
tsi	Tsimshian	possibly anomalous	trendline debatable but not implausible	leave data unchanged
tsr	Akei	possibly anomalous	trendline debatable but not implausible	leave data unchanged
tud	Tuxá	may not be anomalous	trendline looks plausible (very small language)	leave data unchanged
tzc	Tzotzil, Chamula	probably anomalous	non-census datapoints make trendline highly variable	disregard E69, E88, and E92; construct 2-datapoint trendline from E78 and E00 (both based on census data)
urd	Urdu	probably anomalous	unlikely increase from 1969 to 1971 datapoints	disregard E69
waz	Wampur	probably anomalous	unlikely decrease from 1969 to 1970 datapoints	disregard E69 as outlier
wic	Wichita	may not be anomalous	trendline looks plausible (very small language)	leave data unchanged

wir	Wiraféd	may not be anomalous	trendline looks plausible (very small language)	leave data unchanged
wll	Wali (Sudan)	probably anomalous	unlikely increase from 1977 to 1978 datapoints	disregard E92
xrw	Karawa	possibly anomalous	trendline highly variable, but E05 comments suggest that it may be plausible	leave data unchanged
xsy	Saisiyat	possibly anomalous	debatable increase from 1969 to 1973 datapoints	disregard E69
yee	Yimas	possibly anomalous	debatable decrease from 1977 to 1981 datapoints, but possible	leave data unchanged
yig	Yi, Guizhou	possibly anomalous	E00 datapoint appears to supersede E92	disregard E92
ykm	Yakamul	possibly anomalous	debatable increase from 1978 to 1981 datapoints	leave data unchanged
yra	Yerakai	possibly anomalous	debatable decrease from 1969 to 1971 datapoints	leave data unchanged
yuy	Yugur, East	probably anomalous	unlikely decrease from 1990 to 1991 datapoints	disregard E92
ywt	Yi, Western	probably anomalous	unlikely decrease from 1990 to 1991 datapoints	disregard E92
zat	Zapotec, Tabaa	possibly anomalous	trendline highly variable	disregard E88 as outlier
zav	Zapotec, Yatzachi	probably anomalous	unlikely increase from 1969 to 1977 datapoints	disregard E69
zeg	Zenag	possibly anomalous	debatable decrease from 1979 to 1980 datapoints	disregard E92
zen	Zenaga	probably anomalous	unlikely decrease from 1992 to 1998 datapoints	disregard E78 and E92
zkp	Kaingáng, São Paulo	may not be anomalous	trendline looks plausible (very small language)	leave data unchanged
zro	Záparo	may not be anomalous	trendline looks plausible (very small language)	leave data unchanged

7. REMOVAL OF DISCREPANCIES IN ESTIMATES FROM THE SAME YEAR. In a few instances, different editions of *Ethnologue* report different estimates for a language, but attribute them to the same year. As an example, for Guerrero Nahuatl [ngu], the 1978 edition gives high and low figures of 180,000 and 160,000, respectively (average = 170,000), attributing the estimates to SIL 1977. However, the 1988 edition gives figures of 90,000 and 80,000 (average = 85,000) and also attributes the estimates to SIL 1977. In such instances we used the most recent estimate in the calculation and dropped the older.

8. TREATMENT OF EXTINCT LANGUAGES WITHIN TIME SERIES. Languages that are extinct as mother tongues are shown as zero values in the database of numbers of speakers. A time series of zeros would imply no overall trend in the status of that language and therefore those languages were taken out of the sample from the year after that in which they were recorded as going extinct. Eleven languages in our sample were reported as having zero speakers before the end of a non-zero time series, i.e., subsequent editions of *Ethnologue* reported one or more speaker. This may have occurred because a language was believed to have gone extinct, but later found to be still in use among a small community of speakers. In such cases the zero values were removed from the ILD. In instances where the zero value was the first of only two datapoints, leaving only a single datapoint, then both were removed from the ILD.

9. TREATMENT OF SPLITS AND MERGERS. If pre-2005 editions of *Ethnologue* considered a language in our sample as a dialect of a larger language, we excluded any datapoints for that larger language from the database. However, if these editions gave separate speaker totals for these putative dialects, those totals were included as datapoints. Conversely, if pre-2005 editions of *Ethnologue* considered a language in our sample as comprising two or more distinct languages, and if separate speaker totals were available for those putatively distinct languages, we aggregated the totals into a single datapoint.

10. SPECIAL CHALLENGES OF *ETHNOLOGUE* DATA ANALYSIS. Any retrospective analysis of language demographic data presents certain challenges. Next, we discuss three that are particular to *Ethnologue*.

Changes in the number of languages. Since its inception in 1951, each new edition of *Ethnologue* has reported a higher number of languages. In the earliest editions, some of this increment could be explained by the addition of "new" languages of which Western linguistic science was previously unaware. In more recent editions, the editors explain that the increment is, except for a very few cases, no longer due to such "discoveries" but rather to dialects of single languages being reclassified as separate languages. At the same time, as part of the editorial process of preparing each new edition, a number of entries are dropped from the roster. Prior to the advent of ISO codes and roster-change documentation, it may be presumed that most often such entries were expunged because the speech form in question was redefined as a dialect, with its old entry being merged with that of the parent language in the new edition. (See above for our treatment of such cases.) Additionally, in some instances entries appear to have been dropped because they were determined to simply be alternative names for another language already on the roster, or because they were names for an entire ethnic group, not a language.

The 15th edition of *Ethnologue* listed 7,299 languages for the world. The sifting of speech forms to determine whether or not they should be considered discrete languages will no doubt continue, but, significantly, the 16th edition lists 7,296 languages—essentially the same number as the previous edition. As M. Paul Lewis writes, "the rate of languages being split off from existing ones and previously separate languages being re-classified as a single larger language is about equal. In the 16th edition, the count of living languages has diminished (by 3) for the first time. This is largely the result of mergers of existing languages, though we raise the possibility (in the Introduction) that it could also be the result of our having re-classified a good number of 'Nearly Extinct' languages as 'Extinct'..." (Lewis, pers. comm., 25 May 2009).

Evidence of moribundity and vigor. Two important components of the ILD database are the fields that record *Ethnologue's* qualitative assessments of whether use of a particular language is moribund or vigorous. *Ethnologue* uses a number of standard locutions to indicate that a language has or may become moribund (i.e., it is no longer being learned by young people). *Ethnologue* descriptions were considered to be indicative of moribundity if they point to a decline in the use of the language by, or its disfavor among, young people, or if they make some general reference to language loss (e.g., "It is reported that the language appears to be dying out"). In such cases the Moribundity checkbox was checked.

Ethnologue also has a number of standard locutions that indicate that a language is vigorous. The most common is the simple notation "Language use vigorous." *Ethnologue* descriptions were considered to be indicative of vigor if they point to robustness in the use of the language, its acceptance by young people, its being taught in school, there being a language revitalization program in place, etc., or if they make some general reference to language vigor (e.g., "The people have a positive attitude toward the language"). In such cases the Vigor checkbox was checked.

It is important to understand that these characterizations of language moribundity and vigor are descriptive, not diagnostic. Not infrequently, a language entry may contain evidence of both moribundity and vigor. Examples include cases where indigenous languages are still suffering declines in acceptance/use by young people, but for which language revitalization efforts have begun.

Is the language spoken primarily/entirely by indigenous people? This question, which is an important concern of the ILD, is problematic because there is no standard list of indigenous peoples/languages. In fact, as the United Nations' State of the World's Indigenous Peoples concludes, there is no standard definition of "indigenous peoples," no definition of the term has ever been adopted by a U.N.-system body, and indigenous peoples themselves have "rejected the idea of a formal definition of indigenous peoples at the international level to be adopted by states. Similarly, government delegations expressed the view that it was neither desirable nor necessary to elaborate a universal definition of indigenous peoples" (UNDESA 2009:4–5).

Absent definitive guidance, we used our experience with the dataset and knowledge of the ethnographic literature to determine which languages to check as "indigenous." We used the definition given in the International Labour Organisation's Convention 169 on Indigenous and Tribal Peoples (1989) as a general guide to which groups should be considered indigenous:

> (a) tribal peoples in independent countries whose social, cultural and economic conditions distinguish them from other sections of the national community, and whose status is regulated wholly or partially by their own customs or traditions or by special laws or regulations;

> (b) peoples in independent countries who are regarded as indigenous on account of their descent from the populations which inhabited the country, or a geographical region to which the country belongs, at the time of conquest or colonisation or the establishment of present state boundaries and who, irrespective of their legal status, retain some or all of their own social, economic, cultural and political institutions.

The process was straightforward for the Americas and Europe, where our knowledge of the ethnographic literature made identification of indigenous groups relatively simple. For example, the only European languages in our sample that we marked as "indigenous" were Ume Sami [sju], spoken in Sweden, and a handful of North Caucasian and Altaic languages whose *Ethnologue* "Main Region" is European Russia. Some minority languages in our sample, such as Welsh [gym], although they might also be considered indigenous, were not so marked because it would have meant the inclusion of most European languages. In Africa, although some have questioned the application of "indigenous" to any of the continent's languages (UNDESA 2009:6), we feel the situation is fairly clear-cut: almost every small language group in sub-Saharan Africa was marked as indigenous (the exceptions being sign languages and creoles). In Australia, it is easy to identify the Aboriginal/Torres Strait Islander languages, and in the Pacific (including Papua New Guinea) it is also obvious which languages are indigenous.

In Asia, our relative lack of ethnographic knowledge made the process more difficult (and again, some have challenged the use of the concept in at least some areas of Asia). In India, for example, we generally marked as "indigenous" only those languages that *Ethnologue* listed as being spoken by a Scheduled Tribe, or those that are well-known to be indigenous (such as Andamanese languages). In China, we marked those languages spoken by groups listed as official minority nationalities; in Japan, the Ryukyuan languages; in Taiwan, the small non-Chinese languages; and so forth. Virtually all languages in Indonesia with fewer than 10,000 speakers were so marked.

Admittedly, this process is ad hoc and inevitably we will have made mistakes. However, given the strong interest on the part of various international bodies in the status of indigenous languages globally (e.g., UNPFII, the Convention on Biological Diversity), we feel it is important to make a start at identifying them.

In the final analysis, of the 1,500 languages in our sample we considered 1,285, or 85.6%, to be indigenous. This estimate is supported by the fact most of the world's languages are endemic to a single country (i.e., spoken there and nowhere else). In our ILD sample, 1,453 out of the 1,500 languages had a speaker-number estimate; of these, 1,187, or 81.6%, were endemic. This is very close to the results of Harmon's earlier study of the 1992 edition of *Ethnologue*, in which he found that 83.3% of the world's languages are endemic (Harmon 1995:10). It seems logical to assume that there is a very large overlap, probably on the order of 95%, between indigenous and endemic languages. If that assumption is correct, and if we conservatively posit our figure of 85.6% to represent a high-end estimate of the proportion of indigenous languages in the world, then we can derive a low estimate of 81% by multiplying 85.6 by the 95% overlap. Rounding off this range, we therefore believe it reasonable to estimate that 80–85% of the world's languages are spoken by indigenous people.

Appendix B. Calculating the Index of Diversity

MEASURING THE GEOMETRIC MEAN SHARE OF A POPULATION IN TERMS OF NUMBERS OF SPEAKERS OF LANGUAGES. The ILD uses language evenness in conjunction with language richness as a proxy for linguistic diversity. Because the goal of the index is to measure trends in linguistic diversity, it must account for changes in richness and evenness: that is, changes in the relative distribution of mother-tongue speakers among discrete languages within the total population, as measured from the starting point of the index to its ending point. For any given grouping of languages at a particular starting point in time—call it Time 0—the way we measure their relative distribution is to calculate each one's share of the total population of the grouping and then find the average of those shares; this average share becomes the numerical benchmark for relative distribution at Time 0. We then move to a subsequent point in time—Time 1—and redo the calculations. This yields a new average share. We then compare the change in average share from Time 0 to Time 1, thus producing a trendline of changes in the relative distribution of the languages in that grouping.

As an example, consider languages grouped at the global level. In any given year, each language in the world holds a particular share of the global population: languages with a large number of mother-tongue speakers have greater shares, while languages with a smaller number have lesser ones. With each passing year the shares held by individual languages change—and thus the average share changes—because (1) the world's languages are growing at different rates and (2) speakers are shifting between languages. Tracking those changes in average share across the years produces a trend in the distribution of the world's languages, and the simplest way to show the trend graphically is by depicting changes in the average share as a single line that goes either up or down from one year to the next. That is what the ILD Global trendline does.

It is important to specify what we mean by "average," because in mathematics there are actually several kinds of averages, some of which are more appropriate for analyzing certain sets of numbers than others. When most people use the word "average," they usually mean a simple calculation in which one adds a set of numbers and then divides by a count of numbers in the set; thus, the average of 2 and 8 is 5 (2 + 8 = 10 divided by 2). Technically, this calculation is called the arithmetic mean, and it works fine for simple sets of numbers. But another kind of average, the geometric mean, is more appropriate for analyzing data sets with skewed distributions such as the size distribution of languages. Wikipedia happens to have a very clear explanation:

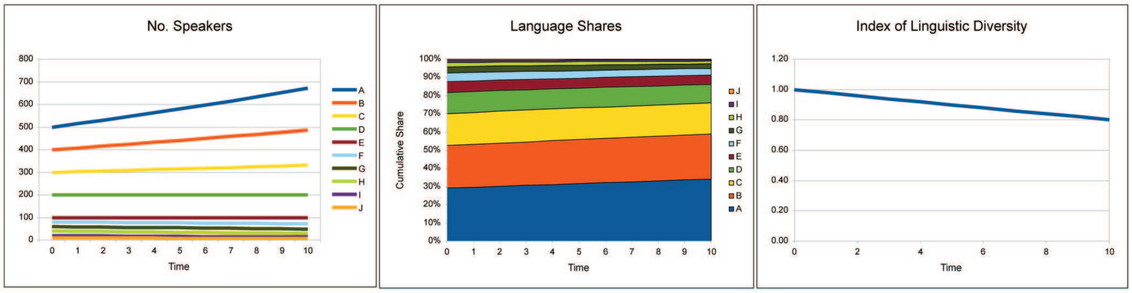
> The geometric mean, in mathematics, is a type of mean or average, which indicates the central tendency or typical value of a set of numbers. It is similar to the arithmetic mean, which is what most people think of with the word "average," except that instead of adding the set of numbers and then dividing the sum by the count of numbers in the set, n, the numbers are multiplied and then the nth root of the resulting product is taken.... The geometric mean ... is ... often used for a set of numbers whose values are meant to be multiplied together or are *exponential in nature, such as data on the growth of the human population* or interest rates of a financial investment. (emphasis added; http://en.wikipedia.org/wiki/Geometric_mean, accessed March 2010)

In our simple example above, the arithmetic mean was 5, but the geometric mean is 4 (2 x 8 = 16, and thence the square root of 16, since there are 2 numbers in the set; if there had been 3 numbers one would take the cube root, etc.).

When we use the term "average" with respect to the ILD, we refer to the geometric mean, not the arithmetic mean. The reason we use the geometric mean is precisely because of the consideration that we have highlighted in the quoted definition above: we are analyzing language data in a world where the numbers of speakers are unevenly distributed among languages: more than 94% of the world's people speak one of the 389 largest languages, each of which has more than a million speakers, while the other 6,520 non-extinct languages account for the fewer than 5% of the world's population (http://www.ethnologue.com/ethno_docs/distribution.asp?by=size, accessed March 2010; Lewis 2009 includes the same analysis but with slightly differing figures). If we were to construct an index using the conventional notion of "average"—i.e., by calculating trends in the arithmetic mean—we would be unable to accurately reflect shifts in evenness because the arithmetic mean would not give a meaningful measure of the extremely skewed distributions. Indeed, the arithmetic mean share of the world's population is constant over time for any distribution of speaker numbers as long as they remain greater than zero.

Figures B-1 provides an example that illustrates this, and also shows how the ILD is calculated. We have set up the example as a simplified model of the real world where there are a certain number of languages with different numbers of mother-tongue speakers. In our example the world consists of 10 languages, A through J, each having a different number of speakers at Year 0, the starting point of the index calculation. We then set a different growth rate for each language—just as in the real world each language is growing (or declining) at a different rate—through Year 10, the endpoint of the index. In Figure B-1, we made Languages A, B, and C grow; Languages D and E stay the same; and Language F, G, H, I, and J decline. And, just like in the real world, the overall population is growing, going from 1,710 in Year 0 to 1,962 in Year 10.

FIGURE B-1



Language	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Annual Rate**	Percent change
A	500	515	530	546	563	580	597	615	633	652	672	1.03	3%
B	400	408	416	424	433	442	450	459	469	478	488	1.02	2%
C	300	303	306	309	312	315	318	322	325	328	331	1.01	1%
D	200	200	200	200	200	200	200	200	200	200	200	1.00	0%
E	100	100	100	100	100	100	100	100	100	100	100	1.00	0%
F	80	79	78	78	77	76	75	75	74	73	72	0.99	-1%
G	60	59	58	56	55	54	53	52	51	50	49	0.98	-2%
H	40	39	38	37	35	34	33	32	31	30	29	0.97	-3%
I	20	19	18	18	17	16	16	15	14	14	13	0.96	-4%
J	10	10	9	9	8	8	8	8	7	7	7	0.96	-4%
Total (N)	**1710**	**1732**	**1754**	**1777**	**1801**	**1826**	**1851**	**1878**	**1905**	**1933**	**1962**		
pA	0.29	0.30	0.30	0.31	0.31	0.32	0.32	0.33	0.33	0.34	0.34		
pB	0.23	0.24	0.24	0.24	0.24	0.24	0.24	0.24	0.25	0.25	0.25		
pC	0.18	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17		
pD	0.12	0.12	0.11	0.11	0.11	0.11	0.11	0.11	0.10	0.10	0.10		
pE	0.06	0.06	0.06	0.06	0.06	0.05	0.05	0.05	0.05	0.05	0.05		
pF	0.05	0.05	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04		
pG	0.04	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.02		
pH	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02		
pI	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01		
pJ	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		
Arithmean	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10		
Geomean	**0.054**	**0.053**	**0.052**	**0.051**	**0.050**	**0.049**	**0.048**	**0.047**	**0.045**	**0.044**	**0.043**		
ILD	**1.00**	**0.98**	**0.96**	**0.94**	**0.92**	**0.90**	**0.88**	**0.86**	**0.84**	**0.82**	**0.80**		

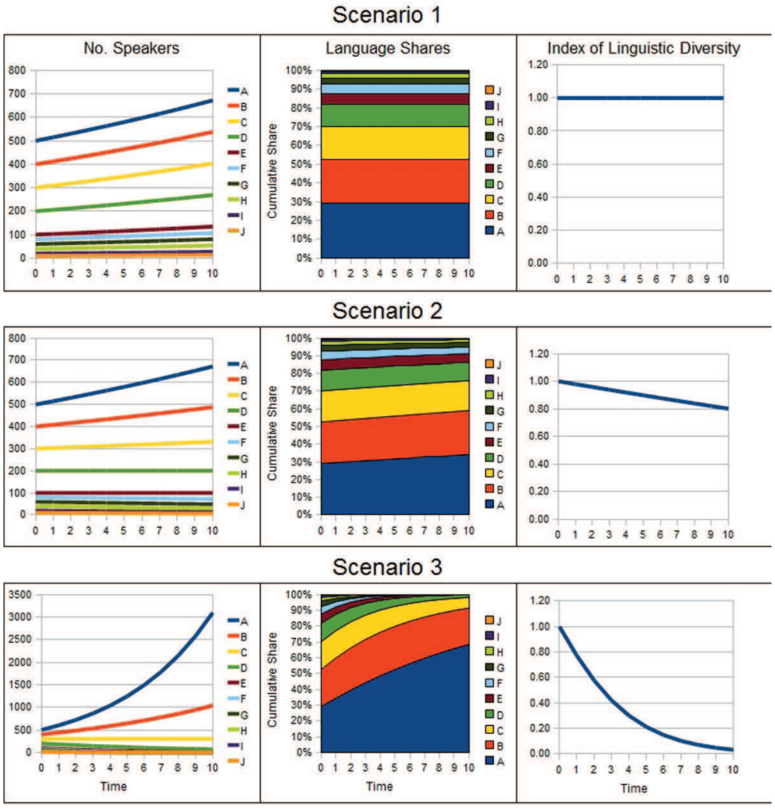
However, when we look at the geometric mean we see a different story: the average is decreasing, going from 0.54 in Year 0 to 0.43 in Year 10. This is because the geometric mean is not just indicating a raw average, as the arithmetic mean does; rather, it indicates the average share of the global population held by each language in a world where the size distribution of languages is highly skewed and languages are growing at different rates (positive or negative). That is why we use the geometric mean to measure the situation of languages in the real world. And in this example, the geometric mean correctly indicates that a loss of distributional diversity is occurring.

The ILD simply compares changes in the geometric mean—the average share—over time by dividing the geometric mean at the endpoint of the index by that at the starting point. Here, the calculation is 0.43 divided by 0.54 = 0.80. So in Figure B-1 the ILD declines from 1.00 to 0.80—just as ILD Global did in the real world over the period 1970–2005.

This was by design, of course: we set up the growth rates such that Figure B-1 would approximate the situation in the real world. This gives us a point of comparison to which we can add two other scenarios that show how the ILD behaves under different extremes.

In Figure B-2 we have taken the graphs from Figure B-1 and flanked them by similar graphs that illustrate these additional scenarios. On the top row (Scenario 1), we begin with the same simplified world as before: the same ten languages with the same starting populations as in Figure B-1. But this time all the languages grow at the same rate—that is, each one holds its share of the population. This is the hypothetical situation of stability that produces a perfectly flat ILD trendline, as can be seen from the upper righthand graph. On the bottom row (Scenario 3), we have the same starting conditions, but this time the annual growth rates reflect a sharp decline in most of the ten languages. Here, we see that the ILD trendline declines sharply, reflecting the steep loss of diversity under this scenario. If one compares the three middle graphs (Language Shares), we see that as diversity declines as we go from the top to the bottom row, the area of the graph taken up by the largest languages shifts or begins to "bulge," with the area taken up by the smaller languages being "squeezed out." This is a graphical depiction of shifts in the distribution (or concentration) of the world's speakers.

