

The size distribution of conspecific populations: the peoples of New Guinea

Vojtech Novotny^{1*} and Pavel Drozd²

¹*Institute of Entomology, Czech Academy of Sciences and Biological Faculty, University of South Bohemia, Branisovska 31, 370 05 Ceske Budejovice, Czech Republic*

²*University of Ostrava, Department of Biology, Brafova 7, 701 03 Ostrava, Czech Republic*

The size distribution of the language populations in New Guinea, which represent over 15% of the world's languages, is analysed using models analogous to the resource division models of species abundance distribution in ecological communities. A model distribution of resource segments reflecting population size is created by repeated selection of an existing resource segment and its division into two. We found that any dependency of the selection probability on the size of the segment generated negatively skewed abundance distributions after log transformation. Asymmetric segment division further exacerbated the negative skewness. Size-independent selection produced lognormal abundance distributions, irrespective of the segment division method. Size-dependent selection and asymmetric division were deemed reasonable assumptions since large language populations are more likely to generate isolates, which develop into new populations, than small ones, and these isolates are likely to be small relative to the progenitor population. A negatively skewed distribution of the log-transformed population sizes was therefore expected. However, the observed distributions were lognormal, scale invariant for areas containing between 100 and over 1000 language populations. The dynamics of language differentiation, as reflected by the models, may therefore be unimportant relative to the effect of variable growth rates among populations. All lognormal distributions from resource division models had a higher variance than the observed one, where half of the 1053 populations had between 350 and 3000 individuals. The possible mechanisms maintaining such a low variance around a modal population size of 1000 are discussed.

Keywords: size distribution; New Guinea; resource division model

1. INTRODUCTION

Language is worth a thousand pounds a word!
Lewis Carroll, *Through the Looking Glass*

The distribution of species abundance in ecological communities has been modelled by various resource division models, each stipulating a particular way of partitioning a common, limiting resource between species (Tokeshi 1993, 1997). Such a resource can be schematically represented by a line of unit length, which is divided into two segments and then one of them chosen and divided again, etc. A community of N species is modelled by $N-1$ such selection and division cycles; the length distribution of the resulting segments is then compared with the population sizes of species. The probability of segment selection can be independent of its size (random fraction model) (Tokeshi 1993), a linear function of its size (broken stick model) (MacArthur 1960; Tokeshi 1997) or a power function of its size (power fraction model) (Tokeshi 1996). The point of division can be chosen at random, as in all the above-mentioned models, or it can produce new segments with a fixed length ratio, such as 0.75:0.25 in the sequential breakage model (Sugihara 1980).

As an alternative to resource division models, species abundance distributions have been fitted by statistical distributions, in particular the lognormal distribution, which can result from a normal distribution of population growth rates among species (the central limit theorem) (May 1975).

* Author for correspondence (binatangi@datec.com.pg)

The premises of both resource division and lognormal models are realistic for conspecific populations that are likely to have a common limiting resource and whose population growth rates vary, as a result of numerous ecological factors, around a species-specific value. However, the abundance distribution models have only been applied to the analysis of communities, where their assumptions are more tenuous. This is probably because few species consist of discrete populations which function as demographic entities or entities competing with one another for resources (cf. Hughes *et al.* 1997).

Homo sapiens L., which is subdivided into numerous language groups, approximates such a species. Human culture groups have been postulated to evolve at least partly according to a hierarchical branching process (Cavalli-Sforza *et al.* 1988; Durham 1991; Mace & Pagel 1994). Language barriers cause and later contribute to maintaining the partial reproductive isolation (Barbujani 1991), which may result in genetic and morphological differences among language populations (Sokal *et al.* 1988; Smith *et al.* 1994). Language populations often divide up habitable areas into a mosaic of exclusive, non-overlapping and defended territories (e.g. Matthiessen 1962; Mace & Pagel 1995; Nettle 1999). We suggest that these characteristics make resource division models a useful tool in the study of the processes of origin, fragmentation, extinction and growth of language populations. This approach may also be applicable to other similarly structured population systems, such as Indian castes (Gadgil & Malhotra 1983).

