

The distribution of cultural and biological diversity in Africa

Joslin L. Moore^{1,2*}, Lisa Manne³, Thomas Brooks^{1,4}, Neil D. Burgess^{1,5}, Robert Davies⁶, Carsten Rahbek², Paul Williams³ and Andrew Balmford¹

¹Conservation Biology Group, Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK

²Zoological Museum, University of Copenhagen, Universitetsparken 15, DK-2100, Copenhagen Ø, Denmark

³Biogeography and Conservation Laboratory, The Natural History Museum, Cromwell Road, London SW7 5BD, UK

⁴Center for Applied Biodiversity Science, Conservation International, 1919 M Street NW, Suite 600, Washington, DC 20036, USA

⁵Conservation Science Programme, World Wildlife Fund, 1250 24th Street NW, Washington, DC 20037-1193, USA

⁶Llanunwas, Upper Solva, near Haverfordwest, Pembrokeshire SA62 6UJ, UK

Anthropologists, biologists and linguists have all noted an apparent coincidence in species diversity and human cultural or linguistic diversity. We present, to our knowledge, one of the first quantitative descriptions of this coincidence and show that, for $2^\circ \times 2^\circ$ grid cells across sub-Saharan Africa, cultural diversity and vertebrate species diversity exhibit marked similarities in their overall distribution. In addition, we show that 71% of the observed variation in species richness and 36% in language richness can be explained on the basis of environmental factors, suggesting that similar factors, especially those associated with rainfall and productivity, affect the distributions of both species and languages. Nevertheless, the form of the relationships between species richness and language richness and environmental factors differs, and it is unlikely that comparable mechanisms underpin the similar patterns of species and language richness. Moreover, the fact that the environmental factors considered here explain less than half of the variation in language richness indicates that other factors, many of which are likely to be historical or social, also influence the distribution of languages.

Keywords: species richness; language richness; biogeography; Africa; vertebrates; humans

1. INTRODUCTION

Species are not evenly distributed. Species richness tends to be higher at low latitudes, in tropical regions, in less seasonal climates and in areas of habitat or topographical heterogeneity (Pianka 1966; Currie 1991; Rohde 1999). Similarly, the density of different human cultures also differs markedly from place to place. Recent studies suggest that language density increases closer to the equator (Nichols 1990; Mace & Pagel 1995). Nichols (1992, 1997) has also noted that, globally, linguistic diversity at a family or stock level is higher in coastal regions, at lower latitudes and in wetter and less seasonal climates. The similarity of these patterns suggests that cultural and biological diversity may have similar geographical distributions (Mace & Pagel 1995; Nettle 1998; Oviedo *et al.* 2000; Pagel 2000). Here we provide a quantitative test of this hypothesis for sub-Saharan Africa.

Few previous studies have directly compared cultural and biological diversity. Pagel (2000) showed that indigenous North American languages have a similar latitudinal gradient to that of mammals. Oviedo *et al.* (2000) mapped the distribution of the world's ethnolinguistic groups (based on the ethnologue data; Grimes 1996) onto Global 200 ecoregions (areas identified by the World

Wildlife Fund USA as having the highest conservation priority due to a combination of high species richness and high endemism; Olson & Dinerstein 1998). They noted that 67% of the ethnolinguistic groups are in one of the Global 200 ecoregions, reflecting the concentration of languages and species in tropical regions. Similarly, Smith (2001) has looked for correlations between biological diversity and cultural diversity across 10 native North American 'culture areas'. Results were equivocal; he found good correlations between tree richness and three aspects of cultural diversity but very poor correlations between plant, bird or mammal richness and cultural diversity. However, the small sample size and lack of control for area effects make these results difficult to interpret.

In this study, we focus on language richness and vertebrate species richness as our measures of diversity. The difficulty of defining culture and consistently measuring cultural diversity (Smith 2001), especially over large geographical areas, means that discussion of cultural diversity tends to focus on linguistic diversity or language richness as proxies (Mace & Pagel 1995; Nettle 1996; Oviedo *et al.* 2000). Clearly, language richness is insufficient to describe all cultural diversity. Numerous cases exist where the distinctiveness of language does not correspond to the distinctiveness of culture and/or ethnicity (Oviedo *et al.* 2000). In addition, it can often be difficult to differentiate a language from a 'family of languages' or a dialect (Grimes 1996). Similarly, vertebrate species richness does

* Author for correspondence (jlm18@cam.ac.uk).