FIGURE B-2



CALCULATING THE ILD IN A REAL WORLD OF MISSING DATAPOINTS. The calculation of the ILD works in three steps. The description below differs from the simplified version given in the main text in that one does not need to know the number of speakers of every language in every year in order to calculate the index. It allows for missing datapoints by interpolating between datapoints, assuming a constant annual rate of growth (or decline). This is the simplest assumption one can make in the absence of data. However, no datapoints were extrapolated using this method, as the assumption of a constant annual rate of change beyond the first and last data years is not always reasonable. Therefore missing datapoints remained prior to the first data year and after the last data year for each language. To allow for this, the index was calculated by finding the average change in share from one year to the next for all languages for which actual or interpolated datapoints existed, and then chaining together the average changes for each year into an index starting at one in the baseline year. This method is adapted from that of the Living Planet Index (Loh et al. 2005).

1. The fraction F of the total population (global or regional) represented by each datapoint (a datapoint means N speakers of language *l* in year *y*) was calculated.

$$F_{ly} = (N_{ly}+1)/P_y$$

where N_{ly} is the number of speakers of language *l* in year *y*, and
P_y is the total population in year *y*.

To avoid taking the log of zero or dividing by zero in step 2, each *N* value was increased by 1. The total populations from 1950 to 2005 of the world and five regions—Africa, Asia, Pacific, Europe and the Americas—were taken from UN Population Division (2006 revision). Downloaded from http://esa.un.org/unpp/index.asp.

Missing annual values between consecutive N_{ly} values were interpolated. This was done by assuming a constant annual rate of change between two datapoints. The intermediate values were calculated using a simple log-linear interpolation.

$$N_i = N_p(N_q/N_p)^{(i-p)/(q-p)}$$

where *i* = year of intermediate datapoints,
p = year of the preceding datapoint, and
q = year of the subsequent datapoint.

For example if $N_{1980} = 1000$ and $N_{2000} = 100$,
then $N_{1990} = 1000 \text{ x } (100/1000)^{(10/20)} = 1000 \text{ x } 0.1^{1/2} = 316$

2. The geometric mean of the ratio of fraction of speakers from one year to the next across all languages in the sample was calculated. This was done by log-transforming the ratio of consecutive F values such that:

$$d_y = log_{10}(F_{ly}/F_{ly-1})$$

F_{ly} = fraction of population speaking language *l* at year *y*,
F_{ly-1} = fraction of population speaking language *l* the preceding year.

The mean d value for all languages with data in a single year was then calculated

$$\bar{d}_y = \frac{1}{n_y}\sum_{i=1}^{n_y} d_{iy}$$

where

n_y = number of languages with some value (actual or interpolated) for F in the year y (not all languages in the sample have data for every year of the index because the earliest datapoint may be after 1970 or the most recent before 2005, and no values were extrapolated).

3. Finally, the geometric means in each year were antilogged and chained together to form an index, such that

$$I_y = I_{y-1}\, 10^{dy}$$

and the index value in 1970 was set to unity.

$$I_{1970} = 1.0$$

where I_y = the Index of Linguistic Diversity in year y.

In this way, the ILD shows the trend in the fraction of the total population that speaks a language that is average or typical of all languages in the sample. Note that the interpolation was not linear but log-linear, and that the average change in numbers across all languages was taken as the geometric mean and not the arithmetic mean. This means that increases and decreases in the ILD are equivalent to each other for the purpose of calculating the index. For instance, using log-linear interpolation and log-transforming all the data in this way, a doubling of the fraction of a population speaking language A between 1970 and 2005 would be cancelled out by a halving of the fraction of the population speaking language B over the same period. This is because doubling means multiplying by 2, whereas halving represents multiplying by 0.5. The arithmetic mean of 2 and 0.5 is 1.25, whereas the geometric mean is 1.

References

Adelaar, Willem. 2007. Latin America. In Moseley 2007a, 97–100.

Adelaar, Willem & J. Diego Quesada. 2007. Meso-America. In Moseley 2007a, 197–209.

Black, Paul. 1983. Aboriginal languages of the Northern Territory. Darwin, NT, Australia: School of Australian Linguistics, Darwin Community College.

Bradley, David. 2007. East and Southeast Asia. In Moseley 2007a, 349–422.

Canonge, Elliott & Dick Pittman. N.d. [ca. 1958]. *The fifth edition of the Ethnologue of Bibleless tribes for prayer intercessors, Bible translators, missionaries, prospective missionaries, mission councils*. Glendale, CA: Wycliffe Bible Translators.

Crevels, Mily. 2007. South America. In Moseley 2007a, 103–194.

Dimmendaal, Gerrit J. & F. K. Erhard Voeltz. 2007. Africa. In Moseley 2007a, 579–634.

Dorian, Nancy C. 1981. *Language death: The life cycle of a Scottish Gaelic dialect*. Philadelphia: University of Pennsylvania Press.

Evans, Nicholas. 2001. The last speaker is dead—long live the last speaker! In Paul Newman and Martha Ratcliff (eds.). *Linguistic fieldwork*, 250–281. Cambridge: Cambridge University Press.

Evans, Nicholas. 2010. *Dying words: Endangered languages and what they have to tell us*. Chichester, UK: Wiley-Blackwell.

Garza Cuarón, Beatriz & Yolanda Lastra. 1991. Endangered languages in Mexico. In Robert H. Robins and Eugenius Uhlenbeck (eds.), *Endangered languages,* 93–134. Oxford and New York: Berg.

Golla, Victor. 2007. North America. In Moseley 2007a, 1–95.

Gordon, Raymond G., Jr. (ed.). 2005. *Ethnologue: Languages of the world*. 15th edn. Dallas: SIL International.

Grimes, Barbara F. (ed.). 1974. *Ethnologue*. 8th edn. Huntington Beach, CA: Wycliffe Bible Translators.

Grimes, Barbara F. (ed.). 1978. *Ethnologue*. 9th edn. Huntington Beach, CA: Wycliffe Bible Translators.

Grimes, Barbara F. (ed.). 1984. *Ethnologue: Languages of the world*. 10th edn. Dallas: Wycliffe Bible Translators.

Grimes, Barbara F. (ed.). 1988. *Ethnologue: Languages of the world*. 11th edn. Dallas: Summer Institute of Linguistics.

Grimes, Barbara F. (ed.). 1992a. *Ethnologue: Languages of the world*. Vol. 1. 12th edn. Dallas: Summer Institute of Linguistics.

Grimes, Barbara F. (ed.). 1992b. *Ethnologue index*. Vol. 2. 12th edn. Dallas: Summer Institute of Linguistics.

Grimes, Barbara F. (ed.). 1996a. *Ethnologue: Languages of the world*. 13th edn. Dallas: Summer Institute of Linguistics.

Grimes, Barbara F. (ed.). 1996b. *Ethnologue language family index*. 13th edn. Dallas: Summer Institute of Linguistics.

Grimes, Barbara F. (ed.). 1996c. *Ethnologue name index*. 13th edn. Dallas: Summer Institute of Linguistics.

Grimes, Barbara F. (ed.). 2000a. *Ethnologue, Volume 1: Languages of the world*. 14th ed. Dallas: SIL International.

Grimes, Barbara F. (ed.). 2000b. *Ethnologue, Volume 2: Maps and indexes*. 14th ed. Dallas: SIL International.

Gunnemark, Erik & Donald Kenrick. 1985. *A geolinguistic handbook*. Gothenburg, Sweden: The authors.

Gunnemark, Erik & Donald Kenrick. 1983. *What language do they speak? A geolinguistic handbook covering languages, countries and peoples in the whole world*. Gothenburg, Sweden: The authors.

Harmon, David. 1995. The status of the world's languages according to *Ethnologue*. *Southwest Journal of Linguistics* 14:1/2. 1–28.

Harmon, David. 2002. *In light of our differences: How diversity in nature and culture makes us human*. Washington, DC: Smithsonian Institution Press.

Haspelmath, Martin, Matthew S. Dryer, David Gil, and Bernard Comrie (eds.). 2005. *The world atlas of language structures*. Oxford: Oxford University Press.

ILO [International Labour Organisation]. 1989. *C169 Indigenous and Tribal Peoples Convention, 1989*. http://www.ilo.org/ilolex/english/convdisp1.htm.

Krauss, Michael. 1992. The world's languages in crisis. *Language* 68. 4–10.

Krauss, Michael. 2006. Classification and terminology for degrees of the language endangerment. In Matthias Brenzinger (ed.), *Language diversity endangered,* 1–8. Berlin: Mouton de Gruyter.

Lewis, M. Paul (ed.). 2009. *Ethnologue: Languages of the world*. 16th edn. Dallas: SIL International.

Loh, Jonathan, R.E. Green, T. Ricketts, J. Lamoreaux, M. Jenkins, V. Kapos & J. Randers. 2005. The Living Planet Index: Using species population time series to track trends in biodiversity. *Philosophical Transactions of the Royal Society of London B* 360. 289–295.

Maffi, Luisa. 2005. Linguistic, cultural, and biological diversity. *Annual Review of Anthropology* 29. 599–617.

Meillet, A[ntoine] & M[arcel] Cohen. 1952. *Les langues du monde*. Revised edn. Paris: Centre National de la Recherche Scientifique / H. Champion.

Meillet, A[ntoine] & M[arcel] Cohen. 1924. *Les langues du monde*. Collection Linguistique publiée par La Société de Linguistique de Paris 16. Paris: Librairie Ancienne Édouard Champion.

Moseley, Christopher (ed.). 2007a. *Encyclopedia of the world's endangered languages*. London and New York: Routledge.

Moseley, Christopher. 2007b. General introduction. In Moseley 2007a, vii–xvi.

Nettle, Daniel. 1999. *Linguistic diversity*. Oxford: Oxford University Press.

Pittman, R[ichard] S. (ed.). 1965. *Ethnologue*. 6th edn. Santa Ana, CA: Wycliffe Bible Translators.

Pittman, Richard S. (ed.). 1970. *Ethnologue*. 7th edn. Santa Ana, CA: Wycliffe Bible Translators.

Salminen, Tapani. 2007. Europe and North Asia. In Moseley 2007a, 211–280.

Skutnabb-Kangas, Tove. 2000. *Linguistic genocide in education—or worldwide diversity and human rights?* Mahwah, NJ: Lawrence Erlbaum Associates.

Tesnière, L. 1928. *Statistique des langues de l'Europe*. Appendix in A[ntoine] Meillet, *Les langues dans l'Europe nouvelle*, 291–484. Paris: Payot.

UNDESA [United Nations Department of Economic and Social Affairs]. 2009. *State of the world's indigenous peoples*. New York: United Nations.

UNESCO [United Nations Educational, Scientific, and Cultural Organisation]. 2009a. A methodology for assessing language vitality and endangerment. http://www.unesco.org/culture/ich/index.php?pg=00142.

UNESCO. 2009b. *UNESCO interactive atlas of the world's languages in danger*. http://www.unesco.org/culture/ich/index.php?pg=00206. (Accessed February 2010.)

UNPFII [United Nations Permanent Forum on Indigenous Issues]. 2008. *International Expert Group Meeting on Indigenous Languages*. http://www.un.org/esa/socdev/unpfii/en/EGM_IL.html. (Accessed February 2010.)

van Driem, George. 2007. South Asia and the Middle East. In Moseley 2007a, 283–347.

Voegelin, C. F. & F. M. Voegelin. 1977. *Classification and index of the world's languages*. New York: Elsevier.

WBT [Wycliffe Bible Translators]. 1951. *Missionary ethnologue for intercessors, translators, missionaries, and mission councils*. 1st edn. Berwick, Victoria, Australia: Wycliffe School of Linguistics. (9 pp., mimeograph.)

WBT [Wycliffe Bible Translators]. 1952. *Translator's ethnologue for intercessors, translators, missionaries, and mission councils*. 2nd edn. Grand Forks, ND, and Glendale, CA: Wycliffe Bible Translators. (25 pp., mimeograph.)

WBT [Wycliffe Bible Translators]. 1953. *Missionary ethnologue for intercessors, translators, missionaries, and mission councils*. Revised and extended edn. of 1st (1951) edn. Berwick, Victoria, Australia: Wycliffe School of Linguistics. (3 pp., mimeograph.)

WBT [Wycliffe Bible Translators]. N.d. [ca. 1953]. *Translator's ethnologue for intercessors, translators, missionaries, and mission councils*. 4th edn. Norman, OK, and Glendale, CA: Wycliffe Bible Translators. (28 pp., mimeograph)

Whalen, D.H. & Gary F. Simons. 2009. Endangered language families. Paper presented at the 1st International Conference on Language Documentation and Conservation, University of Hawai´i, 12–14 March.

Winter, J.C. 1981. Die Khoisan-Famille. In Bernd Heine, Thilo C. Schadeberg, and Ekkehard Wolff, eds., *Die Sprachen Afrikas: Ein Handbuch*, 329-374. Hamburg: Helmut Buske Verlag.

Wurm, Stephen A. 2007. Australasia and the Pacific. In Moseley 2007a, 425–577.

Wurm, S. A. & Shirô Hattori (eds.) 1981. Language atlas of the Pacific area. Part 1, New Guinea area, Oceania, Australia. Canberra: Pacific Linguistics C-66.

WWF [Worldwide Fund for Nature] 2008. *Living Planet Report 2008*. Gland, Switzerland; WWF.

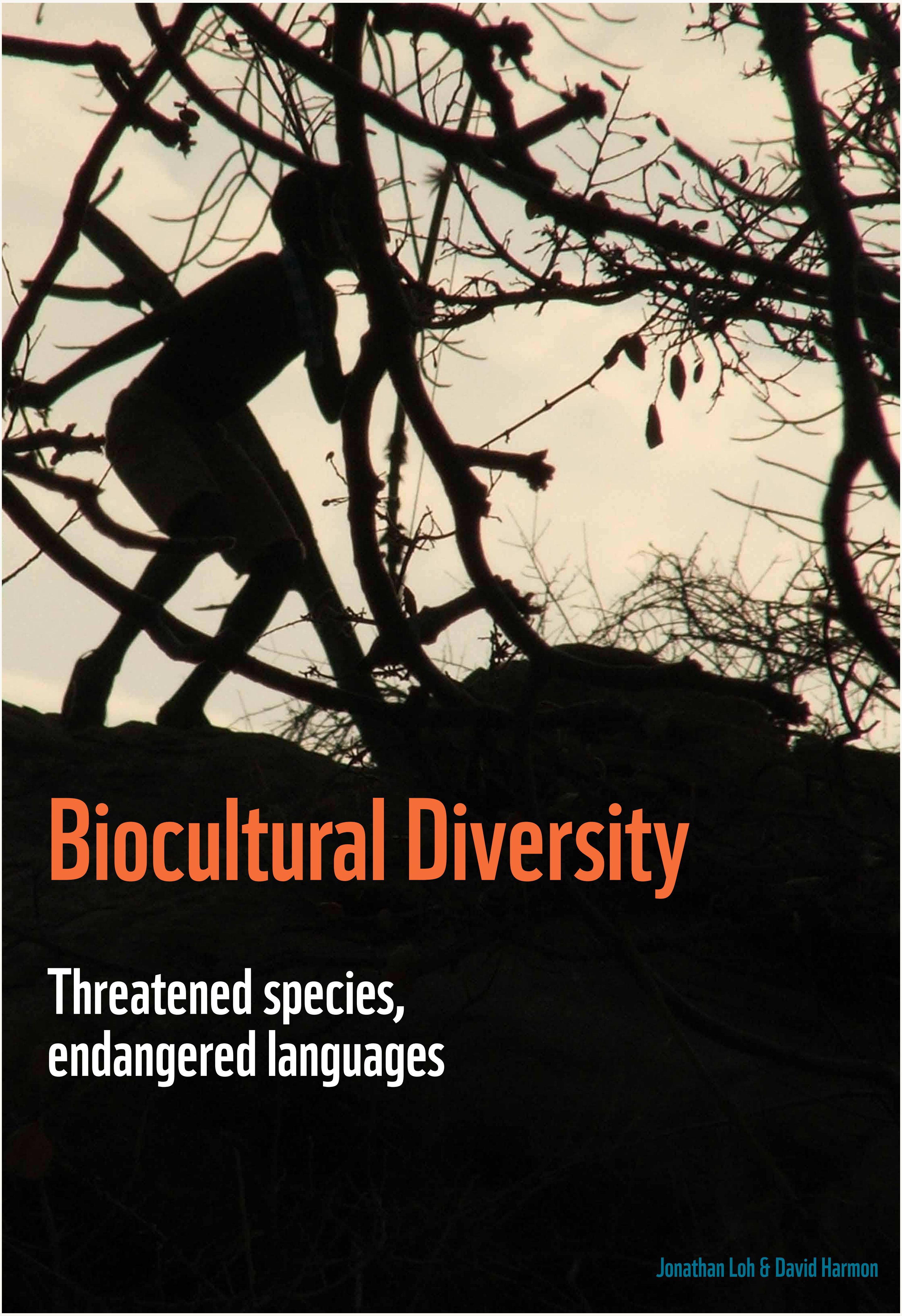
David Harmon
dharmon@georgewright.org

Jonathan Loh
jonathan@livingplanet.org.uk

Biocultural Diversity: Threatened species, endangered languages

Jonathan Loh and David Harmon

WWF Netherlands, Zeist, The Netherlands
(2014)



Biocultural Diversity

Threatened species, endangered languages

Jonathan Loh & David Harmon

This research and report has been commissioned by WWF Netherlands

About the Authors

Jonathan Loh is a biologist specializing in monitoring, assessment and indicators of biological and cultural diversity. He works for several conservation NGOs and is an Honorary Research Associate of the Zoological Society of London.

David Harmon is executive director of the George Wright Society, a professional association of people who work in parks, protected areas, and cultural sites. He also maintains an active research interest in the relationship between biological and cultural diversity, having co-founded the NGO Terralingua, which is devoted to that subject.

Acknowledgements

We would like to thank Johan van de Gronden, Monique Grooten, Natasja Oerlemans and Natascha Zwaal of WWF-Netherlands for their enthusiasm, support and comments on earlier drafts of this report. We are also grateful to Terralingua and The Christensen Fund for supporting precursors to the work presented here. We should point out that the opinions expressed in this report are those of the authors and not necessarily those of WWF or any other of the organizations we work with.

WWF

WWF is one of the world's largest and most experienced independent conservation organizations, with over 5 million supporters and a global network active in more than 100 countries.

WWF's mission is to stop the degradation of the planet's natural environment and to build a future in which humans live in harmony with nature, by conserving the world's biological diversity, ensuring that the use of renewable natural resources is sustainable, and promoting the reduction of pollution and wasteful consumption.

Reproduction

Citation

Loh, J. & D. Harmon. 2014. Biocultural Diversity: threatened species, endangered languages. WWF Netherlands, Zeist, The Netherlands.

Design

peer&dedigitalesupermarkt (www.pdds.nl) and Jonathan Loh

Front cover photograph: Hadza boy, speaker of a Khoi-san language, hunting. Fewer than 1,000 speakers of Hadza remain. Lake Eyasi, Tanzania © Jonathan Loh

ISBN 978-90-74595-00-1

BIOCULTURAL DIVERSITY: THREATENED SPECIES, ENDANGERED LANGUAGES

BIOCULTURAL DIVERSITY

Nature and culture as dual aspects of a single entity

The Oxford English Dictionary offers the following definitions (OED Online):

Nature
The phenomena of the physical world collectively; esp. plants, animals, and other features and products of the earth itself, as opposed to humans and human creations.

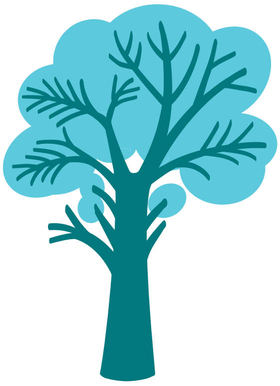
Culture
The distinctive ideas, customs, social behaviour, products, or way of life of a particular nation, society, people, or period. Hence: a society or group characterized by such customs, etc.

It is customary to think of nature and culture as being quite different, belonging to entirely separate domains, one contains items such as butterflies, the Amazon rainforest and photosynthesis, while the other contains items such as Beethoven's piano sonatas, wedding ceremonies or sushi. Yet nature and culture often interpenetrate and overlap. What is wine-making, bee-keeping or gardening: nature or culture? They are undoubtedly human activities, and each has its own culture, but there is a strong element of nature involved. What about varieties of domesticated plants and animals? They are human creations because their genomes have been altered by thousands of generations of selective breeding, and particular breeds may be associated with particular places or peoples, so they are as much a product of culture as of nature. What about landscapes? Is there anywhere left in the world that is entirely natural, untouched by human intervention? The deep sea bed perhaps, and possibly Antarctica; but most landscapes are, to a greater or lesser extent, the product of human culture too. Even the Amazon rainforest is what it is not just because of the natural evolution of its ecosystems, but also because of centuries of human manipulations to those ecosystems. So would it make more sense to think of all the myriad manifestations of nature and culture as expressions of a single concept, a nature-culture nexus?

BOTH NATURE AND CULTURE ARE WHAT THEY ARE AS A RESULT OF EVOLUTION, AND THEY HAVE EVOLVED IN SIMILAR WAYS

We can think of nature and culture as being dual aspects of a single entity, biocultural diversity; but not just because the two concepts are blurred at their interface. It is because both nature and culture, as defined above, are what they are as a result of evolution, and they have evolved in similar ways. So similar, in fact, that in this report we will describe culture and cultural evolution in the same terms as nature and natural evolution, using concepts borrowed from genetics, ecology and population biology. We will go on to examine the extinction crisis facing both biological and cultural diversity, and use methods developed in conservation biology to assess and compare the state of biodiversity with the state of cultural diversity, and contrast recent trends in the two.

In order to assess status and trends we need a unit of measure. Biodiversity and cultural diversity are such broad concepts that we need to focus on something specific and measureable, so we have chosen two fundamental units or classifiers of nature and culture: species and languages. Species are the basic units of biodiversity; languages are a useful proxy to stand for the world's diverse cultures. Other elements of biodiversity such as ecosystems or genes, and other aspects of culture such as religions, arts, or livelihood and subsistence strategies, are much harder to define and very much harder to measure.