In this study, we develop a new, generalized resource division model and apply it to the size distribution of the

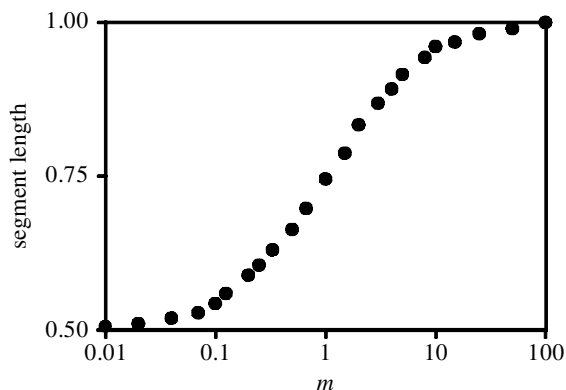


Figure 1. The average length of the larger segment resulting from a division of a resource unit into two segments using the generalized power fraction model with various values of m . The averages from 1000 replicates are given for each data point.

language populations of New Guinea. The island of New Guinea is one of a few areas where linguistic diversity was not drastically reduced during the Holocene period (Diamond 1997). With less than 0.001% of the world's populace speaking over 15% of its languages, it is an area particularly inviting for language population studies.

2. METHODS

In Tokeshi's (1996) power fraction model, the resource unit line is divided into two segments at random and in each next step one of the segments is chosen and divided at random again until N segments are obtained. The probability of an i th segment being selected for division is $p_i = ax_i^k$, where a is a constant ($\sum ax_i^k = 1$), x_i is the length of the segment and k is a parameter of the model. The point of division of the i th segment, delimited by the $x_{i\min}$ and $x_{i\max}$ bounds on the resource unit line, is determined as $x_{i\min} + 0.5x_i z$, where z is a random number drawn from a uniform (0, 1) distribution. The length of the larger of the two segments created by this division can be any value from 0.5 to 1.0 with equal probability, so that its average length is 0.75. The power fraction model with $k = 0$, i.e. the size-independent selection probability, is identical to the random fraction model, while for $k = 1$ the probability of selection is a linear function of the segment length and the model is identical to the broken stick model.

The generalized version of the model proposed here retains the segment selection procedure, but also introduces a power function to the division algorithm. The point of division of an i th segment is determined as $x_{i\min} + 0.5x_i z^m$, where m is a parameter of the model ($m \geq 0$). The average length of the larger segment converges to 0.5 for small m -values and to 1.0 for large m -values (figure 1). For $m = 1$, the average length is 0.75 and the model is identical to Tokeshi's original version (Tokeshi 1996). For $m = 0$, the division becomes deterministic as each segment is split in the middle. PowerNiche software (Drozd & Novotny 1999) was used for all calculations.

The size distribution of the language populations was derived from census data reported for Papua New Guinea (PNG) by Grimes (1996) and for Irian Jaya by Silzer & Clouse (1991). These two data sets were also combined to analyse the island of New Guinea. The most linguistically diverse province, Madang, was also analysed separately. The data from PNG and Madang

were used for most of the analyses because of their higher precision (Silzer & Clouse 1991; Grimes 1996).

The size distribution of areas occupied by individual languages was derived from maps published by Wurm & Hattori (1981) for PNG and Zgraggen (1975) for Madang. This analysis did not include populations from the North Solomons Province as well as several other populations where information on area was not available.

The population and area sizes were \log_{10} transformed, fitted by a lognormal distribution and compared with various niche division models. Differences between the empirical and model distributions were tested by a Kolmogorov–Smirnov test.

The growth rates of the PNG language populations are not known. Population censuses which were completed in 1980 and 1990 (Papua New Guinea National Statistical Office 1988, 1993) report on the numbers for 595 census districts; however, they ignore language boundaries. The growth rate in each district was expressed as the ratio of the 1990 population to the 1980 population. Five districts with average population sizes of less than 250 individuals were excluded.