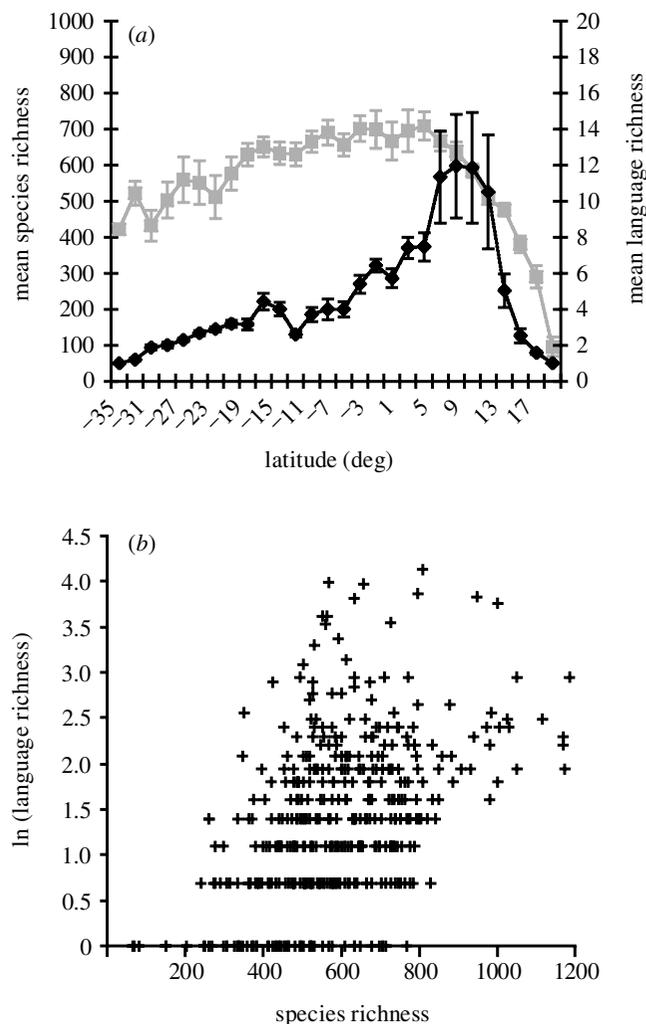


Figure 1. (a) Latitudinal trends in species richness (grey squares) and language richness (black diamonds) for sub-Saharan Africa. Each point represents mean richness per cell for all 2° cells with the central coordinate at the stated latitude. Note that the number of cells varied with latitude. Error bars are ± 1 s.e. (b) The correlation between species richness and language richness for the 421 2° grid cells containing language data.

not represent all biological diversity. Vertebrates comprise only 3% of named species (Groombridge & Jenkins 2000); however, comprehensive distribution data for other taxa are unavailable for sub-Saharan Africa. In addition, species richness does not of course represent all aspects of diversity (Magurran 1988). Nevertheless, language richness and vertebrate species richness do represent important and quantifiable aspects of cultural and biological diversity.

Another important attribute of distributions are patterns of endemism. A species (or language) is endemic if it is unique to a specific area or region. Clearly, simply counting the number of endemic species or languages per region or area will produce a measure of endemism that depends strongly on the size and positioning of the areas or regions considered. To circumvent this difficulty, we measure endemism as the number of the 25% most range-restricted species or languages in each grid cell.

(a) *Why would we expect a correlation between cultural and biological diversity?*

Although many have posited a correspondence between cultural and biological diversity (Mace & Pagel 1995; Harmon & Maffi 2002), much less has been written about the possible causes of such a relation. One simple explanation is that factors thought to be important in the generation, or maintenance, of species richness may also be important for the generation, or maintenance, of language richness. We outline several such factors below.

One obvious factor is the total availability of resources or energy (measured as temperature, potential evapotranspiration, actual evapotranspiration or net primary productivity) which, by limiting total biomass production, may limit the number of viable populations and hence the number of species that can be reasonably expected to persist (Currie 1991). A similar rationale can also be applied to language richness (Smith 2001). In addition, it has been suggested that energy (measured as temperature or potential evapotranspiration) may also have a direct effect on the rate of biological evolution through increased mutation rates and shortened generation times in areas with high energy (Rohde 1999); clearly this mechanism cannot apply to human language richness.

Climatic variability on a variety of temporal and spatial scales may be important for both species (Fjeldså & Lovett 1997; Fjeldså *et al.* 1997) and languages (Nettle 1996, 1998). Environmental stability, predictability or low seasonality could all facilitate specialization through decreasing the range of conditions that a species need tolerate, as well as providing refuges in times of climatic change. Similarly, one could expect that the size of a group's social network (and hence need for shared language) will decrease with food security, which one would expect to be greatest in regions with long growing seasons and climatic predictability (Nettle 1996, 1998).

Factors that increase the degree of (or opportunities for) isolation over time, thereby enhancing the chance of allopatric speciation, such as large areas (Rosenzweig 1995) or topographic heterogeneity (Pianka 1966), can also be expected to facilitate diversification of both species and languages. In addition, for humans, a key factor may be how the landscape facilitates transport (Nettle 1996).

Habitat heterogeneity is another covariate of species richness (Currie 1991; Rahbek & Graves 2001) and has been suggested as facilitating language diversity (Pagel 2000; Mace & Pagel 1995). One explanation is that habitat heterogeneity could allow a variety of specialist cultures to coexist (Smith 2001). Alternatively, one could argue that habitat heterogeneity allows a single group more options for subsistence and so a viable cultural group can be smaller, or more narrowly distributed, while still maintaining self-sufficiency.

History may also play an important part, particularly in terms of time since previous glaciations or other transformative geological events (Rohde 1999). In addition, the role of chance events such as dispersal or local extinctions has also been suggested as leaving discernible signatures in current distribution patterns (Schluter & Ricklefs 1993). History may play an especially important part in determining the pattern of language distribution, because humans generate historical effects through interactions between themselves and the kinds of societies and technologies that