There are striking parallels between species and languages (Harmon 2002). A species is a group of similar individual organisms that is capable of interbreeding. The ability to produce fertile offspring is fundamental to the biological definition of a species. Horses and donkeys belong to different species, even though they are closely related, as their offspring, mules, are infertile. Humans all belong to a single species, *Homo sapiens*. The genetic variation among humans is remarkably small, reflecting the fact that the modern human species is relatively young, only about 200,000 years old, and yet there is a staggering amount of cultural and linguistic variation among the human population (Pagel & Mace 2004). Linguists identify around 7,000 languages spoken worldwide (Lewis *et al*. 2013). By analogy with the definition of a species, two human individuals can be said to speak the same language if they can understand one another. If they find each other unintelligible, they are speaking different languages.[1] Dialects, by this definition, are analogous to subspecies: communication is possible between two individuals, although it may not be as easy. There are several subspecies of tiger, Siberian, Bengal or Sumatran for example, which can interbreed successfully in zoos, but their geographic ranges do not overlap in the wild. Given time, sadly something which is not on the tiger's side, the geographically isolated subspecies would evolve into reproductively separate species, a process known as speciation. New languages can evolve through a process that is akin to biological speciation, and the formation of dialects is the first step along the path to the evolution of two separate languages, provided that there is limited intercommunication between the two dialect populations.[2]

By using species to stand for all biological diversity and languages to stand for cultural diversity we are taking a narrow view, but making a useful simplification at the same time. Biological diversity is broader than species richness. It spans across scales from genes and proteins at the microscopic level to ecosystems and landscapes at the macroscopic level. Species lie somewhere in the middle, but as the carriers of genes and the components of ecosystems, they can fairly represent all biological diversity. In the same way, languages will stand as a proxy for all of cultural diversity, from the micro level of words, ideas and behaviours to the macro level of peoples and societies.

Fisherman, a speaker of a Trans-New Guinean language, hanging nets up to dry. Western Province, Papua New Guinea.

© Brent Stirton / Getty Images / WWF

EVOLUTION OF SPECIES AND LANGUAGES

Explaining biocultural diversity in terms of the Tree of Life

The parallels between species and languages have been noted and commented upon since the 19th century, famously by Charles Darwin in The Descent of Man (1874).

> *The formation of different languages and of distinct species, and the proofs that both have been developed through a gradual process, are curiously parallel…. We find in distinct languages striking homologies due to community of descent, and analogies due to a similar process of formation…. The frequent presence of rudiments, both in languages and species, is still more remarkable….In the spelling also of words, letters often remain as the rudiments of ancient forms of pronunciation. Languages, like organic beings, can be classed in groups under groups; and they can be classed either naturally according to descent, or artificially by other characters. Dominant languages and dialects spread widely, and lead to the gradual extinction of other tongues. A language, like a species, when once extinct, never, as Sir C. Lyell remarks, reappears.*

THE BIOCULTURAL TREE DIFFERS FROM THE USUAL VERSION IN THAT IT HAS GONE THROUGH NOT ONE BUT TWO DISTINCT TYPES OF BRANCHING OR DIVERSIFICATION

To illustrate the analogy between species and languages, picture that well-known Darwinian metaphor, the Tree of Life. The biocultural version of the tree differs from the usual version in that it has gone through not one but two distinct types of branching or diversification; the second diversification took place near the end of one the myriad outer twigs of the first tree (Figure 1). The first diversification was the evolution of multicellular organisms on Earth today, and the second diversification represents the evolution of human cultural diversity. Both of these evolutionary diversifications can be represented as trees, or phylogenies, but with one tree growing from the end of one branch of the other.[3] Figure 1 shows these two great radiations – the biological and the cultural – on a log scale. The first radiation took place near the bottom of the tree, between around 550 million years ago and the second radiation occurred about halfway up the tree at around 70-80 thousand years ago.

Life first appeared around 3.5-4.0 billion years ago; the earliest fossils of the simplest cells, bacteria, date back about 3.5 billion years. More complex life did not evolve until 1.8 billion years ago when the first plant and animal cells appeared in the form of algae and protozoa. These eukaryotic cells contained a nucleus to hold their DNA and had a more complex internal structure, which arose from the symbiotic union of prokaryotic (bacterial) cells, but remained unicellular. For the first three billion years of the evolution of life on earth the most complex organisms were single-celled organisms. It was not until about 550 million years ago that colonies of cells grouped together into the first multicellular life forms, known as the Ediacaran fauna, which resembled quilted discs and pillows.

The Ediacaran period lasted only a few millions of years before the quilted pillows were blown away in a massive, unprecedented and unrepeated diversification of animal biota that happened around 540 million years ago, known as the Cambrian Explosion. This explosion, or radiation, produced new life forms or species more rapidly than at any time before or since. Multicellular organisms appeared of enormous complexity by comparison with the Ediacaran fauna, including some of the most bizarre animals in the fossil record: many had hard body armour and possessed a range of formidable weaponry. Within a geological blink it was all over, but the Cambrian Explosion had produced myriad life forms including all known basic body plans of animals. The ancestors of arthropods, molluscs, annelid worms, echinoderms and all other modern phyla including the chordates (and therefore us) were there in one form or another, and all animal species since that time have conformed to the basic blue prints that evolved in that sudden burst of activity.

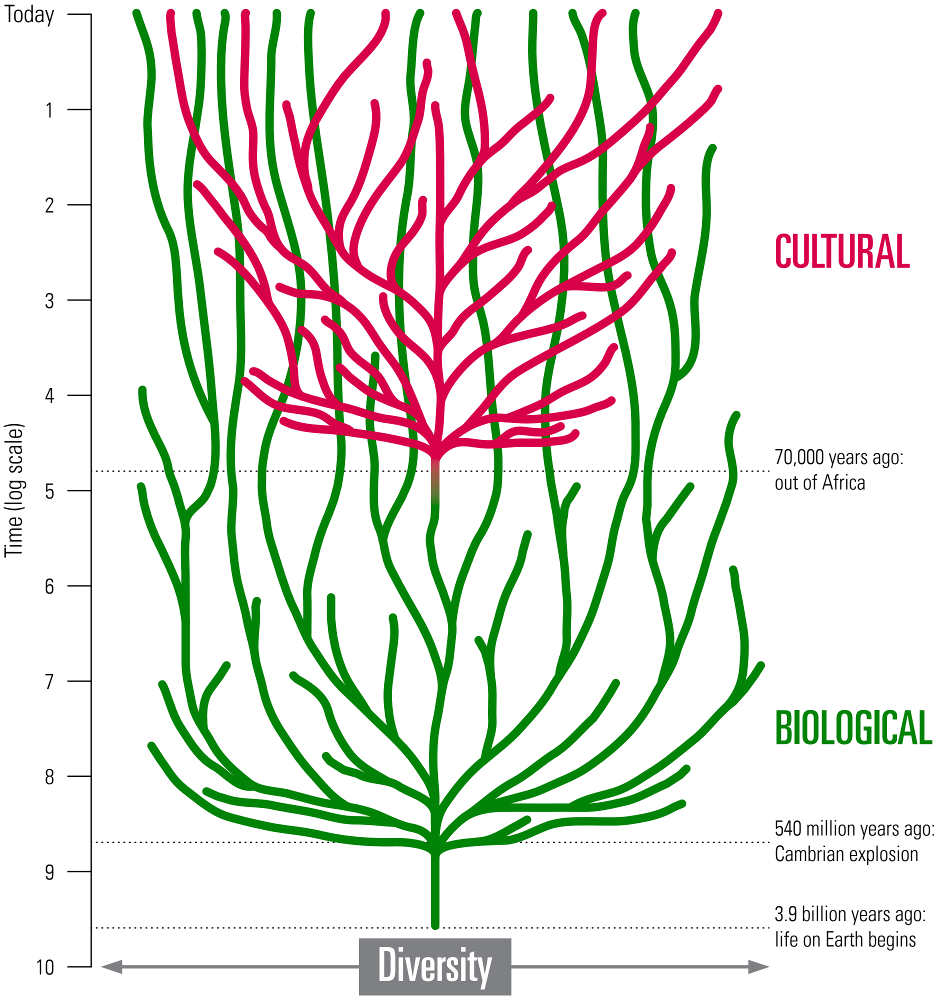


Figure 1: The Biocultural Tree of Life
The biological tree (in green) of species diversity began its diversification with the Cambrian Explosion around 540 million years ago; the cultural tree (in red) of linguistic diversity began to diversify about 70-80,000 years ago, near the end of one of the myriad branches of the biological tree.

The reason for the Cambrian Explosion is unknown, but a number of possibilities have been proposed. The entire planet during the time immediately preceding the Cambrian was glaciated, a period known as Snowball Earth (Walker 2003). The warming that ended Snowball Earth seems to have jump-started a new phase of multicellular evolution: initially the Ediacaran, followed by the Cambrian Explosion.

Another physical change at that time in Earth's history was a rise in the atmospheric oxygen content to its current level of around 21%, which would have aided the evolution of complex multicellular organisms. A third possibility is that a new type of gene that controls morphological development in the embryo, known as Hox genes, first appeared at the time of the Cambrian Explosion, enabling a plethora of new body plans to evolve.

The Cultural Explosion

The Tree of Life continued to branch and grow, continually evolving new species and losing old ones through extinction. Where whole groups of species died out, those branches stopped growing. This process went on for more than half a billion years, until the number of individual twigs at the outer edge of the tree numbered in the millions. Then an extraordinary, unparalleled event occurred at the end of one of its twigs. To an external observer, that twig would not have appeared exceptional, for although it represented a large mammalian species, it was by no means the biggest, or fastest, or the one with the most impressive body armour or weapons. But at some point, for reasons that are still unknown, the species on that twig began to talk. That species was our own, and as a result of our remarkable and unique innovation, language, the tree began a second massive evolutionary radiation, as significant as the Cambrian Explosion 540 million years earlier.[4]

THE HEREDITARY TRANSMISSION OF CULTURE IS MEDIATED NOT BY PASSING DNA FROM PARENT TO OFFSPRING, BUT BY ONE INDIVIDUAL LEARNING SOMETHING FROM ANOTHER, AND THIS TRANSMISSION IS GREATLY FACILITATED AND ACCELERATED BY MEANS OF LANGUAGE

Modern *Homo sapiens* first appeared only around 200,000 years ago, but we can trace our lineage back to the last common ancestor that we share with our closest living cousins, the chimpanzee and the bonobo, who lived about six million years ago. Exactly how or when language evolved is not known. But once it had taken hold it enabled an entirely new mode of evolution to take off – cultural evolution. Cultures evolve like species in many ways. Cultural items or traits are subject to hereditary transmission, variation by mutation and selection: the prerequisites of evolutionary change. Heredity in biology involves passing genetic information encoded in DNA from parent to offspring. The hereditary transmission of culture is mediated not by passing DNA from parent to offspring, but by one individual learning something from another, be it an idea, a behaviour, custom or another aspect of a way of life, and this transmission is greatly facilitated and accelerated by means of language. One can think of cultural information being transmitted as memes – the cultural analogue of genes (Dawkins 2006, Dennett 2002. See box "What is a Meme?" for further explanation). A meme, such as a song for example, existing in the brain of one individual is passed out of the mouth and via the ear into the brain of another individual. The meme has been copied, and can be replicated again and again in the brains of more individuals. Variation among memes occurs in a manner similar to mutation in genes. The tune of the song can be altered, the words can change, verses added or dropped. Selection is carried out by the individuals who come in contact with the meme, sometimes deliberately, sometimes unconsciously. A memorable, useful or otherwise interesting meme will be replicated many times, and spread successfully through a population – it becomes an element of a culture. Less memorable memes will be less successful; unmemorable memes will be forgotten.

© James Frankham / WWF-Canon

Ashaninka woman, a speaker of an Arawakan language. Nueva Victoria, Yurua River, Ucayali Province, Peru.

What is a Meme?

The word meme was coined by Richard Dawkins in his book The Selfish Gene (2006), and the word itself has become a successful meme. Dawkins proposed memes as the basic units of cultural evolution, and the idea has been developed by other thinkers such as Daniel Dennett (2002), although it remains controversial and is not widely accepted by sociologists and cultural theorists. Dawkins introduces the concept of cultural evolution with a linguistic example:

> *Geoffrey Chaucer [c.1343-1400] could not hold a conversation with a modern Englishman, even though they are linked to each other by an unbroken chain of some twenty generations of Englishmen, each of whom could speak to his immediate neighbours in the chain as a son speaks to his father. Language seems to evolve by non-genetic means, and at a rate which is orders of magnitude faster than genetic evolution.*

Just three conditions are necessary and sufficient for evolution to occur: replication (or heredity), variation (or mutation) and competition (or selection). If these three conditions are met, evolution will happen. In nature, the first two conditions are met by DNA, the molecule that encodes genetic information in all plants and animals, which replicates itself, but not perfectly (because of mutation). The third condition is provided by the fact that resources are finite, so individual organisms (and therefore their DNA) must compete for them in the Darwinian struggle for existence. But is there any other material apart from DNA which can replicate with variation and competes in the struggle for life? Yes, but not a material in the literal sense. Ideas. Memes are ideas which meet the three conditions. Firstly, an idea can be copied from the brain of one individual to the brain of another, so it can replicate. Secondly, ideas show variation from one individual brain to another, either because of imperfect copying or because of an innovation by an individual brain. Finally, ideas compete for finite brain-space in a population of brains, some are very successful and become very common, others are less successful and remain rare, and others still are not copied at all and go extinct. These ideas could be songs, stories, games, recipes, customs, clothing, art, technologies, anything in fact that constitutes culture. The basic unit of cultural heredity is the meme. The song “Happy Birthday to You” is an example of a phenomenally successful meme. As is wearing a tie. How to make an origami frog is a less successful meme, as it is more difficult to learn, and not very useful, but that also makes it a more interesting meme. A language is not a meme, it is a vast collection of memes working together – a meme complex – but which are largely copied as a group. Most elements of culture are in fact meme complexes. The rules of chess, for example, is a very successful set of memes, although not every brain remembers them all correctly. There are many other chess memes, such as opening gambits which are very successful at propagating themselves, but only among serious chess players.

It takes several years for the infant brain to acquire language, but once it has grasped it sufficiently, it enormously accelerates the learning of other types of behaviour, such as how to make tools, to hunt animals and gather edible plants, to cook, to grow crops and raise livestock, to make clothes, to build homes, to tell stories, to paint pictures, to write books, to play games, to do anything, in fact, that constitutes culture. Language is a tool for encoding and transmitting memes. Because of the fundamental importance of language to culture, linguistic evolution usually goes hand-in-hand with cultural evolution, and languages can be viewed as a proxy for cultural groupings in terms of the Tree of Life. It is possible for cultural behaviour to be transmitted between linguistic groups, but the transmission is faster and more accurate within a single group.

The 7,000 languages spoken in the world today represent the outermost twigs of the second tree, but of course there are many more extinct languages whose branches ended before reaching the outer edge of the tree. Like species, some languages can be classified into closely-related families, while others stand alone in families of one. Languages belonging to the same family have a common ancestor, just as families of species do. Germanic languages for example must have evolved from a single, ancestral proto-Germanic language. All Germanic languages belong to the much larger, extended family of Indo-European languages, which includes among others French, Irish, Greek, Russian, Persian and Hindi. Indo-European includes about 430 languages and is among the half-dozen largest families, of which the most diverse are Austronesian (about 1,250 languages spoken in Southeast Asia and the Pacific) and Niger-Congo (about 1,500 African languages) (Lewis *et al*. 2013).

LIKE SPECIES, SOME LANGUAGES CAN BE CLASSIFIED INTO CLOSELY-RELATED FAMILIES, WHILE OTHERS STAND ALONE IN FAMILIES OF ONE

LANGUAGES BELONGING TO THE SAME FAMILY HAVE A COMMON ANCESTOR, JUST AS FAMILIES OF SPECIES DO

Perhaps the biggest difference between biological evolution and linguistic-cultural evolution is speed. Biological evolution is slow while cultural evolution is so rapid that it can be observed taking place within our own lifetimes or even in front of our very eyes and ears: by watching a film made more than 50 or 60 years ago it is possible to hear how much language has changed in terms of pronunciation, accent and some words and phrases. Hence classifying languages into evolutionary families is tricky, as many if not all similarities between related languages can be erased within a few centuries. This is the problem that bugs the reconstruction of ancestral languages such as proto-Indo-European, and makes it paradoxically harder for linguists to draw phylogenetic trees and date the appearance and disappearance of languages than it is for biologists to draw and date phylogenetic trees of species. Another major difference between species and languages is that borrowing occurs far more readily between languages. Borrowing words is the equivalent of different species exchanging genetic material, something bacteria can do easily, but is less common among multicellular organisms. English is an example of a language whose origins lie in one part of the tree, the Germanic branch, but has incorporated a vast number of words from a language belonging to a neighbouring branch, French.[5] So horizontal transmission of language, and hence culture, does take place, but not so much as to destroy the basically tree-like structure of cultural diversity.

Walking the talk

Around 100,000 years ago the modern human population comprised somewhere in the order of 100,000 individuals, largely confined to the African continent, with a few living north of the Sahara and as far east as Palestine. Between 70-80,000 years ago, during the last ice age, people began to migrate out of Africa, probably crossing from the Red Sea to the Arabian Peninsula as sea levels were much lower, rather than moving north across the Middle East where deserts barred the way.[6] From Arabia, the migrants spread inexorably across Asia, probably following coastlines and moving up river valleys. Their descendants colonised South Asia first, and reached East Asia 60-70,000 years ago. About 40-60,000 years ago they succeeded in crossing the straits between mainland Southeast Asia and Australia.

One pathway led the migration north from the Arabian Sea up the Persian Gulf, at that time a broad, forested river valley into which the Tigris and Euphrates flowed. This was the route that led to Mesopotamia, Anatolia and eventually, around 30,000 years ago, to Europe. Another pathway led into northern Asia, and then, about 15,000 years ago, across the Bering Sea into North America. Within another thousand years the descendants of the first migrants to America reached Tierra del Fuego.

The last great migration was not by land but by sea, across the Pacific Ocean from Southeast Asia around 5,000 years ago, finally reaching New Zealand about 1,000 years ago (and, surprisingly, eastward across the Indian Ocean to reach Madagascar about 1,200 years ago). The human colonization of the globe, save for Antarctica, was complete.

Between 160,000 and 80,000 years ago, back in Africa, humans learned to make composite tools such as spears, decorate themselves with beads, catch seafood, use pigments and carry out ritual burials. It is possible that this was the period in which complex language, culture and art first appeared,[7] and people carried these with them as they crossed the globe. As they spread, living in small isolated groups, cultural and linguistic evolution would have rapidly given rise to thousands of local and regional variations, leading ultimately to a vast diversity of human languages and cultures.

Drawing the Family Tree

One way to look at linguistic evolution in action is to compare closely related modern languages. For example, English and Dutch are related languages in the Germanic family, descended from a common ancestral language that would have been spoken somewhere on the northwest coast of mainland Europe around 2,000 years ago. Frisian is a language that is spoken in Friesland in the north of the Netherlands[8] and is closely related to both Dutch and English. Scots is a language also very closely related to English that is still spoken at home by a substantial proportion of the population in Scotland; it contains many words that are closer to their Germanic roots than modern English, and would be unintelligible to most English speakers outside Scotland. It is easy to see how the languages have diverged from a common root by comparing words for the numbers from one to ten (Table 1).

Table 1: The names of numbers 1-10 in four Western Germanic languages

Number	Dutch	Frisian	Scots	English
1	een	ien	ane	one
2	twee	twa	twa	two
3	drie	trije	thrie	three
4	vier	fjouwer	fower	four
5	vijf	fiif	fyve	five
6	zes	seis	sax	six
7	zeven	sân	seiven	seven
8	acht	acht	aicht	eight
9	negen	njoggen	nyne	nine
10	tien	tsien	ten	ten

The similarity between English and Frisian is demonstrated by the saying “Good butter and good cheese is good English and good Fries”, which when spoken sounds virtually the same in both languages. Tellingly, the saying also demonstrates the importance of dairy-based agriculture in both cultures. So it is not difficult to imagine members of the Frisii tribe after the end of the western Roman Empire in 410 CE (along with their cousins the Anglii and the Saxones) crossing the North Sea to Britain, at first as raiders but later as settlers, taking with them their language and perhaps a few of their cows.

© Martin Harvey / WWF-Canon

Baka woman, a speaker of a Niger-Congo language, with collecting basket, gathering plants in the forest of La trinationale de la Sangha. Central African Republic.

Linguists have made comparisons between the thousands of languages spoken in the world in order to work out the evolutionary relationships between them. These comparisons rely on similarities between words which have descended from common ancestral words, like the numbers one to ten in Dutch, English, Frisian and Scots. Such words are termed cognate, and by knowing how sounds have changed systematically in different languages over time, comparative linguists have reconstructed language phylogenies or family trees.

European visitors to India as early as the sixteenth century began to notice similarities between Sanskrit, Latin and Greek, but the most famous of these was William Jones (1746-94), who is considered to be the founding father of comparative linguistics. Jones was a scholar and magistrate living in Calcutta in the 1780s. He was a polyglot fluent in a dozen languages and familiar with two dozen more. He became fascinated with Indian culture and co-founded the Asiatic Society of Bengal in 1784. In a paper he delivered to the Asiatic Society in 1786 he noted that Sanskrit, Latin and Greek bear

> *...a stronger affinity, both in the roots of verbs and the forms of grammar, than could possibly have been produced by accident; so strong indeed, that no philologer could examine them all three, without believing them to have sprung from some common source, which, perhaps, no longer exists... there is a similar reason, though not quite so forcible, for supposing that both the Gothic and the Celtic... and the old Persian might be added to the same family* (Jones 1824).

That common source came to be known as Proto-Indo-European (or sometimes proto-Indo-Germanic) or PIE. It is not difficult – with hindsight - to see how Jones and others reached their conclusion. Look, for example, at the names of the numbers one to ten in Sanskrit, Latin, Ancient Greek, Gothic (an old Germanic language), Welsh and Hindi.

Table 2: The names of numbers 1-10 in selected ancient and modern Indo-European languages

Number	Ancient Greek (c.400 BCE)	Latin (c.100 BCE)	Sanskrit	Hindi (modern)	Gothic (Germanic c.350 CE)	Welsh (modern)
1	oinos	una	eka	Ek	ains	un
2	duo	duo	dvi	do	twai/twos/twa	dau/dwy
3	treis	tres	tri	tin	þreis	tri/tair
4	tessares	quattuor	chatur	car	fidwor	pedwar/pedair
5	pente	quinque	pancan	panch	fimf	pump
6	hex	sex	sash	chhah	saíhs	chwech
7	hepta	septem	saptan	sat	sibun	saith
8	okto	octo	ashta	Ath	ahtau	wyth
9	ennea	novem	navan	nau	niun	naw
10	deka	decem	dasan	das	taíhun	deg

By using this method a family tree for the Indo-European languages was being put together even before Darwin proposed his theory of evolution of species by descent from a common ancestor (see Figure 2).