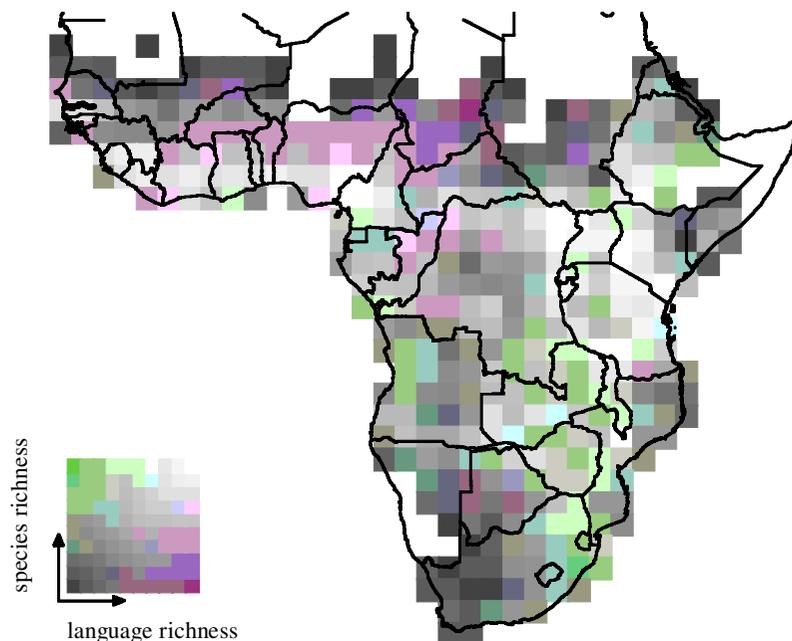


Figure 2. A comparison of language and species richness in sub-Saharan Africa. Each variable has been partitioned into 10 equal-frequency ranks. Cells of equal rank for both variables are coloured in grey scale from black (both low rank) to white (both high rank). The degree of colour saturation indicates that the ranks are different. A cell is green if the species-richness rank is high relative to language richness and blue if language richness is high relative to species richness.

they develop. It has been proposed that language diversity is generally lower in complex societies, in agricultural societies and in areas with a long history of empire (Nichols 1992; Nettle 1996).

These factors need not be mutually exclusive. For example, combining the species richness–energy hypothesis and those that stress the role of history and biome area, recent analyses of South American land birds support the suggestion that the historical interaction between climate and topography is instrumental in generating the species pool from which local assemblages of species are drawn (Rahbek & Graves 2001).

(b) Looking for environmental correlates

After examining overall correlations between vertebrate and linguistic richness, we test the degree to which both patterns are correlated with environmental factors. In addition, we look for evidence that similar factors may be important in determining both species and language richness. We chose 10 environmental variables that have previously been suggested as important in determining either species or language richness and tested their explanatory power in a multiple regression model. The variables were mean length of the growing season, net primary productivity (NPP), mean annual rainfall, within-year rainfall variability, between-year rainfall variability, mean annual temperature, mean annual temperature range, altitudinal range and two measures of habitat variability.

In interpreting the results of these models it is important to consider several limitations. Correlations by themselves do not demonstrate causality. Moreover, even where species and language richness are correlated with similar environmental factors this could well arise through different processes. Humans are a single species and language areas tend to partition a region rather than overlap. In this respect, language diversity resembles genetic and/or

morphological differentiation within species. Moreover, intraspecific diversification typically takes place over shorter time-scales than interspecific differentiation. Hence, it is very likely that even though they may covary in space, the processes determining patterns of species and linguistic richness differ. Nevertheless, with these caveats in mind, cautious interpretation of environmental models can shed light on our understanding of the extent to which such patterns coincide.

2. MATERIAL AND METHODS

(a) Data

We measured biological diversity using total species richness of four vertebrate groups (mammals, birds, snakes and amphibians). Distribution data for 3882 species have been compiled by the Zoological Museum of Copenhagen and mapped onto a one-degree grid (Burgess *et al.* 1998; Brooks *et al.* 2001). Language richness was calculated from the contemporary distribution of 1686 languages (Grimes 1996). Language definitions are based on reciprocal intelligibility measures, particularly as related to literature (Grimes 1996). Widespread languages, such as Arabic (including local variants), Hausa, Swahili and other lingua francas (used predominantly as second languages) and modern European languages (French, Portuguese, English, Greek, Spanish, Italian) were excluded from the analysis. In addition, distributions of immigrant languages are not available outside of their country, or countries, of origin. The majority of countries have associated range maps that show the distribution of each language. However, for some countries, notably Ethiopia, Sudan and Namibia, the location of each language was marked with a point. Hence, in these countries language richness per cell will probably be underestimated.

We analysed the data degraded to a $2^\circ \times 2^\circ$ grid, as this best represented the resolution of the language dataset. Richness was calculated as the number of species or languages present in each

2° × 2° cell. Species distribution data were available for all 515 cells that cover mainland sub-Saharan Africa; however, only 421 cells contained data on languages. While some of the 94 cells without language data represent areas empty of people, many reflect the displacement of languages by modern colonial languages or simply a lack of data. Hence, cells containing no data were excluded from the analyses.

Climate data were derived from annual monthly mean data interpolated onto a 0.5° grid for 1901–1998 (New *et al.* 2000). Data were degraded to 2° by calculating the means of 0.5° grids. Annual rainfall was calculated as the mean of total rainfall and temperature calculated as the mean value of the annual mean of temperature. All temperature calculations were made in K and then converted to °C before further analyses were undertaken. Inter-annual variation in rainfall was measured as the coefficient of variation (CV) of the annual value over the entire 97-year period. Intra-annual variation in rainfall was calculated as the mean of the CV of within-year deviations from the annual mean. Temperature range was calculated as annual maximum minus annual minimum. Mean growing season was calculated as the mean number of months per year for which rainfall was greater than twice the value of mean monthly temperature (in °C; Nettle 1996).

NPP was derived from the DOLY global model (Woodward *et al.* 1995) and was degraded from 0.25° to 2° resolution. Altitudinal range was derived from a 0.05° grid of elevation (Hutchinson *et al.* 1996) and was calculated as the difference between the minimum and maximum value in each 2° cell. Habitat variability was based on White's vegetation map of Africa (White 1983). We defined two measures of habitat variability: habitat richness, the number of different vegetation types in each 2° grid cell; and habitat heterogeneity, the total number of discrete polygons of vegetation types in each 2° grid cell.

We defined three major biome types, each of which were amalgamations of biomes defined in the World Wildlife Fund ecoregion classification (Olson & Dinerstein 1998). These three biomes were forests (combining both wet and dry broadleaf forests), savannahs (including tropical, sub-tropical, montane and flooded grasslands, savannahs and shrublands) and arid areas (combining deserts, xeric shrublands and mediterranean shrublands and woodlands). Each ecoregion has been previously classified as one of these biome types (Olson & Dinerstein 1998). The 2° grid was intersected with the ecoregions map and the area of each biome in each cell was calculated. A 2° cell was classified as a given biome if that biome occupied the largest proportion of the available terrestrial area (i.e. excluding water and mangroves). The intersections were carried out using ARCVIEW GIS 3.2a.

(b) *Statistical methods*

Correlations between species richness, language richness and environmental variables were calculated as Spearman's rank correlations. Correlations of environmental variables with the principal component analysis (PCA, see below) axes were calculated as Pearson's correlation coefficients. The majority of datasets are strongly spatially autocorrelated. Autocorrelation does not affect the value of correlation coefficients (Cliff & Ord 1981), but strong positive autocorrelation in both datasets being compared is likely to overestimate the significance of correlations because the degrees of freedom are overestimated. We used a modified *t*-test (Dutilleul 1993; Legendre 2000) that calculates an effective number of degrees of freedom for autocorrelated data, to test for significance. Note that because the amount of autocorre-

lation differs between variables there is no consistent relationship between the size of the correlation and its significance when different variables are compared.

Regressions were carried out using R (R Development Core Team 2001). Some explanatory variables were transformed prior to analysis to improve their distribution. Intra-annual rainfall variation, inter-annual rainfall variation, temperature range, altitudinal range, habitat richness and habitat heterogeneity were all log transformed and rainfall was square-root transformed. As climatic variables were strongly correlated (table 1), we used PCA to identify two 'climatic' variables that are orthogonal linear combinations of the climate variables (mean growing season, NPP, mean annual rainfall, within-year rainfall variability, between-year rainfall variability and temperature range). General linear models were fitted to each of the species richness and language richness datasets. For species richness, we fitted a model with Poisson errors and log link, and for language richness a model with gamma errors and log link. The language data showed evidence of persistent heteroscedasticity that could not be controlled for by altering the error distribution. For each model, all significant terms were chosen from the two PCA climate variables (named climate1 and climate2), habitat heterogeneity, altitudinal range and their squares, using forward and backward selection procedures in R (stepAIC). In both cases, the forward and backward selection procedure produced the same model. We did not include habitat richness in the models, as it was strongly correlated with habitat heterogeneity but more weakly correlated with language and species richness. Many of the explanatory variables used for the regressions were strongly spatially autocorrelated. Autocorrelation within variables makes interpretation of regression models difficult as significance values may be inflated and apparent correlations between variables may be due to similar patterns of autocorrelation. Here, we wish to examine patterns of association between different variables and so focus entirely on correlation coefficients and R^2 -values, which are unaffected by autocorrelation (Cliff & Ord 1981). Hence, we do not report significance values here and interpret marginally significant terms with considerable caution.

3. RESULTS

Both language richness and species richness show strong latitudinal gradients (figure 1*a*). Moreover, language richness correlates with species richness (figure 1*b*), supporting previous suggestions that cultural diversity mirrors biological diversity. The coincidence is not driven by differences of distribution between taxa—the correlation remains when taxonomic groups are analysed separately (table 2), except for amphibians, which display no significant correlation. Note that the correlations are generally stronger for better-sampled groups (birds and mammals). This suggests that the relation is not simply a reflection of sampling effort being focused close to human settlements, because if that were the case we would expect the most poorly sampled groups to show the strongest correlations (Balmford *et al.* 2001). The correlation of language richness does not hold if only range-restricted species (those in the lowest quartile for range size in their group) are considered (table 2). This is true even if only the 25% most-localized languages are considered ($r_s = 0.10$, $n = 421$, $p > 0.05$).

The strength of the correlation between richness of vertebrates and languages varied between major biomes. The

Table 1. Spearman's rank correlation coefficients between environmental variables and compared with species richness (SR) and language richness (LR) and Pearson's product-moment correlations between the environmental variables and the first two components of the climate PCA analysis (climate1 and climate2).^a

(The environmental variables are mean length of the growing season (GS), net primary productivity (NPP), mean annual rainfall (rain), within-year rainfall variability (WYR), between-year rainfall variability (BYR), mean annual temperature (temp), mean annual temperature range (TR), altitudinal range (ALTR), habitat heterogeneity (HH) and habitat richness (HR).)

	GS	NPP	rain	WYR	BYR	temp	TR	ALTR	HH	HR
SR	0.53*	0.55*	0.48*	-0.46	-0.37	-0.37	-0.44	0.47***	0.59***	0.47***
LR	0.43*	0.37	0.48**	-0.33	-0.49*	0.14	-0.41	0.01	0.20	0.14
GS	—	0.94***	0.91***	-0.90***	-0.80***	-0.16	-0.77**	0.06	0.27	0.14
NPP	—	—	0.90***	-0.84***	-0.74***	-0.13	-0.77**	0.00	0.28	0.14
rain	—	—	—	-0.75***	-0.84***	0.02	-0.78***	0.02	0.118	0.03
WYR	—	—	—	—	0.66**	0.24	0.73***	-0.12	-0.29	-0.18
BYR	—	—	—	—	—	-0.09	0.66***	0.04	-0.17	-0.05
temp	—	—	—	—	—	—	-0.11	-0.46**	-0.33	-0.28
TR	—	—	—	—	—	—	—	-0.04	-0.23	-0.09
ALTR	—	—	—	—	—	—	—	—	0.41**	0.44***
HH	—	—	—	—	—	—	—	—	—	0.85***
Climate1	0.98***	0.95***	0.93***	-0.89***	-0.87***	-0.86**	0	0.03	0.25	—
Climate2	-0.12	-0.06	0.11	0.24	-0.15	-0.18	0.99***	-0.42	-0.28	—

^a Significant correlations are marked with asterisks (*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$); all other correlations are insignificant ($p > 0.05$).

correlation is strongest within forests and also holds within savannahs, but breaks down if arid areas are considered separately (table 2) and this suggests that part of the correspondence may be driven by different mean levels of language and species richness between biomes. It is also consistent with the hypothesis that language diversity is facilitated in areas where travel and transportation are difficult, such as dense tropical forests (Nettle 1996), but the pattern may also simply reflect different mean levels of environmental factors in different biomes affecting language and species richness similarly.