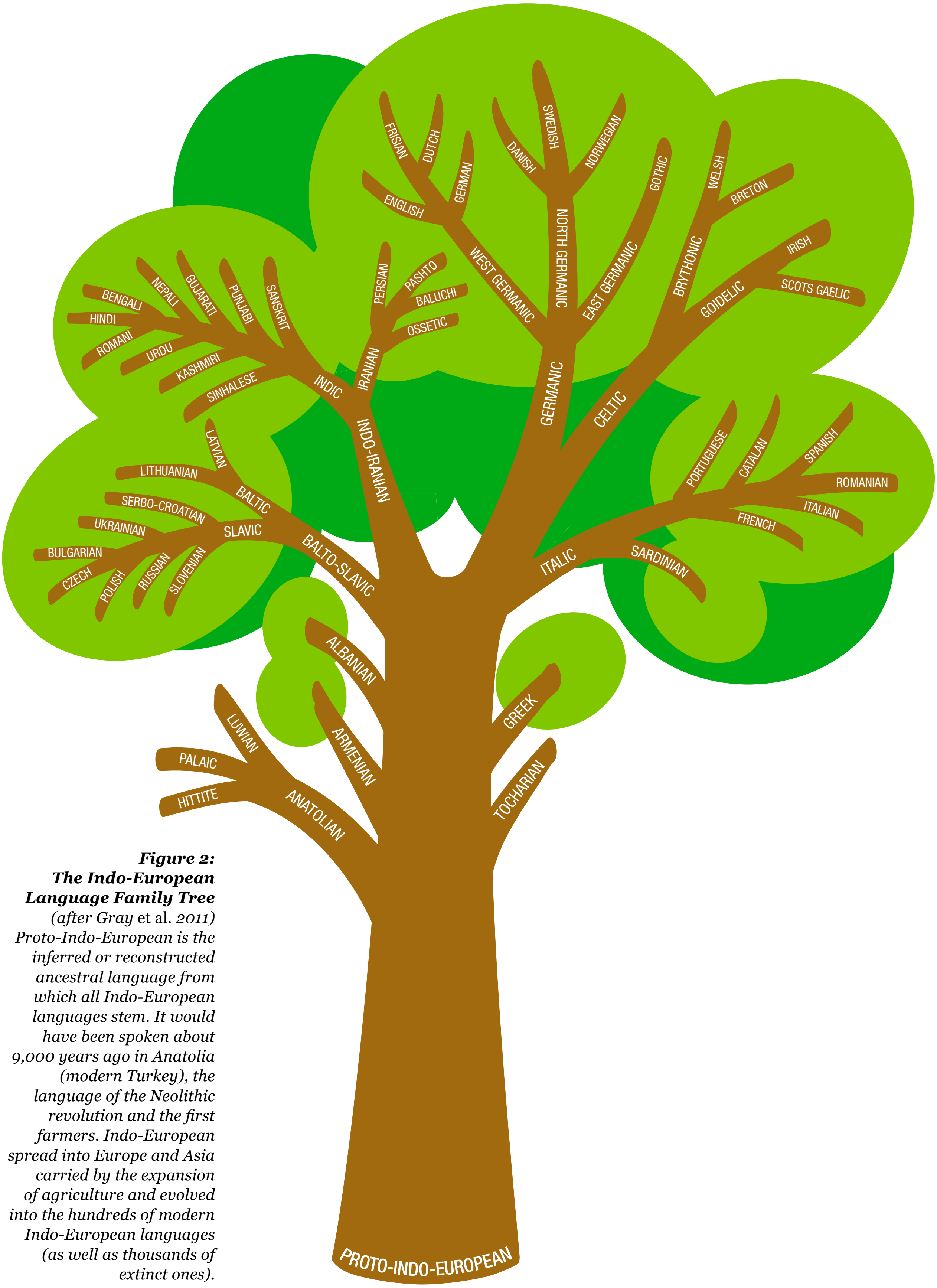


Figure 2:
The Indo-European
Language Family Tree
(after Gray et al. *2011)*
Proto-Indo-European is the inferred or reconstructed ancestral language from which all Indo-European languages stem. It would have been spoken about 9,000 years ago in Anatolia (modern Turkey), the language of the Neolithic revolution and the first farmers. Indo-European spread into Europe and Asia carried by the expansion of agriculture and evolved into the hundreds of modern Indo-European languages (as well as thousands of extinct ones).

So who were the original Proto-Indo-Europeans and where did they come from? One of the most surprising findings of the comparative linguists was that some extinct languages spoken in Bronze Age Anatolia (modern Turkey) such as Hittite belonged to the Indo-European family, even though modern Turkish, a member of the Altaic language family, does not. The Hittite language was known only from cuneiform inscriptions on clay tablets dating from the second millennium BCE and deciphered by the Czech linguist Bedrich Hrozný in the early 20th century (Hrozný 1917). The key breakthrough came when Hrozný found the word *watar* in a sentence alongside an ideogram, or symbol, known to mean 'bread' in Sumerian. The similarity of *watar* to water, or Wasser in German, and the similarities of another word in the sentence to eat or essen, and of another to aqua in Latin, led him to guess the meaning was something like eat bread, drink water. But the shock was that this finding placed Hittite in the Indo-European family, not with the Altaic family of central Asian languages, or the Afro-Asiatic family along with other Middle Eastern languages such as Arabic (see figure 9).

A great many theories have been proposed to locate the homeland and time of the Proto-Indo-Europeans. Recently, scientists have re-examined the linguistic evidence using computational methods developed to determine the evolutionary relationships between species based on their DNA. Instead of comparing DNA, Gray and Atkinson (2003) compared a list of 200 words in 87 languages to reconstruct the Indo-European family tree (see Figure 2). They were able to date its origin to about 8-9,000 years ago (Gray and Atkinson 2003, Gray *et al.* 2011), supporting the theory proposed by the archaeologist Colin Renfrew that the Indo-European languages were carried from Anatolia into Europe and South Asia on a cultural tsunami caused by the greatest seismic event in prehistory – the Neolithic revolution, or the adoption and spread of agriculture (Renfrew 1987).

A number of comparative linguists such as Joseph Greenberg have attempted to construct higher-order language families, which unite several families into a single grouping, equivalent to a class or phylum in zoology (Greenberg 2000). Indo-European has been combined with several other families including Uralic (which covers Finland and western Russian Arctic), Altaic (Siberia, Mongolia, Central Asia and Turkey) and Eskimo-Aleut (North American Arctic) languages into a phylum called Nostratic or Eurasiatic. Greenberg also proposed a super-family called Amerind to include all but two indigenous American languages families. These proposals are not widely accepted among linguists, but there is some support for Eurasiatic, based on quantitative techniques used by evolutionary biologists, who estimate the date of the proposed ancestral language, proto-Eurasiatic, would have been around 15,000 years ago, close to the end of the last ice age (Pagel *et al.* 2013).

Some experts have even attempted to link languages as far apart geographically as Basque and Navajo, along with a few northern Caucasian and Siberian languages, into a phylum called Dene-Caucasian.[9] Such heroic attempts at reconstructing deep historical links between languages are highly controversial among linguists. If such a language as proto Dene-Caucasian ever existed, or more likely but still highly controversially, if not a language then at least some proto-Dene-Caucasian words existed, they would have been spoken right back at the height of the last ice age, by Palaeolithic hunter-gatherers living alongside mammoths and woolly rhinoceroses somewhere in Siberia. However remote this possibility may sound, there is some genetic evidence that could support this theory, from the 24,000-year-old remains of a young boy buried under a stone slab in the village of Mal'ta near Lake Baikal (Raghavan *et al.* 2014). Analysis of his genome revealed European ancestry which suggests that there had been an eastward migration of people into Siberia from Europe.

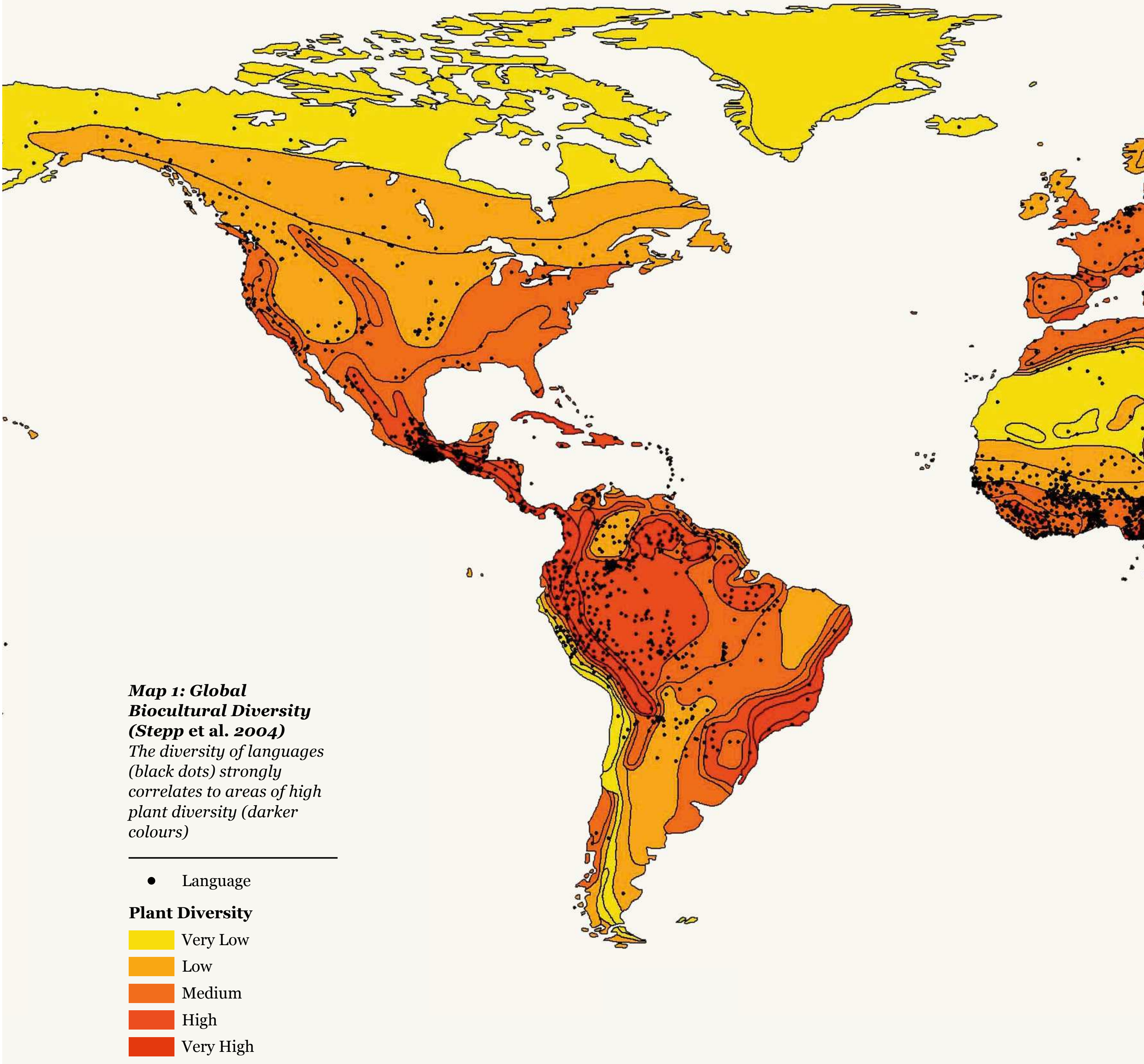
But more surprisingly, his DNA showed he is also ancestral to modern Native Americans, which suggest that some descendants of the Mal’ta population interbred with East Asians in Siberia who then migrated across the Beringian land bridge to the Americas around 15,000 years ago. Is it possible that the Basque and Navajo languages retain some residual cultural imprint from those times?

Basque is one of the few remaining languages left in Europe that is not Indo-European in origin. Genetically, the Basque people also show some differences from other European populations (Cavalli-Sforza 2000). Interestingly, places with Basque-derived names and sites of Palaeolithic cave art overlap geographically in southwest France and northern Spain. The earliest cave paintings date back 30-40,000 years, marking the beginning of western art, and coinciding with the arrival of modern *Homo sapiens* in Europe (long after the arrival of the Neanderthals, who soon disappeared). The cave artists continued to produce their work for another 20,000 years, achieving their *magnum opus* in the extraordinary paintings of horses, bison, aurochs, reindeer, as well as more abstract images of humans, at Lascaux in France and Altamira in Spain, around the time of the last glacial maximum around 17-18,000 years ago. Is it possible that these artists spoke a language that was the ancestor of Basque? Indo-European languages spread into Europe from Anatolia alongside the adoption of agriculture, as settled farming replaced nomadic hunter-gathering as the primary way of life, and the original languages spoken by the first modern humans in Europe fell like dominoes. Basque, for reasons that are unknown, is the last domino standing, a language isolate descended from Palaeolithic hunter-gatherer-artists.

SOME LINGUISTS HAVE ATTEMPTED TO CONSTRUCT HIGHER-ORDER LANGUAGE FAMILIES, WHICH UNITE SEVERAL FAMILIES INTO A SINGLE GROUPING, EQUIVALENT TO A CLASS OR PHYLUM IN ZOOLOGY

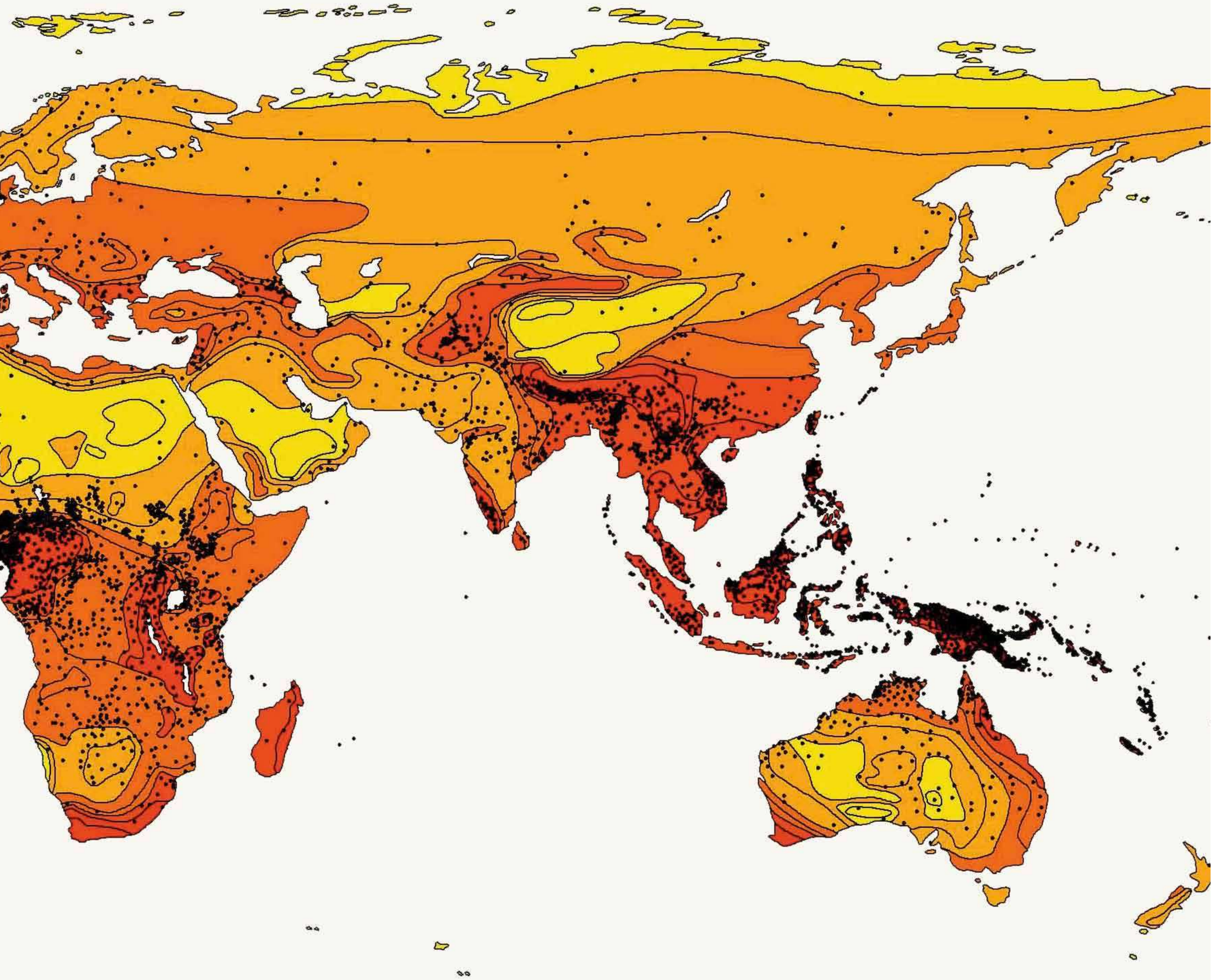
A big coincidence

There is another way in which the evolution of languages mirrors the evolution of species: the similarity in the geographic distributions of languages and of species around the world. Places with high species diversity, especially tropical forests, tend to have high linguistic diversity, and areas of low species diversity, such as tundra and deserts, have low linguistic diversity (Mace & Pagel 1995, Nettle & Romaine 2000, Moore *et al*. 2002, Sutherland 2003, Stepp *et al*. 2004, Loh & Harmon 2005). The island of New Guinea which makes up less than one percent of the Earth's habitable surface, apart from being one of world's biodiversity hotspots with endemic species such as birds of paradise and tree kangaroos, supports around



Map 1: Global Biocultural Diversity (Stepp* et al. *2004)
The diversity of languages (black dots) strongly correlates to areas of high plant diversity (darker colours)

1,000 languages, one seventh of the total. A glance at Map 1 confirms this view, and that it is not just that places with greater population density have greater language density. It is well known to biologists that species density per unit area is highest in equatorial regions and declines towards the poles – a pattern known as Rapaport's rule – and languages obey it too. It is possible that the one causes the other, that in some way higher biodiversity is capable of supporting greater cultural diversity, but the explanation seems to be that both biological and cultural diversity depend on the same environmental factors such as temperature and rainfall (Nettle 1999, Moore *et al*. 2002, Sutherland 2003).



San hunters, speakers of a Khoi-San language, Namibia. Unlike other language families, Khoi-San languages use clicks made by the tongue and sharp intakes of air. The Hadza people (front cover), 2,500 km away in Tanzania, also speak a click language (see figure 9). Genetic evidence suggests that the most recent common ancestor of these two peoples lived as long as 50-70,000 years ago, around the time that modern humans left Africa. It is possible that the very first languages ever spoken were click languages, and that clicks evolved before vocal words as a means of communicating without scaring animals when hunting (Pennisi 2004).

© Martin Harvey / WWF-Canon

DECLINE OF BIOCULTURAL DIVERSITY

The extinction crises facing both species and languages as consequences of similar processes

Until now, the story has been about evolution and diversification. There is another side to the story, decline and extinction. Most species that ever existed have gone extinct. They are the inner branches and twigs of the tree that stopped growing (or evolved into another species) before reaching the outermost edge. Over and above the background extinction rate, there have been at least five biological mass extinction events since the Cambrian explosion, in which global species diversity was suddenly reduced.

The third event was the greatest, 245 million years ago, in which 96% of species went extinct, and the fifth, 65 million years ago, marked the demise of the dinosaurs. Following each mass extinction event, however, biodiversity recovered to or exceeded its previous high level. The present rate of species loss may be in the region of 100-200 times higher than the background rate found in the fossil record (Groombridge and Jenkins 2002), which puts us in the midst of a sixth mass extinction. But this extinction event is cultural as well as biological (Nettle and Romaine 2000).

According to Ethnologue (Lewis *et al.* 2013), a periodic publication dating back to the 1950s which compiles data on the world's languages and speaker numbers, half of the world's population speaks one of only 24 languages, the top ten being Mandarin Chinese, Spanish, English, Hindi, Portuguese, Bengali, Russian, Japanese, Javanese and German.[10] These two dozen languages have speakers numbering in tens or hundreds of millions. The other half of the world's population speak the remaining 7,000 languages (see Figure 3).

The Ethnologue data describe an enormously skewed distribution of speakers among the world's languages. Figure 4 shows that the frequency of languages of different sizes forms a normal, bell-shaped curve, but on a log scale – each category along the horizontal axis of the graph is ten times the size of the previous one. Around half of the world's languages has fewer than 10,000 speakers, and the other half has more than 10,000. But 95% of the world's population are found in the three size classes at the right-hand end of the bell curve, they speak languages spoken by millions, tens of millions or hundreds of millions of people. Forty percent of us occupy the tiny group of languages with 100 million-plus speakers. At the other end of the distribution, just over one percent of the world's population are responsible for maintaining over 5,000 languages, those with fewer than 100,000 speakers. Astonishingly, only about 0.1% of the world population or about 8 million people, equivalent to a city about the size of London, are responsible for keeping one half, or about 3,500, of the world's languages alive.