Looking in more detail, there are clearly areas where the distributions of species and languages differ markedly (figure 2). Notably, language richness is higher relative to species richness in the West African savannah mosaics, the inner Congo basin and the Nama Karoo. By contrast, species richness is apparently high in comparison with language richness throughout the central African plateau, the Angolan escarpment and, to a lesser extent, the Ethiopian highlands. This last result is surprising as Ethiopia has a long and continuous cultural history and hence we would expect it to be a centre of cultural diversity. However, Ethiopia was one of the regions in which language ranges were not fully mapped (Grimes 1996). This will decrease apparent language richness relative to other fully mapped regions and we suspect that it is this, rather than low language richness *per se*, which is driving the appearance of relatively low language richness in Ethiopia.

Many of the environmental variables considered in this study, particularly those associated with climate, were highly intercorrelated (table 1). Hence, we used PCA to generate two climate variables that together account for more than 88% of the variation in the seven climate-related variables (table 1). Examination of the correlations between the principal components and the environmental variables indicates that the first principal component is clearly associated with wetness, potential for growth and

climatic variability. The second principal component is associated with temperature.

Environmental factors explain much of the variation in both species and language richness (table 3). The most important component in explaining either species richness or language richness is the 'wetness' component, climate1. This component explains 48 and 20% (summing the contributions of climate1 and climate1²), respectively, of the total observed variation in species and language richness and constitutes 68 and 56% of the explained observed variation, respectively (figure 3). Nevertheless, the functional forms are clearly different. Species richness shows a strong unimodal relation with climate1 (positively correlated with mean growing season, NPP and mean rainfall) while on a log scale language richness shows a weaker unimodal relation. climate2 (correlated with temperature) and altitudinal range were also included in both models, although they appear to be more important in the language-richness than species-richness model. Interestingly, habitat heterogeneity, while important for species (figure 3), explains little of the variation in language richness.

The environmental variables are better predictors of language richness than is species richness, suggesting that there is no need to invoke a direct causal link between species richness and language richness. A regression model for language richness with species richness as the explanatory variable explains only 20% of the observed variation in language richness, approximately half of that explained by the model using the environmental variables. The converse is also true: language richness accounts for only 24% of the observed variation in species richness when language richness is used as the explanatory variable in a species-richness model.

4. DISCUSSION

This study provides, to our knowledge, some of the first quantitative evidence that not only do language richness

Table 2. Spearman's rank correlations with language richness and results of significance tests corrected for autocorrelation. (n is sample size, adj d.f. gives the degrees of freedom adjusted for autocorrelation, F gives the value of the F -statistic and p indicates the significance of the correlation (** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; ns, not significant.)

variable	r_s	n	adj d.f.	F	p
all vertebrate richness	0.46	421	52	14.02	***
range-restricted richness	0.13	421	75	1.32	ns
bird richness	0.43	421	52	11.78	**
mammal richness	0.49	421	48	15.02	***
snake richness	0.44	421	54	13.08	***
amphibian richness	0.21	421	67	3.07	ns
all vertebrate richness:					
forest regions	0.49	74	50	15.86	***
savannah regions	0.34	305	44	5.80	*
arid regions	0.18	42	—	—	ns

Table 3. Multiple regression results.

(The value of the coefficients and the partial and total R^2 for each minimal model. Variables as in table 1.)

species richness	coefficient	R^2	language richness	coefficient	R^2
intercept	6.83	—	intercept	0.86	—
climate1	0.07	0.28	climate1	0.17	0.18
climate1 ²	-0.02	0.20	climate2	0.35	0.09
HH	0.12	0.15	ALTR ²	0.02	0.07
climate2	-0.05	0.02	climate1 ²	-0.03	0.02
climate2 ²	-0.04	0.03	HH	0.17	0.01
ALTR	-0.25	0.04	total	—	0.36
ALTR ²	0.03	0.01			
total	—	0.71			

and species richness show similar latitudinal gradients, but there is a correspondence between the distributions of species and language richness. This correspondence may arise for any of at least three types of reasons.

First, there may be a causal link between species richness and language richness, with species richness directly facilitating language richness by allowing greater opportunities for resource partitioning. However, people utilize a relatively limited number of species as resources and much biodiversity is either indistinguishable in terms of its value as a human resource, or simply unusable (Smith 2001). This, combined with the lower power of species richness compared with environmental factors to explain variation in language richness, suggests that a direct effect of species richness on language richness is very unlikely to be the major cause of the correlation described here.