Figure 3:
World languages
More than half the world's population speaks one of just two dozen languages (source: Lewis et al. *2013).*

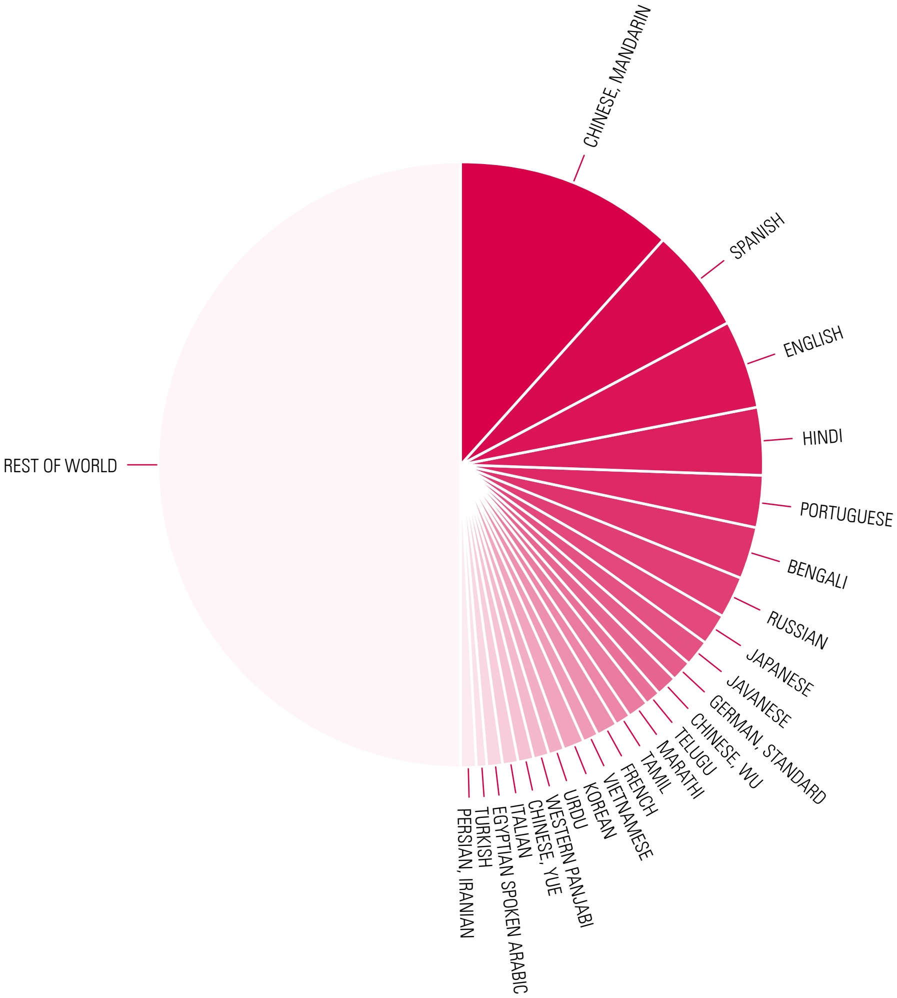
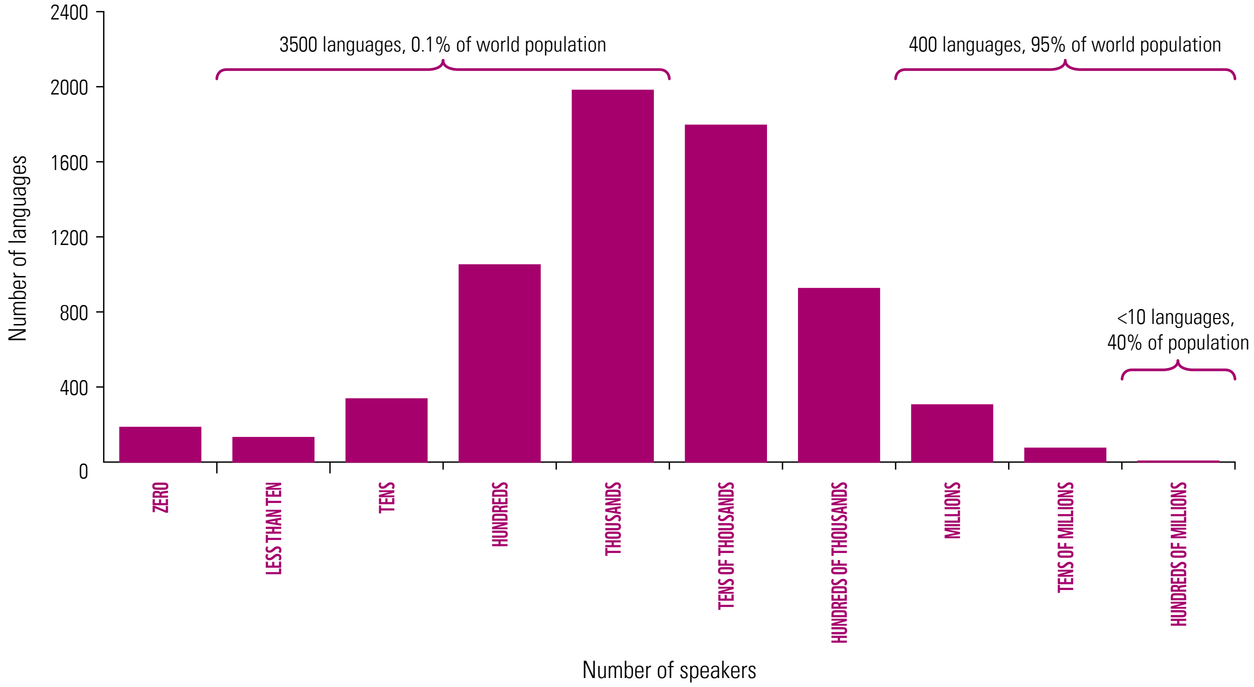


Figure 4:
Sizes of languages
Number of languages by size class (after Harmon 1995, 2002; data source: Lewis et al. *2013).*



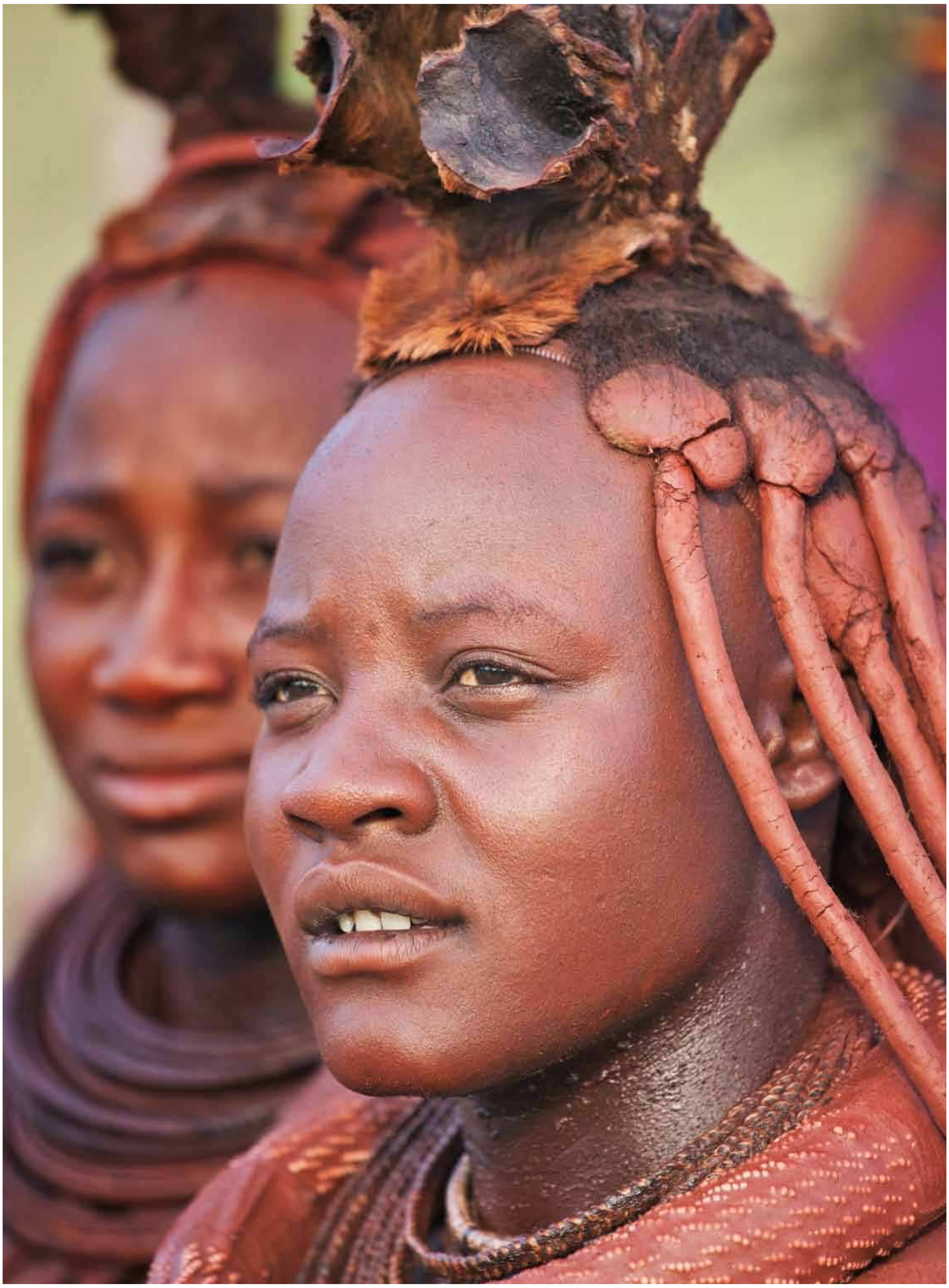
> LANGUAGES GO EXTINCT EITHER BECAUSE THE ENTIRE POPULATION OF SPEAKERS DIES OUT OR, MORE USUALLY, BECAUSE THE SPEAKERS SHIFT TO A DIFFERENT LANGUAGE AND, TYPICALLY WITHIN A FEW GENERATIONS, FORGET THEIR MOTHER TONGUE

A comparison with the distribution of language sizes two decades ago (Harmon 1995) reveals that, while the world's population has grown by about 25%, the number of million-plus languages has expanded but small languages have dwindled away. The distribution is gradually shifting to the right like a wave and becoming even more skewed. The only group on the left of the graph that has grown is the zero (extinct) class. Some linguists predict that 90% of the world's languages will die out this century (Nettle 1999, Nettle & Romaine 2000). Why is this happening?

Languages can go extinct either because the entire population of speakers dies out or, more usually, because the speakers shift to a different language and, typically within a few generations, forget their mother tongue. This can happen for social or economic reasons, such as commerce or migration, or through a deliberate policy of linguistic unification by a dominant group (see box: Linguistic Ecology). The globalization of trade and media, and technological progress in transport and communication, have accelerated the process of language shift, as have nationalization policies that favour a small number of languages, increasing the pressure on languages with thousands or fewer speakers, and boosting the dominance of those with millions. As language is the primary medium of cultural transmission, linguistic diversity and cultural diversity are being diminished simultaneously.

Most of the languages threatened with extinction are evolutionarily quite distinct from the few dominant world languages, and so they also represent very different cultures. Nearly all are spoken by indigenous people, some still living in traditional ways on their ancestral lands, although these are becoming rare. Along with the languages, the traditional knowledge of these indigenous cultures is being forgotten. The names, uses and preparation of medicinal and food species, both plant and animal, and traditional methods of farming, fishing, hunting and natural resource management are disappearing, not to mention the vast array of spiritual and religious beliefs and practices that are often associated with traditional land use and resource management, which are as diverse and numerous as the languages themselves. This vast store of knowledge that has evolved and accumulated over tens of thousands of years could be lost in the course of just two centuries, the 20th and the 21st. While linguists have made great efforts to document, record and archive as many of the endangered languages as possible, and ethnobiologists have attempted to record the traditional uses of plants and animals by indigenous peoples, the most important conservation takes place on the ground, as part of a living culture.

Conserving linguistic and cultural diversity presents a quite different ethical problem compared with the conservation of biodiversity. There are very strong utilitarian and economic arguments for protecting species and maintaining natural ecosystems, but there is also a moral argument that no species should be extirpated for human purposes. Cultures and languages on the other hand can only be maintained by people who choose to, usually but not necessarily the ethnic group with which the culture evolved, nobody should be forced to speak a language or practise a culture if they do not want to. Most indigenous peoples, of course, do want to keep their language and culture alive, but they may not have the opportunities or means or numbers to sustain it.



© Martin Harvey / WWF-Canon

Himba woman, a speaker of a Niger-Congo language. The Himba are semi-nomadic pastoralists who, unlike many indigenous groups in Africa, have managed to maintain much of their traditional lifestyle. Kunene Region, Namibia.

Linguistic Ecology

When linguists discuss the ecology of a region or a country they are not thinking of the relationships between its species and their environment. They are talking about the languages spoken in an area and the dynamics of the interactions between them and the social and political context in which they exist. One of the dominant forces in linguistic ecology is language shift. Language shift occurs when a population of speakers adopts a new language at the expense of their mother tongue, generally over the course of a few generations, and is the biggest driver of language extinction.

A well-documented, on-going example of the process of language shift comes from Britain, where Scottish Gaelic[11] has been losing speakers to English over the last 200 years (MacAulay 1992).[12] In the mid-18th century the population of Scotland was around 1.25 million, consisting of about 300,000 Gaelic speakers, concentrated in the Highlands and Islands, and nearly one million English speakers, concentrated in the Lowlands. The Highland clearances, a programme of removing small-scale farmers from their land to make room for large-scale sheep farmers, and the consequent migration of Gaelic speakers to the Lowlands or away from Scotland altogether led to a steady decline in their number. By the end of the 19th century, monolingual Gaelic speakers had mostly disappeared, and nearly all the remaining Gaelic speakers were bilingual. Today the population of Scotland is around five million, with about 58,000 Gaelic speakers, just above one percent of the total (note that Scottish Gaelic is still above the world median language size, and therefore one of the worlds' larger languages). The Scottish government has made efforts to promote primary education in Gaelic and, although the number of Gaelic speakers continued to decline between 2001 and 2011, the number of speakers aged under 20 remained stable.

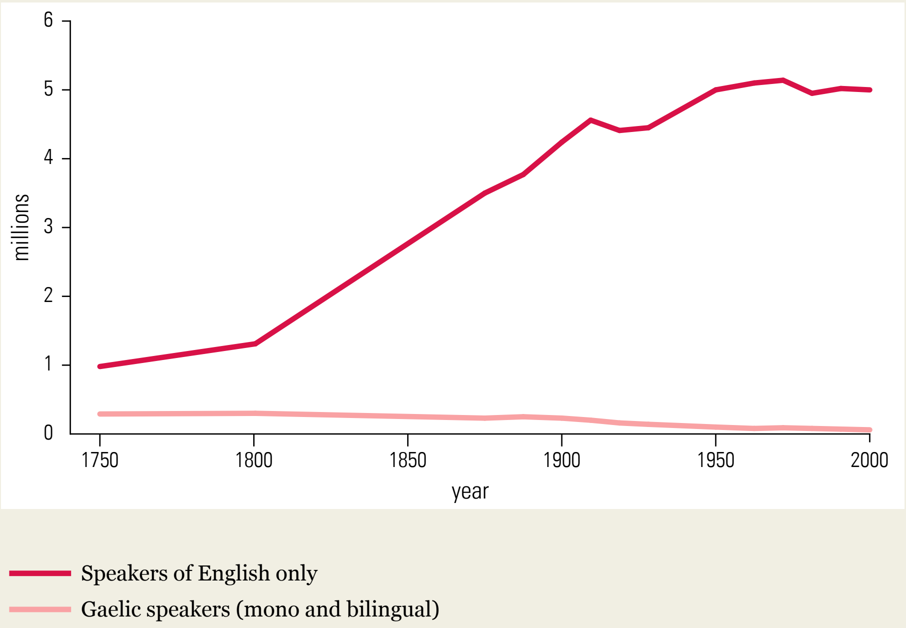


Figure 5: Language Shift in Scotland
(MacAulay 1992)

Of the other Celtic languages, Irish, the closely-related sibling of Scottish Gaelic, is declining alongside it, and Breton (spoken mainly in Brittany, France) is declining faster. The last mother tongue Cornish[13] speaker died in 1777 and the last Manx[14] speaker in 1974, although attempts are being made to keep them alive as second languages.

Welsh is the only Celtic language with a strong speaker base owing to decades of support from the educational system and government policy. Celtic languages have been struggling along beside far larger, socially and politically dominant languages, English and French, for more than a thousand years. The British Isles had an entirely Celtic-speaking population up until the time of the Roman invasion, and remained predominantly Celtic-speaking until the arrival of the Anglo-Saxons in the 5th century. Gradually the Celtic languages were pushed to the western fringes where they survive today. Language shift, to be very clear, does not mean that one population replaces another, but that one language is displaced by another within the same population. The peoples who spoke Celtic languages are still there, genetically the population is still largely Celtic, even in England.

Nenets reindeer herdswoman, a speaker of a Uralic language, eating reindeer meat, Kánin Peninsula, Russia, Arctic.

© Staffan Widstrand / WWF

STATUS OF SPECIES AND LANGUAGES

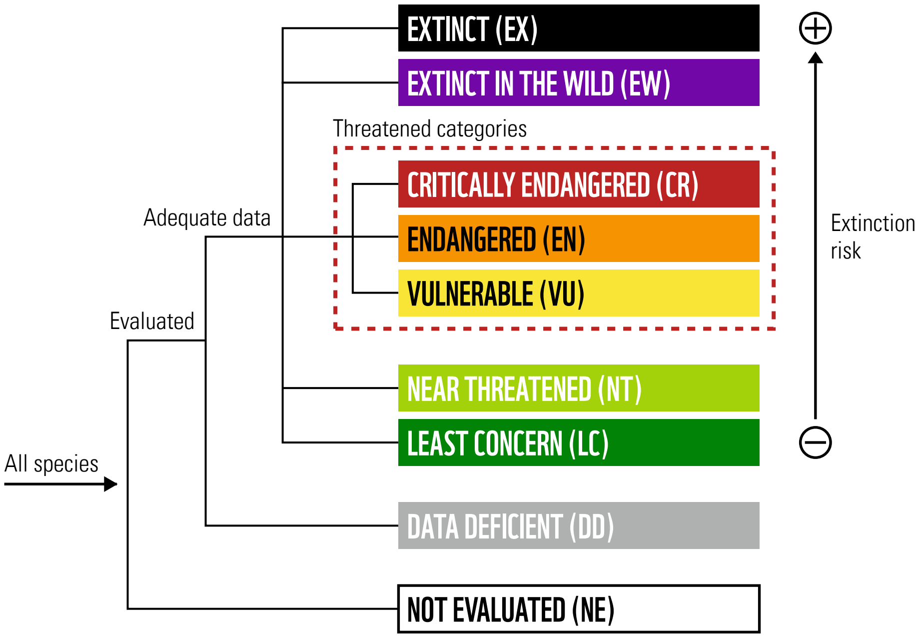
Global similarities and regional differences in the state of biological and linguistic diversity

Because species and languages are alike in terms of their evolution, diversity, and distribution around the world, it is appropriate and feasible to assess their current status in similar ways, and compare the two. We have adapted and applied two methods developed for assessing the state of biodiversity to measure the state of linguistic diversity. The first is the IUCN Red List system which is used to assess the extinction risk to species (IUCN 2013); the second is the WWF/ZSL Living Planet Index which measures the rate at which biodiversity is declining (Loh *et al.* 2005, Collen *et al.* 2009).

Threat Status of Species – Red Listing

The IUCN Red List is a system used by biologists to assess the conservation status of plant and animal species. It is based on a set of categories for ranking species according to their risk of extinction. There are seven categories ranging from Least Concern to Extinct. There is an eighth category for species which have been evaluated but for which there are insufficient data to assess their status. Those species which are categorised as vulnerable, endangered or critically endangered are considered to be threatened.

Figure 6:
IUCN Red List categories
(IUCN 2013)



Only a fraction of all species have been evaluated, but a few taxa (species groups) have been completely evaluated and some have had a random sample of 1500 species evaluated. Among the vertebrates, all mammals (5,501 species known and described to date), birds (10,064 species) and amphibians (6,771 species) and a random sample of 1,500 species of all reptiles (approximately 9,000 species) have been recently assessed. More than 10,000 fish species have been assessed, but as the sample is not random, and the total number of fish species is very large but uncertain (about 32,000 known to date), no firm conclusions can be reached about the status of fish as a group. These assessments are used here to compare the threat status of vertebrate species with the status of languages, based on a random sample of 1500 languages. For the purposes of these comparisons, we combine the category extinct in the wild (EW) with extinct (EX), and the category near threatened (NT) with least concern (LC).

The criteria used to categorise the conservation status of a species into one of the Red List categories include a species' population size, its rate of reduction (if in decline), its range size and rate of decline or fragmentation, existing and future threats, or a combination of these. It is possible to apply some of these criteria to languages and assess their threat status according to either the number of mother-tongue speakers, their rate of decline, or a combination of the two. Range size is harder to apply to languages and therefore was ignored in this analysis, as was existing or projected threat. Because biologists use a wider range of criteria to assess species than has been applied here to languages, the threat status of languages should be considered more conservative.

Threat Status of Languages – UNESCO and Ethnologue

Linguists consider a language to be endangered if it is not being transmitted successfully from one generation of speakers to the next. This is very good reasoning, but it means that the criteria used by linguists to assess the threat status of a language are quite different to the IUCN criteria used by biologists. Ultimately the two sets of criteria, linguistic and biological, are designed to assess extinction risk. Table 3 compares the Red List criteria we have applied to a random sample of 1,500 languages with the criteria used in two systems designed to assess threatened languages, UNESCO's Language Vitality and Endangerment system (UNESCO 2010) and Ethnologue's Expanded Graded Intergenerational Disruption Scale or EGIDS system (Lewis and Simons 2010) The systems are not correlated: critically endangered in the Red List system does not necessarily correspond to critically endangered in the UNESCO system for example; the only category that means the same in all three systems is extinct.

Table 3: Definitions of categories under three systems of assessing the status of languages

Red List (as applied here)	UNESCO	Ethnologue (EGIDS)
Extinct (EX): No speakers remain.	**Extinct:** No one can speak the language.	**Extinct:** The language is no longer used.
Extinct in the Wild (EW): Not applicable.		**Dormant:** The language serves as a reminder of ethnic identity but no proficient speakers remain.
Critically Endangered (CR): Either the number of speakers is observed or projected to decline by 80% or more in three generations (75 years); or speakers number less than 250 and declining by 25% or more in one generation (25 years); or speakers number less than 50.	**Critically endangered:** Youngest speakers are great-grandparents; language not used on a regular basis; language only partially remembered.	**Nearly Extinct:** Only spoken by great-grandparent's generation who have little opportunity to use the language.
Endangered (EN): Either no. speakers observed or projected to decline by 50% or more in three generations (75 years); or no. speakers less than 2,500 and declining by 20% or more in two generations (50 years); or no. speakers less than 250.	**Severely endangered:** Language spoken only by grandparents' and older generations; parents understand but do use it to speak to their children or each other.	**Moribund:** Only speakers are grandparents' generation.
	Definitely endangered: Youngest speakers are parents' generation; children are not using the language at home.	**Shifting:** Parents' generation use the language among themselves but it is not being transmitted to their children.
Vulnerable (VU): Either no. speakers observed or projected to decline by 30% or more in three generations (75 years); or speakers number less than 10,000 and declining by 10% or more in three generations (75 years); or speakers number less than 1,000.	**Vulnerable:** Most children speak their parental language as their mother first language, but usage is restricted to the home or particular social situations.	**Threatened:** The language is used by all generations, but it is losing users.
Near Threatened (NT): The language does not meet the criteria for CR, EN, or VU but is likely to do so in the near future (this category has not been used in this assessment).	**Stable yet Threatened:** The language is spoken by all generations in most contexts, but multilingualism is common and a more dominant language is taking over in some contexts.	
Least Concern (LC): The language does not fall into any of the categories above; speakers are widespread and abundant.	**Safe:** The language is spoken by all generations; inter-generational transmission is uninterrupted.	**Vigorous:** The language is used by all generations, and the situation is sustainable.

Ethnologue further defines a number of higher categories for languages in vigorous use: namely where standardized literature is in use but not widespread (Developing); standardization and literature are in widespread use in education (Educational); the language is used at work and in mass media but without official status as a national or regional language (Wider communication); used in education, work, mass media and government at provincial or national level (Provincial, National); used internationally for trade, knowledge exchange or policy (International).

The UNESCO and Ethnologue EGIDS systems use inter-generational transmission as the principal criterion in assessing a language's vitality, defined according to the number of generations that speak the language: great-grandparents only, grandparents and older, parents and older, or all including children. While there is an undeniable logic to these systems, there are some good reasons for using the IUCN Red List system developed by biologists to assess the status of a language. Firstly, if children are no longer speaking their parental language, unless there is great effort to revitalize the mother tongue, it is inevitable that the language will move up through the categories towards extinction. However, if a language that is close to extinction were to undergo a massive revitalization effort, it would not move back down through the categories as first the grandparents, then parents and finally children learn to speak the language once again. The linguistic categories assume there is one way traffic up the ladder to extinction. But it should be possible to track a reversal in the fortunes of a language dropping back down the categories, which is the case if the Red List criteria are applied. Secondly, the status of a language may change from location to location, or even from family to family, as children could be speaking their mother tongue in some places, while only parents or grandparents use the language in others. The Red List criteria are not concerned with the age of the speakers, only the total numbers. Of course, the end result of a breakdown in inter-generational transmission will be a decline in speaker numbers, so the Red List criteria are focusing on the ultimate effect rather than the direct causes of endangerment.

The linguistic criteria recognize that a language may be safe or vigorous even if it is only spoken by a very small population, as long as inter-generational transmission is uninterrupted. The biological criteria conversely consider a language to be threatened simply if the number of speakers is below a critical threshold (1,000 for vulnerable, 250 for endangered, 50 for critically endangered), even if there is no decline through the generations. This is justifiable as it is precisely when the mass of speakers is small that a language could be threatened by a shift away from the mother tongue towards a more dominant language by means of unforeseen events extraneous to the process of intergenerational language transmission.

Comparison of Conservation Status of Languages and Species

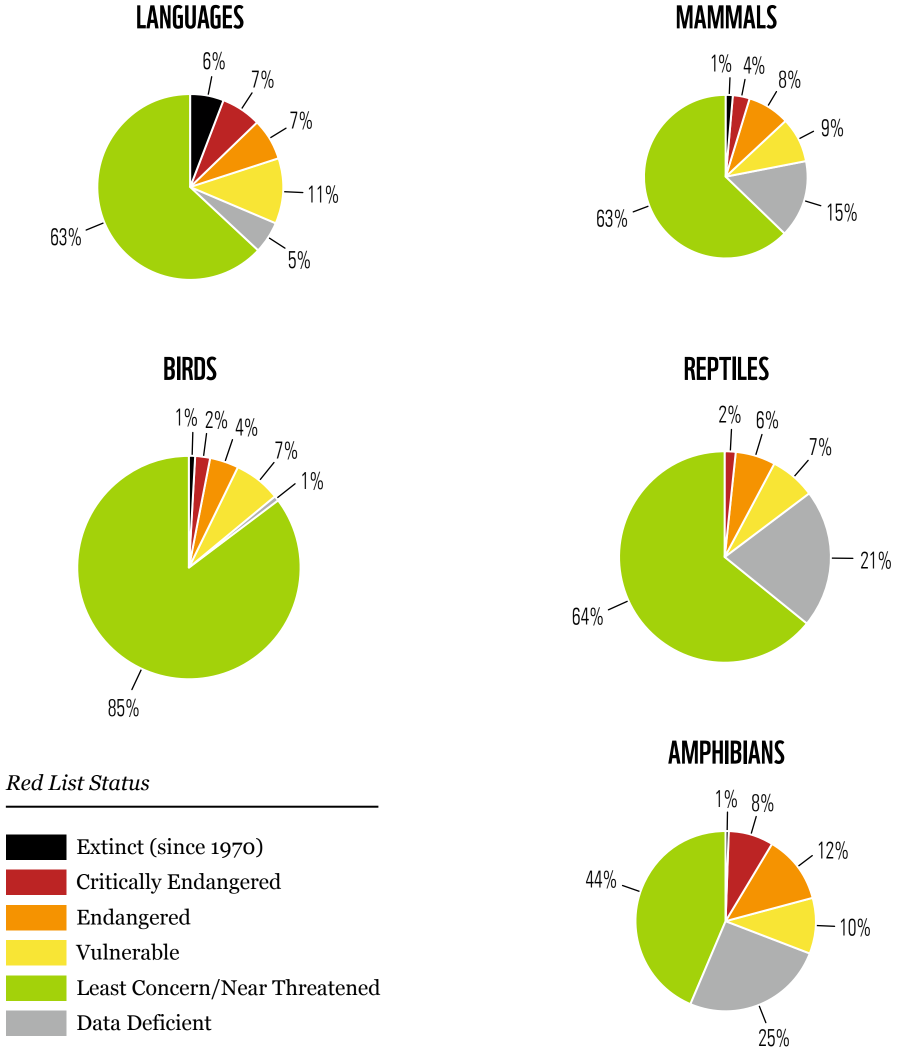
Most importantly for our present purposes, applying the IUCN Red List criteria to languages allows us to assess their threat status on the same basis as species, and make comparisons on a quantified basis. This has been done previously by the ecologist William Sutherland (2003), who used a limited set of the Red List criteria to compare languages with birds and mammals. Sutherland found that a higher percentage of languages was either threatened or recently extinct (32%) than either birds (13%) or mammals (28%). Here we compare the status of languages with that of mammals, birds, reptiles and amphibians, and compare the status of languages between regions of the world and between different language families, using selected Red List criteria. The data on numbers of speakers of languages come from editions of Ethnologue dating from 1951 to 2009, although most of the data come from the 1990s and 2000s (see Table 4).

Table 4: Number of data points on speaker numbers

Period	Data points
1900-1949	19
1950-1959	107
1960-1969	350
1970-1979	601
1980-1989	634
1990-1999	854
2000-2009	1008
Total	3573

Past data are sometimes unreliable, especially in the earlier decades, and therefore the first Red List criterion – rate of decline observed or projected over three generations – has not been used in this analysis. Trends in speaker numbers have only been used in combination with total number of speakers (the second criterion). Therefore the assessment of languages is very much more conservative than that of species groups. The results are shown in Figure 7.

Figure 7:* *Red List conservation status of languages and four vertebrate classes
Size of each pie is proportional to the number of languages or species in each group
Mammal, bird and amphibian data from IUCN (2013), reptile data from Bohm et al. *(2013).*



The analysis indicates that at least a quarter of the world's languages are threatened with extinction (CR, EN or VU), assuming that no data deficient (DD) language is threatened, compared with at least 21% of mammals, 13% of birds, 15% of reptiles and 30% of amphibians, the most threatened class of vertebrate. Furthermore, about 6% of languages have been reported as recently extinct, as opposed to about 1% of vertebrate species. If sufficient data on all the criteria used to evaluate animals were available to assess languages, then the status of languages could be worse than it appears here.

Ethnologue reports figures for the numbers of languages in each EGIDS category except for Extinct. They are Dormant 2.9%, Nearly Extinct 6.0%, Moribund 4.1%, Shifting 6.5%, Threatened 14.8%, Vigorous or better 65.7% (Lewis *et al.* 2013). If the EGIDS categories were translated into Red List categories as in Table 3, the percentages would be quite similar to those given in Figure 7.

Threat Status of Language Families and Regions

Just as the world's population is not evenly distributed among the world's languages, with half the world speaking one of just 24 languages, so the world's languages are not evenly distributed among language families. Figures 8a and 8b show the dominance of a few major language families such as Afro-Asiatic, Austronesian, Indo-European, Niger-Congo and Sino-Tibetan.

Figure 8a:* *The largest language families by number of languages *(source: Lewis* et al. *2013)*

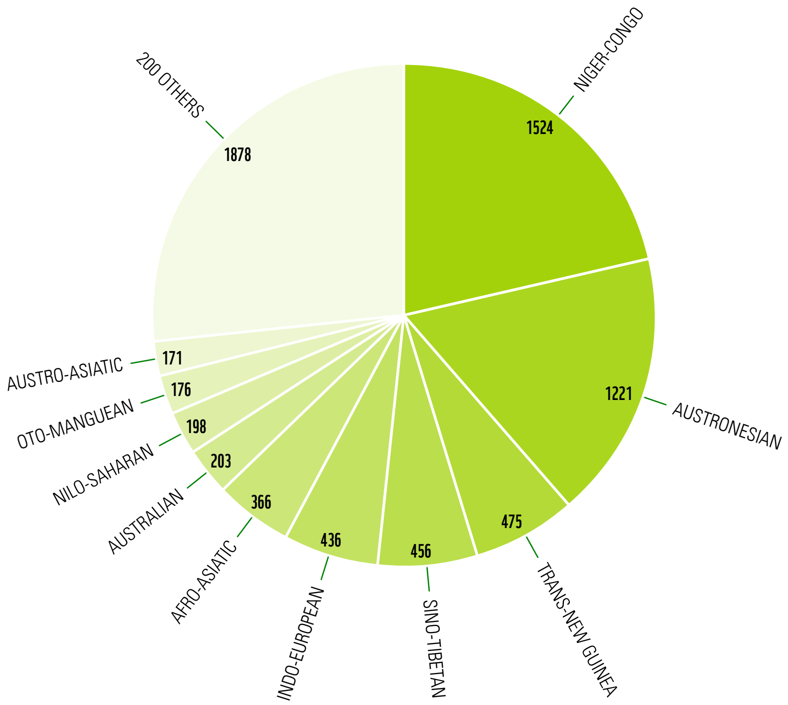
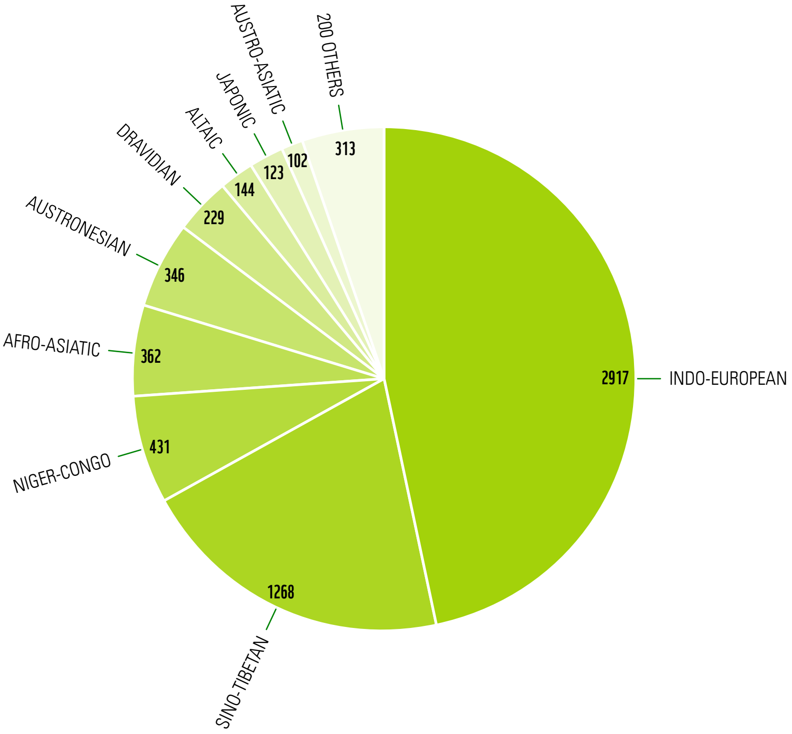


Figure 8b:* *The largest language families by number of speakers (millions) *(source: Lewis* et al. *2013)*



Status of Language Families

Some of the larger language families from different regions of the world have been assessed to compare their conservation status, in exactly the same way as all languages were assessed as a whole. Figure 9 shows the percentage of languages in each Red List category. The status of languages in each family can be compared by looking at the percentage of languages in the extinct (EX) and threatened categories (CR, EN, VU). Note that this assumes that all data deficient (DD) languages are not threatened, so it is a conservative estimate of threat status.

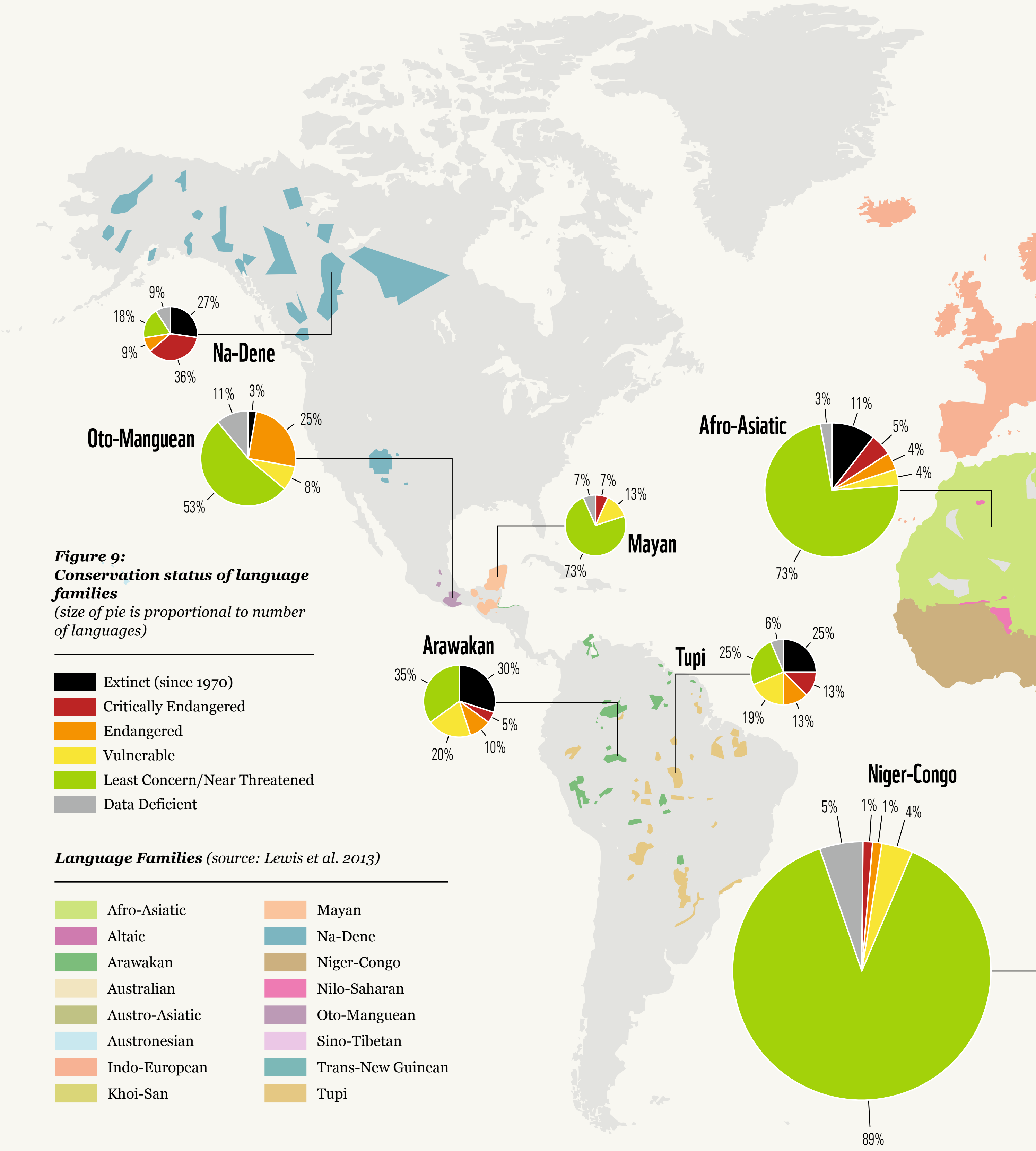


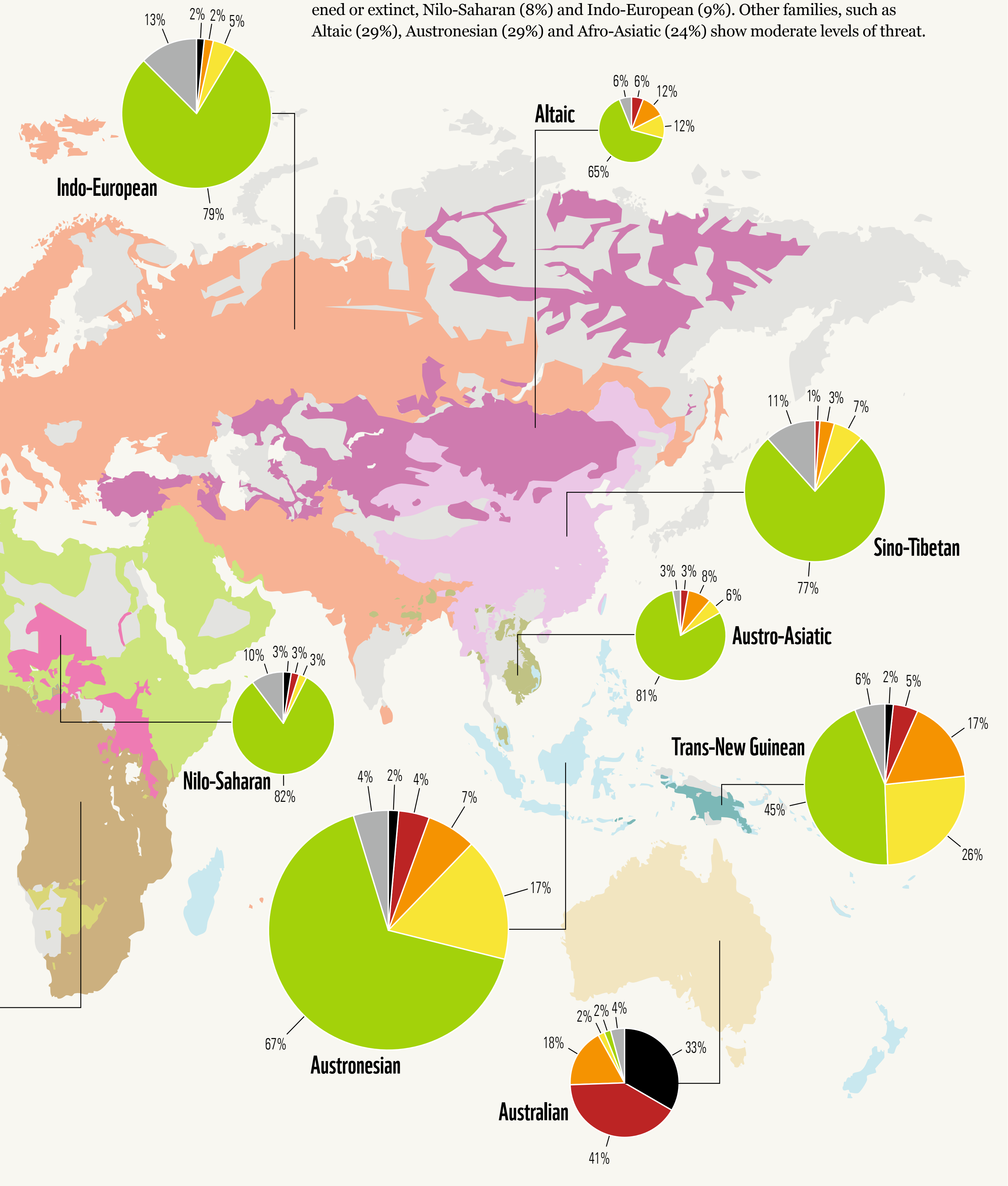
Figure 9:
Conservation status of language families
(size of pie is proportional to number of languages)

- ⬛ Extinct (since 1970)
- 🟥 Critically Endangered
- 🟧 Endangered
- 🟨 Vulnerable
- 🟩 Least Concern/Near Threatened
- ⬜ Data Deficient

Language Families *(source: Lewis et al. 2013)*

- Afro-Asiatic
- Altaic
- Arawakan
- Australian
- Austro-Asiatic
- Austronesian
- Indo-European
- Khoi-San
- Mayan
- Na-Dene
- Niger-Congo
- Nilo-Saharan
- Oto-Manguean
- Sino-Tibetan
- Trans-New Guinean
- Tupi

It is clear that languages in the Australian family are the most severely endangered, with 94% of languages threatened with extinction or extinct (since 1970), followed by some of the American language families such as Na-Dene (73%), Tupian (69%) and Arawakan (65%). The least endangered families are Niger-Congo, with only 6% of languages threatened or extinct, Nilo-Saharan (8%) and Indo-European (9%). Other families, such as Altaic (29%), Austronesian (29%) and Afro-Asiatic (24%) show moderate levels of threat.



Status of Regions

A clear pattern emerges if the data are analysed by region rather than family. Figure 10 show the percentage of languages in each Red List category. The status of regions can be compared by adding up the percentage of languages in the extinct (EX) and threatened (CR, EN, VU) categories. As with families, this assumes that all data deficient (DD) languages are not threatened, so it is a conservative estimate of threat status of each region.