Alternatively, human culture may act directly on both language richness and species richness. One interesting hypothesis along these lines is that large-scale centralized cultures have the effect of reducing both linguistic and biological diversity (Smith 2001). There exists much good linguistic evidence for the process of centralization and empire in reducing linguistic richness, a prime example being the spread of the Romance languages in Europe (Nichols 1992). In addition, the negative effect of western

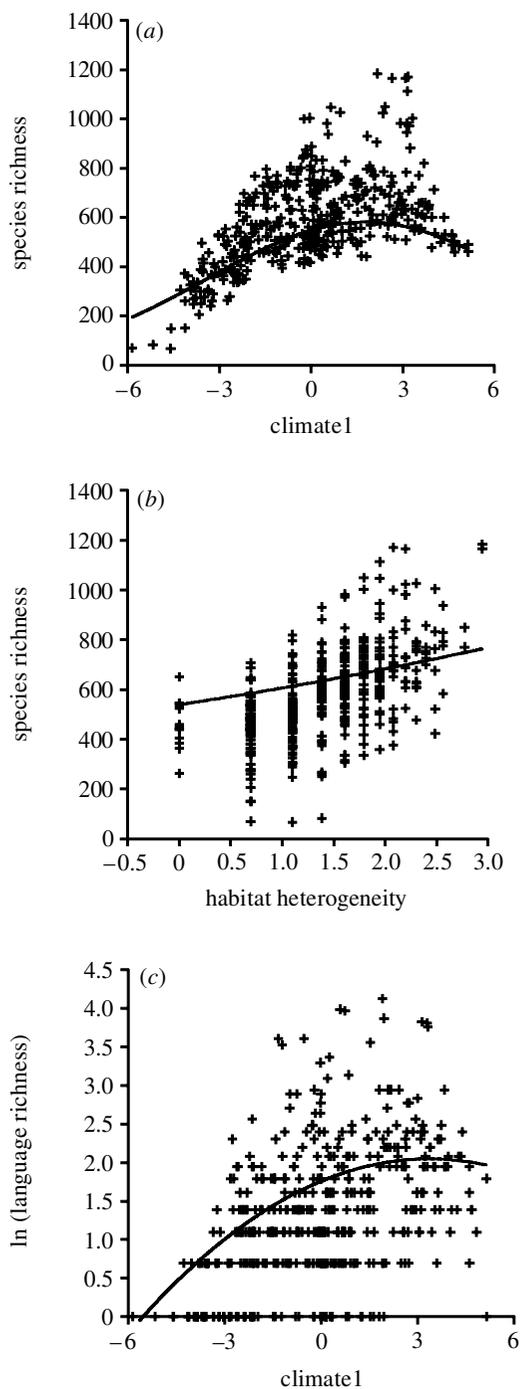


Figure 3. The relationship between (a) species richness and the first climate PCA axis (climate1), (b) species richness and habitat heterogeneity (log scale) and (c) language richness (log scale) and the first climate PCA axis (climate1), based on the models presented in table 3. Increasing values of climate1 correspond to increasing values of mean growing season, NPP and mean annual rainfall and decreasing values of within- and between-year rainfall variability (table 3).

European cultures on biodiversity in Europe has been well documented (Rackham 1986). However, we consider that this explanation is insufficient to account for the pattern uncovered in this analysis, for a number of reasons. First, the comparatively poor performance of language richness as an explanatory variable in a species-richness model suggests that human factors have not driven the pattern

observed in Africa. Second, latitudinal gradients in species richness existed long before the evolution of humans, as attested to by strong latitudinal gradients in the fossil record (Stehli *et al.* 1969; Crane & Lidgard 1989; Jablonski 1993). Third, areas associated with the oldest civilizations in Africa have high language richness (Connagh 1987). Fourth, large-scale cultures associated with language and species loss are typically associated with high population density. Yet, across sub-Saharan Africa, high population density correlates positively with species richness (Balmford *et al.* 2001), contradicting the above hypothesis. Finally, we consider the scale of our work is too coarse to pick up local anthropogenic changes in species distributions due to human activity.

The third, most likely, type of explanation for the correspondence between language and species richness is simply that both variables overlap in their responses to environmental factors. Of course, just because species and language richness show similar associations with a suite of environmental factors does not imply that the processes of species and language diversification and extinction are mechanistically the same.

Both species and language richness were most strongly correlated with the climate1 variable, which correlates with wetness and observed growth. This fits with the idea that species richness is related to the total amount of resources available and the ideas of Nettle (1996) that more languages will develop in areas of low ecological risk, although the hump-backed nature of the relations, especially that of species richness, suggests that other factors are also important, at least at high levels of resource availability.

The inclusion of topographical and habitat heterogeneity terms in the models supports the contention that these kinds of heterogeneity can be important in determining richness patterns (Mace & Pagel 1995; Rahbek & Graves 2001), although the models presented suggest that their influence is relatively weak at the scale of our analysis. It is interesting to note that habitat heterogeneity was more important in the species than the language model. This makes sense when we recall that many species can overlap in their distributions, while peoples who speak different languages tend to partition an area. Hence, we would expect habitat heterogeneity to have a much stronger effect on the number of niches available for a wide range of animal species than on the required range size or potential for overlap of different cultural groups. However, this is unlikely to be a complete explanation as our results contrast with those of Mace & Pagel (1995), who found a strong correlation between language and habitat richness in North America. One possible explanation is that it may be the kind of habitat heterogeneity that matters. For example, habitat heterogeneity may correlate more strongly with highly suitable environmental conditions in North America than it does in Africa. Another potential explanation is simply that the strength of the correspondence may differ between continents or over varying latitudinal ranges.

These models provide some interesting insights into the factors that are associated with language and species richness. However, the evidence for the relative roles of particular factors in determining richness gradients is limited because the environmental variables, particularly NPP,

rainfall and climatic variability are highly inter-correlated at this scale (table 1). This makes it difficult to differentiate between specific variables or processes that appear to affect language or species richness.

Although the environmental factors show a fair correspondence to species and language richness, they do not explain all of the variance in either measure. There are a number of possible reasons for this shortfall in explanatory power. First, we have not incorporated all possible factors into the analysis and the poor overall fits could simply be the result of factors missing from the analysis. Factors not included, but likely to be important, include direct measures of geographical complexity and the nature and strength of biotic interactions. One biotic interaction that could be especially important for humans is the distribution of disease (Crosby 1986; Diamond 1999). It is interesting to note that the central African plateau, one of the areas corresponding to comparatively low language richness, is an area where tsetse flies (*Glossina* spp.) are common, which has limited the opportunities for settlement in these regions.

Second, the underlying null models of language and species-richness distributions can be expected to differ. Whereas a null hypothesis of no latitudinal gradient is reasonable for language richness (since languages tend to partition space), recent work suggests that the null hypothesis for species richness may well be hump-shaped. Recent work on null models of species richness suggests that much of the variation of species richness and, in particular one-dimensional latitudinal gradients, can be largely explained on the basis of geometric constraints (Colwell & Lees 2000). Jetz & Rahbek (2001) have recently shown that a two-dimensional null model of bird species distributions accounts for *ca.* 26% of the observed two-dimensional pattern in sub-Saharan Africa. The choice of null model may have important implications for understanding factors important in determining species richness and also in understanding why the distributions of species and language richness coincide. Hence, testing the effect that incorporating null models has on the power of other explanations is an important project for the future.

Finally, the fit of our environmental models is very probably limited by the different ways in which history may have shaped both species richness and language richness. Major past events, such as mountain building, climate change or catastrophic drought, are all likely to play a part in determining current distributions of both species and languages. The markedly poorer fit of the language-richness model also suggests that human distributions may be more strongly impacted by idiosyncratic historical events, such as key technological innovations, immigration and displacement, cultural shifts and major epidemics. All of these can be expected to strongly affect language richness, since these events are associated with language changes (Nichols 1992). However, they may not be predictable on the basis of existing environmental conditions. The importance of historical contingency is reflected in the power that time since the last cultural shift has to explain current patterns of language richness. For example, the recent spread of the Bantu culture and languages, first through the central forests and then throughout the East African savannahs and central African plateau, has been associated with the development of

novel agricultural techniques that provided new opportunities for resource exploitation (Iliffe 1995). Language richness in these regions corresponds well with the time that the culture group has been present: high richness in the West African forests and savannahs, intermediate in and around the Congo basin and Great Lakes region and low in the East African savannahs and central African plateau (Connagh 1987).

However, our results suggest that, despite these strong, potentially masking historical effects, environmental factors are important in determining the kinds of areas where languages will tend to diversify. Understanding these environmental links will better enable us to understand the processes of language and cultural diversification and the way that the development of culture interacts with the environment.

The authors thank R. Green, B. Adams and J. Fjelds  for stimulating discussions, P. Agnelli, J. L. Amiet, the late W. F. H. Ansell, E. Baker, N. Baker, P. Bates, S. Bearder, W. Bergmans, L. Boitani, B. Branch, D. Broadley, G. Bronner, C. Cabral, M. D. Carleton, A. Channing, C. Claessen, M. Colyn, W. Cotterill, G. Cowlshaw, C. Dardia, G. Davies, F. Dieterlen, N. Dippenaar, B. Dowsett, F. Dowsett-Lemaire, J. M. Duplantier, H. M. de Klerk, J.-P. d'Huart, J. Fahr, J. Fjelds , M. B. Fenton, C. Fitzgibbon, C. Gans, L. Granjon, P. Grubb, T. Halliday, D. C. D. Happold, R. Hoffman, M. E. Holden, P. Howard, B. Hughes, R. Hutterer, N. H. G. Jacobsen, P. Jenkins, W. Jetz, J. Kerbis, D. Koch, A. J. L. Lamberis, M. Languy, M. Largen, H. Leirs, A. Linzey, J. Mutke, J. Oates, B. Patterson, J.-L. Perret, J. C. Poynton, P. A. Racey, J. B. Rasmussen, G. B. Rathburn, L. Robbins, V. Salewski, A. Schi t, D. Schlitter, A. M. Simonetta, J. D. Skinner, S. Spawls, W. Stanley, A. E. Taylor, P. Taylor, V. van Cakenberghe, E. van der Straeten, E. van Dijk, H. van Rompaey, W. Verheyen, V. Wallach, N. Winsler, R. Wirth, D. Yalden and BirdLife International for providing data, and S. Galster, A. Jakobsen, J. B. Larsen, T. Lehmborg, P. V. Nielsen, T. S. Romdal, L. L. S rensen and M. M. Westergaard for entering data. Funding was provided by the Danish Research Council, the Danish Council for Development Research, the Isaac Newton Trust of the University of Cambridge and Conservation International's Center for Applied Biodiversity Science.