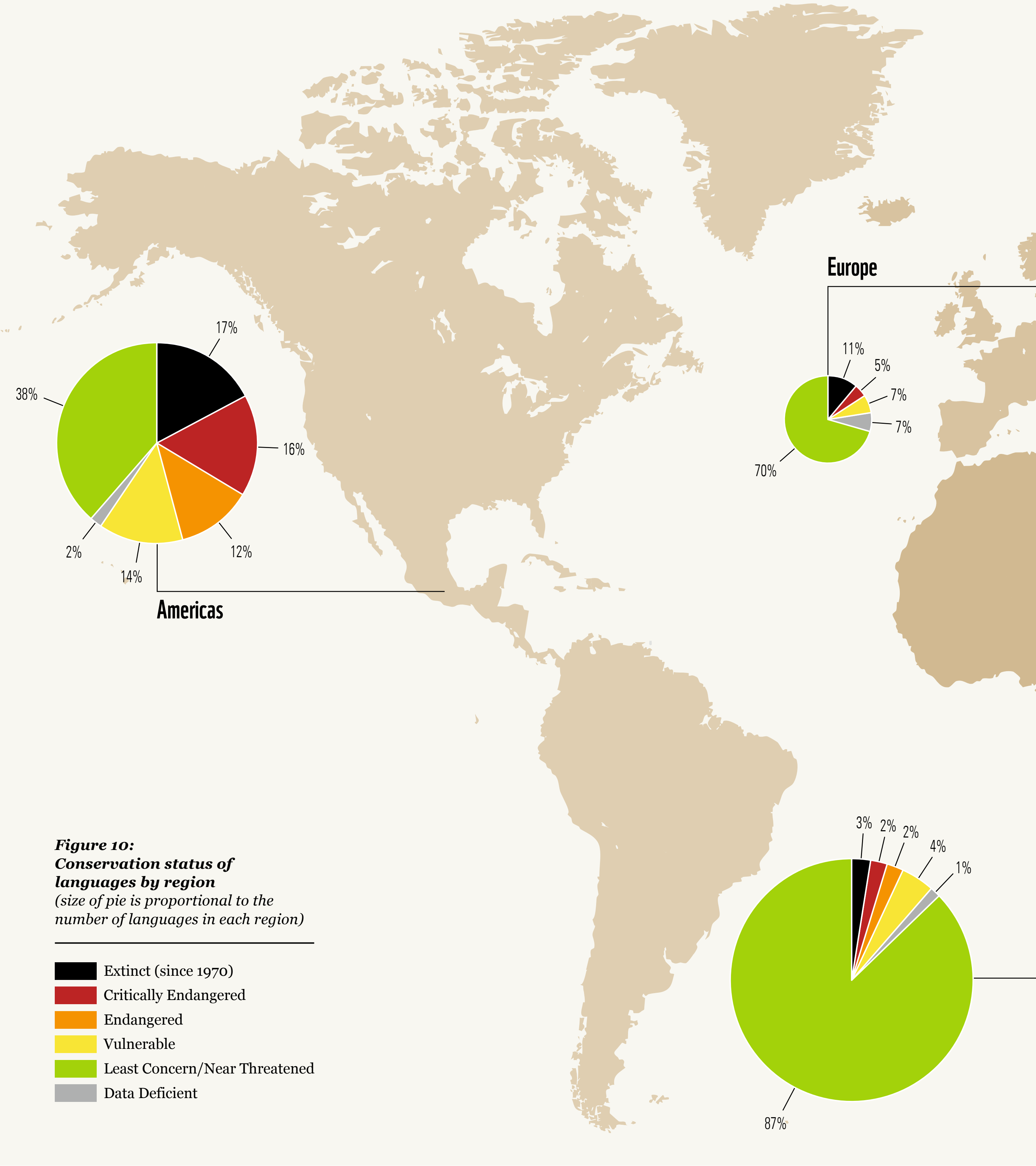
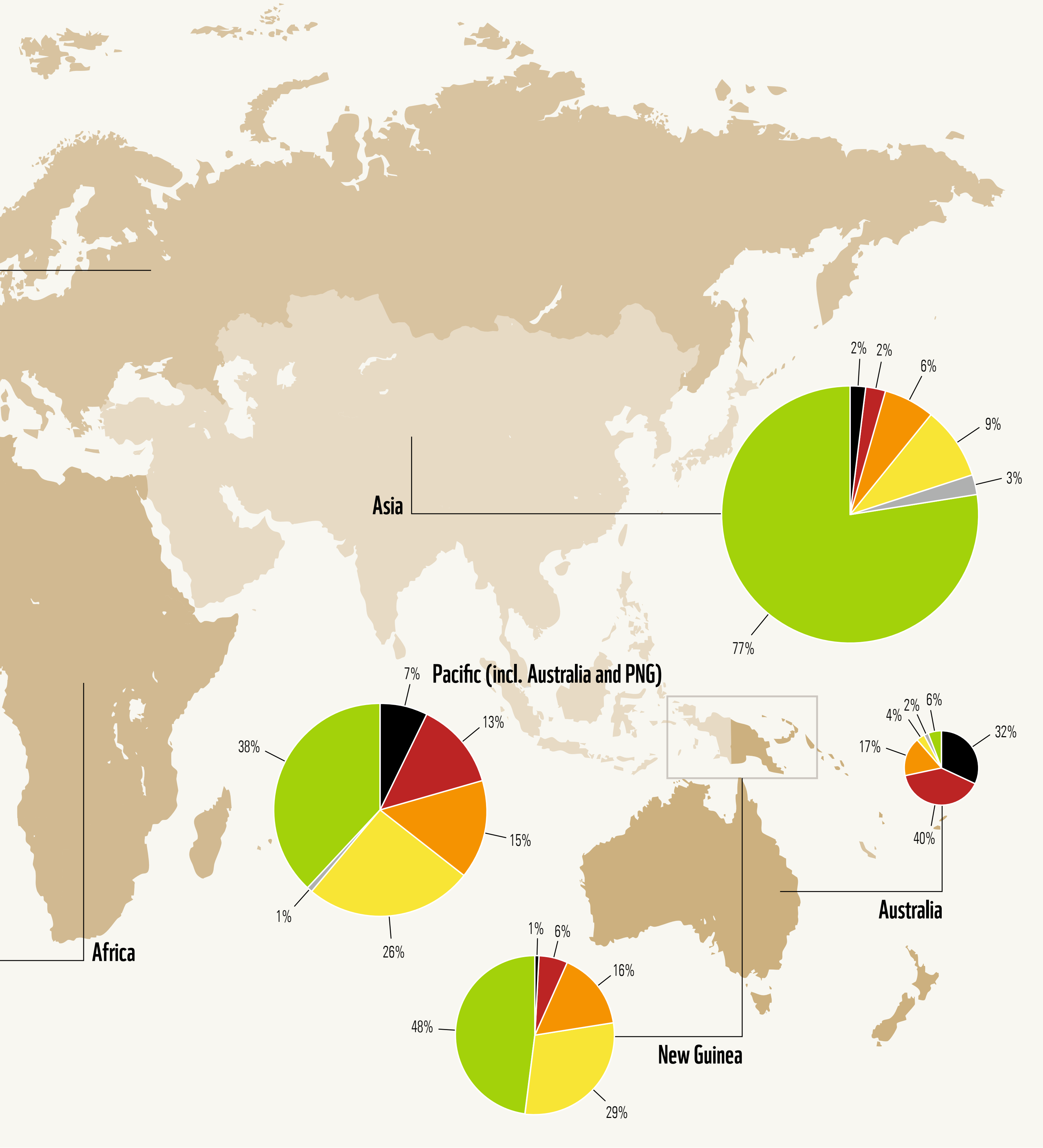


Figure 10:
Conservation status of languages by region
(size of pie is proportional to the number of languages in each region)

- ⬛ Extinct (since 1970)
- 🟥 Critically Endangered
- 🟧 Endangered
- 🟨 Vulnerable
- 🟩 Least Concern/Near Threatened
- ⬜ Data Deficient

Languages of the Pacific and the Americas are the most severely endangered, both regions with about 60% of their languages threatened with extinction or extinct (since 1970). The regions with the lowest level of extinction risk are Africa (11%) and Asia (20%). If Australia is separated from the rest of the Pacific region, it is once more apparent that Australia's languages are the most severely endangered in the world (92%).[15]



Trends in Languages and Species

Another way that biologists assess the state of biodiversity is to use indices based on average trends in the populations of a selection of species, such as the WWF/ZSL Living Planet Index (LPI). Species population indices are essentially like stock market indices such as the Dow Jones or Financial Times Stock Exchange (FTSE) which track trends in market capitalization of a number of companies. The LPI is based on time-series data for approximately 9,000 vertebrate species populations (of about 2,600 different species) from around the world. The index has been published biannually by WWF and ZSL since 1998 (WWF 2012).

In previous work, the authors adapted the LPI method to create an index called the Index of Linguistic Diversity (ILD) (Harmon and Loh 2010) which can be compared with trends in biodiversity as measured by the LPI. The ILD uses trends in the numbers of speakers across a sample of languages to calculate average trends. The same sample of 1,500 languages used in the Red List analysis of languages was also used to calculate the ILD. Data on numbers of mother tongue speakers for each language going back to 1900 were extracted from editions of Ethnologue. After removing all languages with data from only a single point in time (412 languages), and then filtering the data to remove anomalous data points or time series,[16] the remaining dataset contained time series for 985 languages. The ILD calculates the average trend of those languages in the sample. The ILD results for all 985 languages is a measure of trends in linguistic diversity for all languages in the world, and may be compared with the global LPI to see relative trends in linguistic diversity and biodiversity. To facilitate regional comparisons between the two indices, we also calculated the ILD by biogeographic realm to match up with the way the LPI is calculated regionally.

The biogeographic realms used in the analysis are the Afrotropical, the Indo-Pacific, the Nearctic, the Neotropical and the Palearctic realms. These are regions of the world defined according to the shared evolutionary history of their biota. It is a useful way to compare trends with languages, as language families tend to conform approximately to the same biogeographic patterns. Table 5 below shows which families belong to which biogeographic realms.

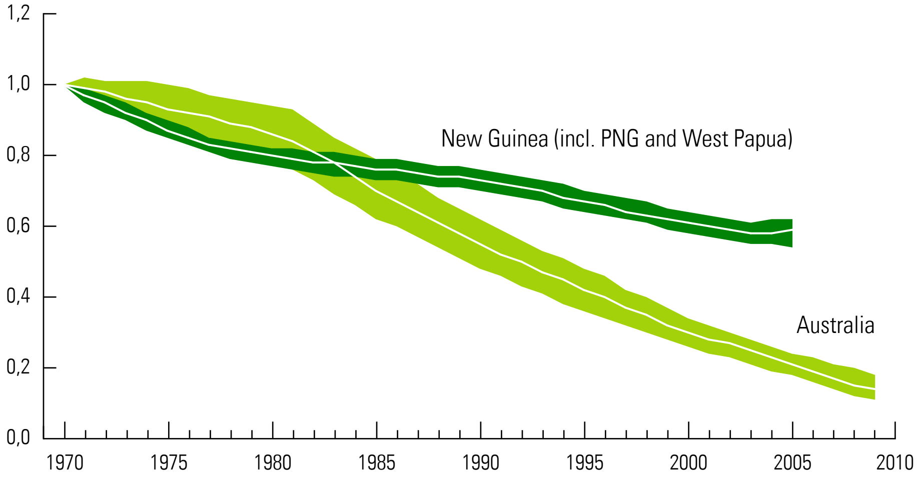
Table 5: Distribution of some language families among biogeographic realms

Biogeographic Realm	Location	Language families analyzed in this report
Afro-tropical	Sub-Saharan Africa	Niger-Congo, Nilo-Saharan
Indo-Pacific	South and Southeast Asia including southern China, Australasia and Oceania	Australian, Austro-Asiatic, Austronesian, Trans-New Guinea
Nearctic	North America and a part of northern Mexico	Na-Dene
Neotropical	Latin America and the Caribbean	Arawakan, Mayan, Oto-Manguean, Tupi
Palearctic	Eurasia, northern Africa and the Middle East, Central Asia, northern China and Japan	Afro-Asiatic, Altaic, Indo-European, Sino-Tibetan

The major difference between the LPI and ILD methodology is that the ILD has been corrected for overall human population growth. Within the period covered by the index, the human population has more than doubled, whereas there is no comparable overall growth in global wildlife populations. Therefore all of the ILD graphs presented here have been corrected for overall population growth. The biogeographic realm ILDs have been corrected for human population growth within the realm boundaries. The ILD therefore is not an index of population trends in quite the same way as the LPI; what the ILD measures is trends in the fraction of the total population belonging to each language.[17] To use another economic analogy, it is like an index of average market share of languages. If the average market share declines, it means that a few languages are increasing their market share at the expense of a greater number of others. This is exactly what we would expect to see if language shift is taking place: as speakers shift away from many small languages to fewer larger languages, then the average market share index falls.

The ILDs and LPIs for each biogeographic realm are shown in figure 12, plotted on the same axes for comparison, as are the global indices. Because the Indo-Pacific realm includes two islands which are particularly important in terms of linguistic diversity, Australia and New Guinea, additional ILDs for the two are shown separately in figure 11. Australia shows the fastest decline in linguistic diversity of any country, with a fall of about 85% in its ILD from 1970 to 2009. The ILD for New Guinea, the number one hotspot for linguistic diversity, which includes the Indonesian half of the island plus the half that is Papua New Guinea (PNG), shows a decline of about 40% between 1970 and 2005. This is a faster decline than the global average ILD, and reflects the Red List status of the island which shows that over 50% of New Guinea's 1000 or more languages are threatened.

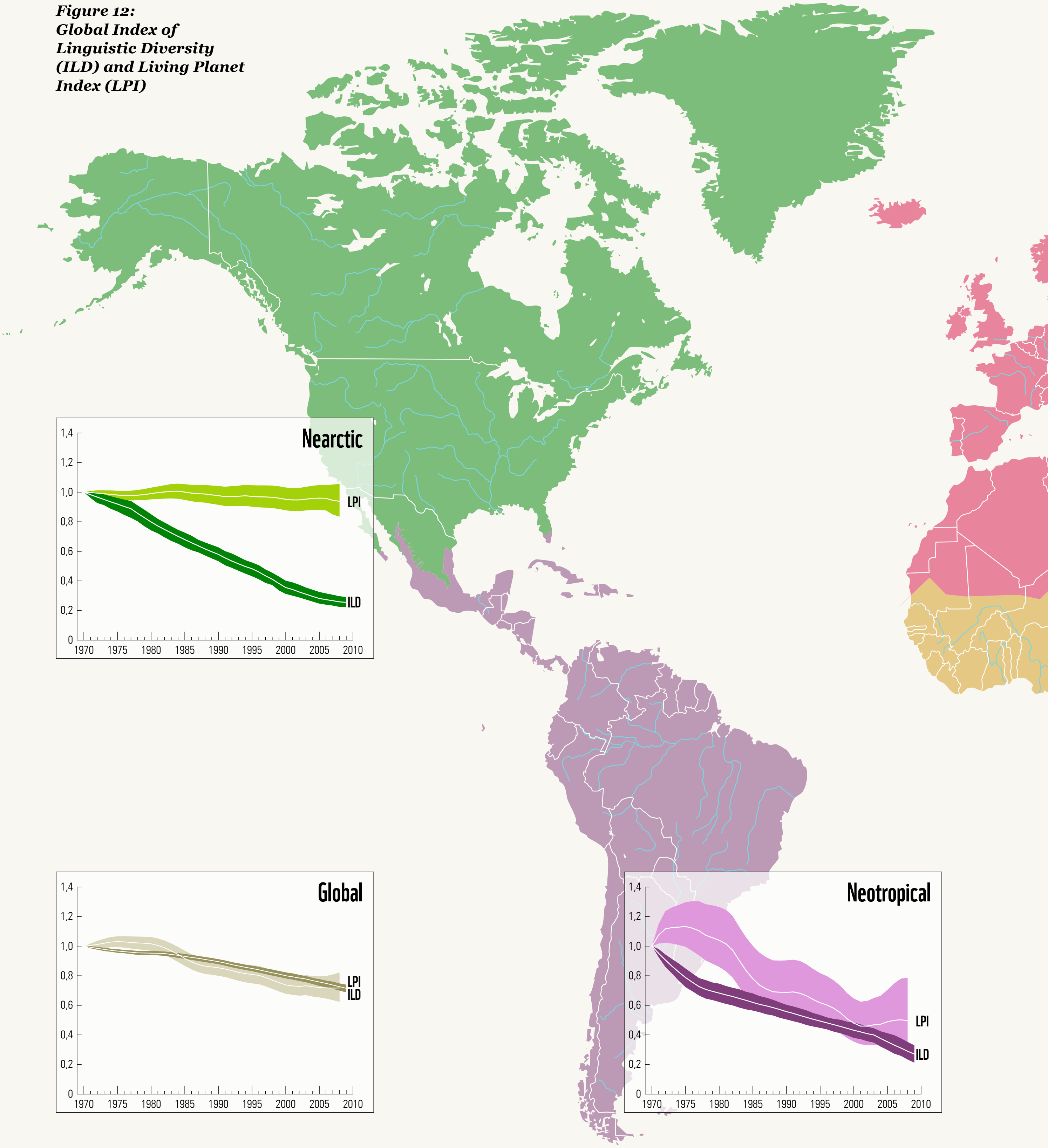
Figure 11:* *Index of Linguistic Diversity. Australia and New Guinea



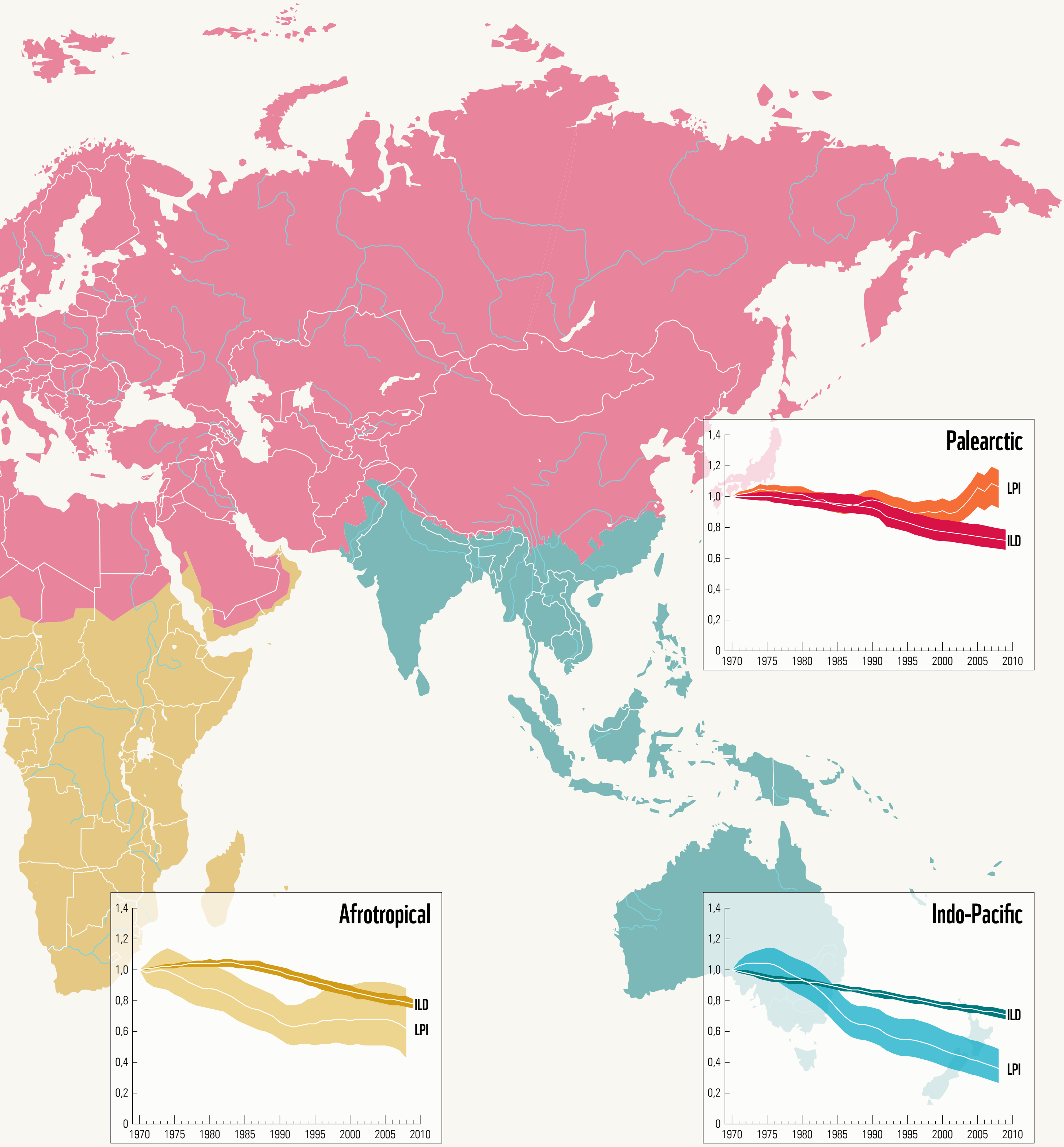
Trends in the LPI and ILD

Globally, the indices of both species populations (LPI) and speakers of languages (ILD) are declining at similar rates, about 30% in 40 years. The most rapid declines in species since 1970 have occurred in the Afrotropics (about 40%), Indo-Pacific (about 65%) and Neotropics (about 50%), whereas the Nearctic and Palearctic have shown little overall change. For languages, the most rapid declines since 1970 have taken

Figure 12: Global Index of Linguistic Diversity (ILD) and Living Planet Index (LPI)



place in the Nearctic and Neotropical realms (both about 75%), whereas the rate of decline in the Afrotropical (about 20%), Indo-Pacific (about 30%) and Palearctic (about 30%) realms has been slower. In summary, biodiversity has declined rapidly in the tropics, but remained steady in temperate realms; linguistic diversity on the other hand has declined rapidly in the new world, but more slowly in the old world.



CONCLUSIONS

It is striking that biodiversity and cultural diversity in general, and species and languages in particular, show extraordinary parallels both in terms of their evolution and the threats they face. Both species and languages have evolutionary histories which can be traced back in time to earlier, ancestral species and languages; both languages and species can be classified in such a way as to show the phylogenetic relationships between those related by descent from a common ancestor. Species may be defined by the ability to interbreed; language may be defined by mutual intelligibility.
By these definitions, subspecies are analogous to dialects. The process of formation of new species, speciation, has its linguistic equivalent. It may even be argued that the evolutionary mechanisms that give rise to both species diversity and linguistic diversity are similar. Biological evolution and, it can be argued, cultural evolution are the result of the action of replication, variation and selection working on hereditary material.
As well as the genetic relationships between languages, linguists talk about the ecology of a region or country, which has nothing to do with flora or fauna.
The global distribution of languages and species show remarkable similarity, with diversity highest in the tropics and declining toward the poles.