REFERENCES

- Balmford, A., Moore, J. L., Brooks, T., Burgess, N., Hansen, L. A., Williams, P. & Rahbek, C. 2001 Conservation conflicts across Africa. *Science* **291**, 2616–2619.
- Brooks, T., Balmford, A., Burgess, N., Fjelds , J., Hansen, L. A., Moore, J., Rahbek, C. & Williams, P. 2001 Toward a blueprint for conservation in Africa. *BioScience* **51**, 613–624.
- Burgess, N., Fjelds , J. & Rahbek, C. 1998 Mapping the distributions of Afrotropical vertebrate groups. *Species* **30**, 16–17.
- Cliff, A. D. & Ord, J. K. 1981 *Spatial processes*. London: Pion.
- Colwell, R. K. & Lees, D. C. 2000 The mid-domain effect: geometric constraints on the geography of species richness. *Trends Ecol. Evol.* **15**, 70–76.
- Connagh, G. 1987 *African civilizations*. Cambridge University Press.
- Crane, P. R. & Lidgard, S. 1989 Angiosperm diversification and paleolatitudinal gradients in cretaceous floristic diversity. *Science* **246**, 675–678.
- Crosby, A. W. 1986 *Ecological imperialism: the biological expansion of Europe 900–1900*. Cambridge University Press.
- Currie, D. J. 1991 Energy and large-scale patterns of animal-species and plant-species richness. *Am. Nat.* **137**, 27–49.
- Diamond, J. 1999 *Guns, germs and steel*. New York: W. W. Norton & Company Inc.
- Dutilleul, P. 1993 Modifying the *t* test for assessing the correlation between 2 spatial processes. *Biometrics* **49**, 305–314.
- Fjelds , J. & Lovett, J. C. 1997 Geographical patterns of old and young species in African forest biota: the significance of specific montane areas as evolutionary centres. *Biodivers. Conserv.* **6**, 325–346.
- Fjelds , J., Ehrlich, D., Lambin, E. & Prins, E. 1997 Are biodiversity 'hotspots' correlated with current ecoclimatic stability? A pilot study using the NOAA-AVHRR remote sensing data. *Biodivers. Conserv.* **6**, 401–422.
- Grimes, B. F. (ed.) 1996 *Ethnologue: languages of the world*. Dallas, TX: Summer Institute of Linguistics.
- Groombridge, B. & Jenkins, M. D. 2000 *Global biodiversity: Earth's living resources in the 21st century*. Cambridge, UK: World Conservation Press.
- Harmon, D. & Maffi, L. 2002 Are linguistic and biological diversity really linked? *Conserv. Biol. Prac.* **3**, 26–27.
- Hutchinson, M. F., Nix, H. A., McMahon, J. P. & Ord, K. D. 1996 *A topographic and climate data base for Africa, version 1.1*. Canberra, Australia: Centre for Resource and Environmental Studies.
- Iliffe, J. 1995 *Africans: the history of a continent*. Cambridge University Press.
- Jablonski, 1993 The tropics as a source of evolutionary novelty through geological time. *Nature* **364**, 142–144.
- Jetz, W. & Rahbek, C. 2001 Geometric constraints explain much of the species richness pattern in African birds. *Proc. Natl Acad. Sci. USA* **98**, 5661–5666.
- Legendre, P. 2000 *MODTTEST*. Montreal: University of Montreal.
- Mace, R. & Pagel, M. 1995 A latitudinal gradient in the density of human languages in North America. *Proc. R. Soc. Lond. B* **261**, 117–121.
- Magurran, A. E. 1988 *Ecological diversity and its measurement*. London: Chapman & Hall.
- Nettle, D. 1996 Language diversity in West Africa: an ecological approach. *J. Anthropol. Archaeol.* **15**, 403–438.
- Nettle, D. 1998 Explaining global patterns of language diversity. *J. Anthropol. Archaeol.* **17**, 354–374.
- New, M., Hulme, M. & Jones, P. 2000 Representing twentieth-century space–time climate variability. II. Development of 1901–96 monthly grids of terrestrial surface climate. *J. Clim.* **13**, 2217–2238.
- Nichols, J. 1990 Linguistic diversity and the first settlement of the new world. *Language* **66**, 475–521.
- Nichols, J. 1992 *Linguistic diversity in space and time*. Chicago, IL: University of Chicago Press.
- Nichols, J. 1997 Modeling ancient population structures and movement in linguistics. *A. Rev. Anthropol.* **26**, 359–384.
- Olson, D. & Dinerstein, E. 1998 The Global 200: a representation approach to conserving the Earth's most biologically valuable ecoregions. *Conserv. Biol.* **12**, 502–515.
- Oviedo, G., Maffi, L. & Larsen, P. B. 2000 *Indigenous and traditional peoples of the World and ecoregion conservation*. Gland, Switzerland: WWF-World Wide Fund for Nature.
- Pagel, M. 2000 The history, rate and pattern of world linguistic evolution. In *The evolutionary emergence of language* (ed. C. Knight, M. Studdert-Kennedy & J. R. Hurford), pp. 391–416. Cambridge University Press.
- Pianka, E. R. 1966 Latitudinal gradients in species diversity: a review of the concepts. *Am. Nat.* **100**, 33–46.
- R Development Core Team 2001 R: a programming environment for data analysis and graphics, version 1.3.1 (August 2001).
- Rackham, O. 1986 *The history of the countryside*. London: Dent.
- Rahbek, C. & Graves, G. R. 2001 Multiscale assessment of patterns of avian species richness. *Proc. Natl Acad. Sci. USA* **98**, 4534–4539.
- Rohde, K. 1999 Latitudinal gradients in species diversity and

- Rapoport's rule revisited: a review of recent work and what can parasites teach us about the causes of the gradients? *Ecography* **22**, 593–613.
- Rosenzweig, M. L. 1995 *Species diversity in space and time*. Cambridge University Press.
- Schluter, D. & Ricklefs, R. E. 1993 Species diversity: an introduction to the problem. In *Species diversity in ecological communities* (ed. R. E. Ricklefs & D. Schluter), pp. 1–10. Chicago, IL: University of Chicago Press.
- Smith, E. A. 2001 On the coevolution of cultural, linguistic and biological diversity. In *On biocultural diversity: linking language, knowledge and the environment* (ed. L. Maffi). Washington, DC: Smithsonian Institution Press.
- Stehli, F. G., Douglas, R. G. & Newell, N. D. 1969 Generation and maintenance of gradients in taxonomic diversity. *Science* **164**, 947–949.
- White, F. 1983 *Vegetation of Africa—a descriptive memoir to accompany the Unesco/AETFAT/UNSO vegetation map of Africa*. Natural Resources Research report XX. Paris: UN Educational, Scientific and Cultural Organization.
- Woodward, F. I., Smith, T. M. & Emanuel, W. R. 1995 A global land primary productivity and phytogeography model. *Global Biogeochem. Cycles* **9**, 471–490.

As this paper exceeds the maximum length normally permitted, the authors have agreed to contribute to production costs.