Two results are immediately apparent when comparing the status and trends in biodiversity and linguistic diversity. Firstly, at the global level, the trends are very similar, both the LPI (species) and ILD (languages) declined by about 30% since 1970, which suggests that biodiversity and linguistic diversity are being lost at similar rates. This supports the conclusion of the Red List analysis comparing the conservation status of languages and species: globally, linguistic diversity is at least as threatened as biodiversity.

THE DIFFERENCE IN REGIONAL TRENDS BETWEEN THE SPECIES AND LANGUAGES CAN BE EXPLAINED BY THE DIFFERENT DIRECT PRESSURES FACED BY BIODIVERSITY AND LINGUISTIC DIVERSITY

The second result is that, while both biodiversity and linguistic diversity are threatened globally, they are declining at different rates in different regions of the world. By far the most rapid losses in linguistic diversity have occurred in the Americas where, according to the Red List analysis, 60% of languages are threatened or have gone extinct since 1970. The ILD plummeted by over 75% between 1970 and 2009 in both the Nearctic and the Neotropical biogeographic realms. The LPI, however, shows that while species populations have fallen in the Neotropics (although with high uncertainty limits), they were almost completely flat in the Nearctic. The LPI fell by more than 60% in the Indo-Pacific, whereas the ILD declined by about 30%, a similar rate to the global average. Of course this masks the catastrophic decline of more than 80% in the ILD of Australia (and more than 40% in New Guinea). The ILDs for the Afrotropical and Palearctic realms both show declines of around 20-30%.

The difference in regional trends between the LPI and the ILD can be explained by the different direct pressures faced by biodiversity and linguistic diversity. Biodiversity decline is the usually the result of one of five main direct threats or pressures: habitat loss and destruction, direct over-exploitation of species from hunting and fishing, competition or predation by invasive alien species, climate change, or pollution. Habitat loss and over-exploitation of species remain the greatest threats for most of the world's biodiversity, and over the last 40 years the strongest pressure has been felt in the tropics, especially in Asia.

THE GROWING FOOTPRINT OF RICH COUNTIES AS WELL AS POPULATION GROWTH IN POOR COUNTRIES ARE DRIVING THE LOSS OF TROPICAL BIODIVERSITY IN AFRICA, ASIA AND LATIN AMERICA

In Europe and North America most of the biodiversity loss from habitat destruction occurred before 1970 and so does not register on the LPI. However, the footprint of natural resource consumption by the developed world is felt increasingly in the developing tropics, and all the more so as China's demand for natural resources grows; so it is the growing consumption by rich counties as well as population growth in poor countries that are driving the loss of tropical biodiversity in Africa, Asia and Latin America.

The decline in linguistic diversity is normally a result of the process of language shift away from small indigenous languages toward larger, national or regional languages. Language shift is driven by a number of social, political and economic factors including migration, urbanization, national unification, colonization, and the globalization of trade and communications. Migrant communities often undergo a process of language shift, whether moving from one country to another, or from a rural to an urban area within the same country. Governments in many developed and developing countries actively promote a single national language at the expense of other, usually minority, languages for political reasons. This has been the case with Mandarin in China, French in France and Amharic in Ethiopia for example. Migration, urbanization and political nationalization have been the primary drivers in Africa, Asia and Europe, where language shift has tended to occur between languages within the region.
In the Americas and the Pacific, especially Australia, the primary driver has also been migration, but the migrants, mainly European, vastly outnumbered the indigenous populations, and so it was the migrants' languages, primarily English, Spanish and Portuguese, that became politically and economically dominant. It is in these regions where indigenous languages are most highly threatened.

Australia and the island of New Guinea deserve particularly close attention: Australia because its indigenous languages are the most highly threatened in the world, and New Guinea because it is the most linguistically diverse place on Earth. Most of the 1,000 or so languages of New Guinea are threatened, but their decline is not as rapid as in Australia where more than 90% are threatened with extinction. The difference between the two islands is of course due to the fact that the vast majority of the Australian population is of European descent, whereas the population of New Guinea is largely indigenous. Australian languages are spoken by minority indigenous communities, and among these communities English is taking over, or has taken over, as the first language. In New Guinea indigenous languages are faring better, although the English-derived lingua-franca Tok Pisin is gaining ground at their expense.

MIGRATION, URBANIZATION AND POLITICAL NATIONALIZATION HAVE BEEN THE PRIMARY DRIVERS OF LANGUAGE LOSS IN AFRICA, ASIA AND EUROPE. IN THE AMERICAS AND AUSTRALIA, THE PRIMARY DRIVER HAS ALSO BEEN MIGRATION, BUT THERE THE MIGRANTS, MAINLY EUROPEAN, VASTLY OUTNUMBERED THE INDIGENOUS POPULATIONS.

A Bajau ("sea gypsy") woman and children, speakers of an Austronesian language, Kusungan Island, Sabah, Malaysia. Bajau people originated from the Philippines and traditionally lived on boats, making their living from the sea, but most are now settled.

© WWF-Malaysia / Mazidi Abd Ghani

EPILOGUE

A future for biocultural diversity?

Ultimately, both linguistic diversity and biodiversity are diminishing as a result of human population growth, increasing consumption, and globalization which erodes differences between one part of the world and another. At the regional level, these fundamental drivers of diversity loss are manifest in different ways. For biodiversity, the biggest threat in modern times has been and still is habitat destruction, followed by over-exploitation (fishing and hunting) and invasive species. Since 1970, habitat loss has been most rapid in the developing world, particularly in Asia, whereas habitat loss in Europe and North America has slowed down and levelled off. Consequently the most rapid decline in biodiversity is now happening in the tropics, the part of the world with the greatest diversity.

LINGUISTIC DIVERSITY AND BIODIVERSITY ARE DIMINISHING AS A RESULT OF HUMAN POPULATION GROWTH, INCREASING RESOURCE AND ENERGY CONSUMPTION, AND ECONOMIC GLOBALIZATION WHICH ERODES DIFFERENCES BETWEEN ONE PART OF THE WORLD AND ANOTHER

For languages and culture, 'habitat' means the human population, which has doubled since 1970, so habitat loss is not the problem. Nor does direct 'consumption' threaten culture (there is no equivalent of cultural over-consumption). It is the cultural analogue to alien invasive species – language shift – that is the greatest threat to linguistic and cultural diversity. It is not that one human population replaces another population, as is the case with invasive species, it is that one language displaces another language within the same population. When an indigenous language goes extinct, often the indigenous culture follows. This process has been happening for the last two hundred years or more in the linguistic ecology of the Americas, Australia and parts of the Pacific, where indigenous languages have been severely threatened by the dominance of European languages, particularly English, Spanish and Portuguese. In Africa, Asia and Europe, where the main drivers of language shift have been migration, urbanization and political unification policies, language shift has tended to occur between languages of those regions, and diversity is being lost, but not as rapidly.

Why do we need so much diversity? Would it not be better for the sake of world peace and the global economy, it is sometimes argued, if we spoke fewer languages in the world? Are languages or cultural diversity really worth conserving as much as species or biological diversity? The logical conclusion of this type of argument is that, ideally, we should speak just one world language. But then we would all become more similar, and the differences between one part of the world and another, or between one culture and another, would rapidly erode away. In the end, we would speak the same language, wear the same clothes, eat the same food, listen to the same music, consume the same brands and hold the same beliefs. One city would look much the same as another. The world would become homogenized. This counter-argument may sound absurd, but it is already happening: the world is already losing its extraordinary biocultural diversity, as the findings of this report demonstrate. No doubt the global economy would continue to grow just as well, or even better, with just a few world languages and cultures. It is even possible that global ecosystems could continue to provide basic life support functions – although probably not as well – with less biodiversity, and humanity would still survive. But this is not just a question of survival, or even global economic productivity. A diverse world is a culturally and naturally richer world. With less diversity, humanity is poorer. It is a question of the kind of world we want to live in.

The science of biocultural diversity is in its infancy, and more research is needed to examine and understand the processes and mechanisms that underpin and unite biological and cultural evolution and ecology. Most importantly, we need a better understanding of how to slow down and reverse the loss of diversity. While the outlook is not bright for many of the world's smaller languages, especially those no longer being learned by children, there is plenty of scope to improve, develop and promote biocultural conservation. There is an opportunity for biodiversity conservation and the conservation of indigenous languages and cultures to go hand in hand. Most of the world's linguistic diversity is found in areas of high species richness and endemism. If biodiversity conservation organizations on the ground in areas of high biocultural diversity were to invest resources in the conservation of indigenous languages and traditional knowledge there would be a double pay-off. Field linguists working on indigenous languages often lack the ecological knowledge needed in order to understand and translate the vast lexicon of terms for species and natural phenomena. Field biologists could benefit from the immense wealth of traditional ecological knowledge of indigenous people. Not only would biological and cultural diversity be conserved together in the environment in which they both evolved, so protecting the full range of living biocultural diversity, but also the traditional resource management systems, a fundamental component of the cultural identities that are now in retreat, could be applied to conserving the landscape, its component species and its languages.

But it is not only the rarest languages and species that we should conserve. Relatively common languages, spoken by tens of thousands of people, and common species are in decline too. Maintaining diversity is not just a question of protecting endangered languages and species in remote hotspots of biocultural diversity such as the Amazon or New Guinea, vitally important though that is, conservation is also a matter of allowing diversity to thrive in those parts of the world where humans have already had a profound impact on the biological and cultural landscape, in the more densely populated parts of the planet. Recognizing and exploring the parallels between nature and culture, and understanding the processes that underlie their evolution, ecology and extinction, is a first step towards ensuring that we can continue to inhabit a world of incredible diversity.

Mongolian herder, speaker of an Altaic language, Baga Lake, Khar Us Nuur National Park, Mongolia.

© Frans Schepers / WWF-Netherlands

END NOTES

[1] It is possible for a person to speak more than one language, of course, whereas an individual animal or plant cannot belong to more than one species. In this report, we refer to mother-tongue speakers of a language only. This means one's native language, or the language one most strongly identifies with as a native speaker, which is usually, although not necessarily, the language one learned first. When we report the numbers of speakers of a language, we mean mother-tongue speakers, so each speaker is only counted once.

[2] There are actually quite a number of different definitions of "species" and "languages," and multiple processes of speciation and language genesis. For details, see Harmon 2002.

[3] This description of the evolution of biocultural diversity builds on JL's contribution "The third flowering of the Tree of Life" to Jorgen Randers, 2052 - A Global Forecast for the Next Forty Years (2012).

[4] Ours is not the only species to have culture – many species of songbird, for example, show regional and local variations in their repertoire of songs which are not inherited genetically but learned from other individuals of the same species, and some species of primate even have a limited repertoire of calls with specific meanings such as "snake" or "leopard" – but we are the only species to have language.

[5] French belongs to the Italic branch, which is also part of the Indo-European family.

[6] There is still controversy surrounding the exact dates and routes of the human diaspora out of Africa. The description here is based on Oppenheimer (2004).

[7] It is likely that proto-language, consisting of sounds, gestures and expressions, had begun evolving long before that time.

[8] Strictly speaking, this is West Frisian, as there are two other Frisian languages spoken in northern Germany.

[9] Basque is a language isolate (in a family of its own) in northern Spain and southwest France, while Navajo is a language in the indigenous Na-Dene family spoken in the southwest United States.

[10] Arabic is not included in the top ten as it is classified as many different languages, such as Algerian spoken Arabic, Egyptian spoken Arabic, etc. If all speakers of Arabic languages were counted together, Arabic would appear in the top ten.

[11] Scottish Gaelic, a Celtic language related to Irish, should not be confused with Scots, a Germanic language related to English.

[12] Native Scots Gaelic speakers and bilingual English-Gaelic speakers are counted together in this example.

[13] Cornish, a Celtic language related to Welsh and Breton, spoken in Cornwall.

[14] Manx, a Celtic language related to Irish and Scottish Gaelic, spoken on the Isle of Man.

[15] A few of Australia's languages are not in the Australian language family, hence the difference between the percentage of Australia's languages that are extinct or threatened with extinction and that for Australian languages.

[16] Languages with 1,000 or more speakers which grew or declined at a rate greater than 10% per year.

[17] For more detailed discussion of the ILD, see Harmon and Loh (2010).

DATA TABLES

Table 6: Conservation status of languages and species
Data for mammal, bird and amphibian species from IUCN 2013, data for reptiles from Bohm et al. *2013, percentage in each Red List category.*

Language or Species Group	EX	CR	EN	VU	LC	DD	Total assessed
Languages (sample of 1500)	6%	7%	7%	11%	63%	5%	1.500
Mammals (all)	1%	4%	8%	9%	63%	15%	5.506
Birds (all)	1%	2%	4%	7%	85%	1%	10.065
Reptiles (sample of 1500)	0%	2%	6%	7%	64%	21%	1.500
Amphibians (all)	1%	8%	12%	10%	44%	25%	6.409

Table 7: Conservation status of selected language families
Percentage of languages in each Red List category.

Language Family	EX	CR	EN	VU	LC	DD	Total assessed
Afro-Asiatic	11%	5%	4%	4%	73%	3%	75
Altaic	0%	6%	12%	12%	65%	6%	17
Australian	33%	41%	18%	2%	2%	4%	51
Austro-Asiatic	0%	3%	8%	6%	81%	3%	36
Austronesian	2%	4%	7%	17%	67%	4%	245
Indo-European	2%	0%	2%	5%	79%	13%	104
Niger-Congo	0%	1%	1%	4%	89%	5%	296
Nilo-Saharan	3%	3%	0%	3%	82%	10%	39
Sino-Tibetan	0%	1%	3%	7%	77%	11%	87
Trans-New Guinean	2%	5%	17%	26%	45%	6%	119
Arawakan	30%	5%	10%	20%	35%	0%	20
Mayan	0%	7%	0%	13%	73%	7%	15
Na-Dene	27%	36%	9%	0%	18%	9%	11
Oto-Manguean	3%	0%	25%	8%	53%	11%	36
Tupi	25%	13%	13%	19%	25%	6%	16

Table 8: Conservation status of languages by region
Percentage of languages in each Red List category.

Region	EX	CR	EN	VU	LC	DD	Total assessed
Africa	3%	2%	2%	4%	87%	1%	402
Americas	17%	16%	12%	14%	38%	2%	255
Asia	2%	2%	6%	9%	77%	3%	464
Europe	11%	5%	0%	7%	70%	7%	44
Pacific (incl. Australia and PNG)	7%	13%	15%	26%	38%	1%	282
Australia	32%	40%	17%	4%	6%	2%	53
New Guinea (incl. PNG and West Papua)	1%	6%	16%	29%	48%	0%	220

Table 9: Index of Linguistic Diversity global and by biogeographic realm
Percentage decline 1970-2009.

Biogeographic Realm	Index	Lower confidence limit	Upper confidence limit
Global	28%	31%	26%
Afrotropical realm	22%	25%	19%
Indo-Pacific realm	29%	32%	26%
Nearctic realm	74%	78%	71%
Neotropical realm	73%	79%	67%
Palearctic realm	28%	34%	21%
Australia	86%	89%	82%
New Guinea (1970-2005)	41%	46%	38%

REFERENCES

Böhm, M., B. Collen, J.E.M. Baillie, *et al*. 2013. The conservation status of the world's reptiles. *Biological Conservation*, **157**, 372–385

Cavalli-Sforza, L.L. 2000. *Genes, People and Languages*. Penguin, London.

Collen, B., J. Loh, S. Holbrook, L. McRae, R. Amin & J.E.M. Baillie. 2009. Monitoring Change in Vertebrate Abundance. *Conservation Biology*, **23** (2), 317-327.

Darwin, C. 1874. *The Descent of Man, and Selection in Relation to Sex. Second edition*. John Murray, London. Downloaded 12 February 2014 from van Wyhe, J. (ed.) 2002. The Complete Work of Charles Darwin Online http://darwin-online.org.uk

Dawkins, R. 2006. *The Selfish Gene*. 30th Anniversary Edition. Oxford University Press, Oxford.

Dennett, D. 2002. The New Replicators. In: M. Pagel (ed.-in-chief) *Encyclopedia of Evolution*, 1, E83-92. Oxford University Press, Oxford.

Gray, R.D. & Q.D. Atkinson. 2003. Language-tree divergence times support the Anatolian theory of Indo-European origin. *Nature*, **426**, 435-439.

Gray, R.D., Q.D. Atkinson & S.J. Greenhill. 2011. Language evolution and human history: what a difference a date makes. *Phil. Trans. R. Soc. B*, **366**, 1090-1100.

Greenberg, J. 2000. *The Eurasiatic Language Family: Indo-European and its Closest Relations. Volume 1: Grammar*. Stanford Univ. Press, Stanford, California.

Groombridge, B. & Jenkins, M.D. 2002. World Atlas of Biodiversity. Prepared by UNEP World Conservation Monitoring Centre. Univ. California Press, Berkeley.

Harmon, D. 1995. The Status of the World's Languages as Reported in Ethnologue. *Southwest Journal of Linguistics*, **14** (1-2), 1-28.

Harmon, D. 1996. Losing Species, Losing Languages: Connections between Biological and Linguistic Diversity. *Southwest Journal of Linguistics*, **15** (1-2), 89-108.

Harmon, D. 2002. *In Light of Our Differences: how diversity in nature and culture makes us human*. Smithsonian Institution Press, Washington and London.

Harmon, D. & J. Loh. 2010. The Index of Linguistic Diversity: A New Quantitative Measure of Trends in the Status of the World's Languages. *Language Documentation and Conservation*, **4**, 97-151. Available online http://nflrc.hawaii.edu/ldc/2010/

Hrozný, B. 1917. *Die Sprache der Hethiter: ihr Bau und ihre Zugehörigkeit zum indogermanischen Sprachstamm: ein Entzifferungsversuch* (The Language of the Hittites: its Structure and its Membership in the Indo-European Linguistic Family). J.C. Hinrichs, Leipzig, Germany.

Huffman, S. no date. *Mapping the Genetic Relationships of the World's Languages*. http://www.worldgeodatasets.com/language/huffman/ Downloaded 22 February 2012.

IUCN 2013. *The IUCN Red List of Threatened Species*. Version 2013.2. http://www.iucnredlist.org. Downloaded 25 February 2014.

Jones, W. 1824. *Discourses delivered before the Asiatic Society: and miscellaneous papers, on the religion, poetry, literature, etc., of the nations of India*. C. S. Arnold. p.28.

Lewis, M. P. & G. F. Simons. 2010. Assessing endangerment: Expanding Fishman's GIDS. *Revue Roumaine de Linguistique,* **55** (2), 103-120

Lewis, M.P., G.F. Simons & C.D. Fennig (eds.). 2013. *Ethnologue: Languages of the World, Seventeenth edition*. SIL International, Dallas, Texas. Online version: http://www.ethnologue.com.

Loh, J., R.E. Green, T. Ricketts, J. Lamoreux, M. Jenkins, V. Kapos & J. Randers. 2005. The Living Planet Index: using species population time series to track trends in biodiversity. *Phil. Trans. R. Soc. B*, **360**, 289-295.

Loh, J. & D. Harmon. 2005. A global index of biocultural diversity. *Ecological Indicators*, **5**, 231-241.

MacAulay, D. 1992. *The Celtic Languages*. Cambridge University Press, Cambridge.

Mace, R. & M. Pagel. 1995. A latitudinal gradient in the density of human languages in North America. *Proc. R. Soc. Lond. B*, **261**, 117-121.

Moore, J.L., L. Manne, T. Brooks, N.D. Burgess, R. Davies, C. Rahbek, P. Williams & A. Balmford. 2002. The distribution of cultural and biological diversity in Africa. *Proc. R. Soc. Lond. B*, **269**, 1645-1653.

Nettle, D. 1999. *Linguistic Diversity*. Oxford University Press, Oxford.

Nettle, D. & S. Romaine. 2000. *Vanishing Voices: The Extinction of the World's Languages*. Oxford University Press, Oxford.

OED Online. Oxford University Press, March 2014.

Oppenheimer, S. 2004. *Out of Africa's Eden. The peopling of the world*. Jonathan Ball, Jeppestown.

Pagel, M., Q.D. Atkinson, A.S. Calude & A. Meade. 2013. Ultraconserved words point to deep language ancestry across Eurasia. *Proc. Nat. Ac. Sci.*, **110** (21), 8471-8476.

Pagel, M. & R. Mace. 2004. The cultural wealth of nations. *Nature*, **428**, 275-278.

Pennisi, E. 2004. The First Language? Science, 303, 1319-1320.

Raghavan, M., P. Skoglund, K.E. Graf, M. Metspalu, A. Albrechtsen, I. Moltke, S. Rasmussen, T.W. Stafford Jr, L. Orlando, E. Metspalu, M. Karmin, K. Tambets, S. Rootsi, R. Mägi, P.F. Campos, E. Balanovska, O. Balanovsky, E. Khusnutdinova, S. Litvinov, L.P. Osipova, S.A. Fedorova, M.I. Voevoda, M. DeGiorgio, T. Sicheritz-Ponten, S. Brunak, S. Demeshchenko, T. Kivisild, R. Villems, R. Nielsen, M. Jakobsson & E. Willerslev. 2014. Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. *Nature*, **505**, 87–91.

Randers, J. 2012. *2052 - A Global Forecast for the Next Forty Years*. Chelsea Green, Vermont.

Renfrew, C. 1987. *Archaeology and Language: the Puzzle of Indo-European Origins*. Cape, London.

Stepp, J.R. *et al.* 2004. Development of a GIS for Global Biocultural Diversity. *Policy Matters*, **13**, 267-171.

Sutherland, W.J. 2003. Parallel extinction risk and global distribution of languages and species. *Nature*, **423**, 276-279.

UNESCO. 2010. *Atlas of the World's Languages in Danger. Third edition.* (C. Mosely, ed.-in-chief). 11-12.

Walker, G. 2003. *Snowball Earth*. Bloomsbury, London.

WWF. 2012. *Living Planet Report 2012* (M. Grooten, ed.-in-chief). WWF International, Gland, Switzerland.

© André Bärtschi / WWF-Canon

Cashinahua girl, a speaker of a Panoan language with around 1,000 speakers, her face painted with dye from huito fruit. Near the Alto Purus Reserved Zone, Ucayali, Peru.

BIOCULTURAL DIVERSITY · THREATENED SPECIES, ENDANGERED LANGUAGES

WWF

Why we are here
To stop the degradation of the planet’s natural environment and to build a future in which humans live in harmony with nature.

www.wnf.nl