3. RESULTS

The language population size and area were correlated (\log_{10} -transformed data from PNG $r = 0.575$, $p < 0.001$ and $n = 722$). The distribution of the language population sizes on all spatial scales, from Madang to the whole of New Guinea, was lognormal or close to lognormal, with the geometric mean of the population size ranging from 543 in the Madang data to 1112 in the PNG data (table 1 and figure 2). The distribution of the language areas was also lognormal, with geometric means of 40 km² in Madang and 164 km² in PNG (table 1). The distributions of the population size had a lower coefficient of variance than the distributions of the population area (0.23 and 0.32 for PNG and 0.21 and 0.34 for Madang, respectively) (cf. table 1). The differences in the p -values of the Kolmogorov–Smirnov test for lognormality between the data sets reflect more increasing sensitivity of the test for higher sample sizes than the magnitude of differences of skewness or kurtosis from zero (table 1).

The distribution of the population sizes in PNG was approximated by a series of models with various values of the k and m parameters (figure 3). All models with $k = 0$ produced lognormal distributions (figure 3a) (none of the distribution was different from lognormal, $p > 0.05$) with skewness and kurtosis close to zero. The variance increased and the mean decreased rapidly with increasing m . The $m = 0$ distribution, i.e. the one with a deterministic 0.5 : 0.5 division of segments was the closest, although still significantly ($p < 0.05$) different from the empirical data on the population sizes (figure 2) and population areas (V. Novotny and P. Drozd, unpublished data).

Distributions with $k = 1$ exhibited an increasing variance and kurtosis and a decreasing mean and skewness with increasing values of m (figure 3b). Distributions with $m = 1$ had an increasing mean and kurtosis and a decreasing variance and skewness with increasing values of k (figure 3c). None of these models produced a distribution similar to the empirical data (the differences between the model and empirical data were significant in all cases, $p < 0.001$) (cf. figure 2). The model distributions were close to lognormal for only either k or m close to zero;

Table 1. Distribution of the language population sizes (N , in individuals) and areas (A , in km^2)

(Variable, variable tested; n , sample size; maximum difference, the maximum difference between the observed and the lognormal cumulative distribution functions; p , the probability with which a hypothesis on lognormality can be rejected (Kolmogorov-Smirnov test). The data are \log_{10} transformed.)

| | variable | n | mean | variance | skewness | kurtosis | maximum difference | p |
|------------|----------|------|-------|----------|----------|----------|--------------------|-------|
| Madang | N | 172 | 2.735 | 0.331 | 0.231 | -0.202 | 0.057 | 0.181 |
| PNG | N | 816 | 3.046 | 0.488 | -0.050 | 0.625 | 0.030 | 0.081 |
| Irian Jaya | N | 242 | 2.880 | 0.600 | -0.030 | 0.399 | 0.054 | 0.084 |
| New Guinea | N | 1053 | 3.009 | 0.520 | -0.070 | 0.562 | 0.034 | 0.006 |
| Madang | A | 172 | 1.600 | 0.298 | 0.372 | -0.181 | 0.050 | 0.342 |
| PNG | A | 730 | 2.214 | 0.486 | -0.107 | -0.449 | 0.034 | 0.050 |

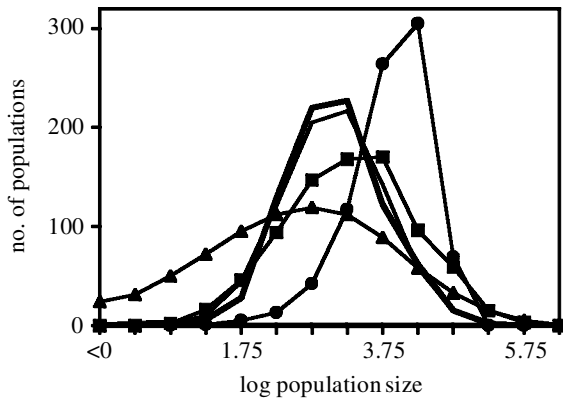


Figure 2. Distribution of the language population sizes (thick line), lognormal (thin line) and some of the resource division models (lines with markers). Triangles, sequential division model ($k = 0$ and deterministic 0.75:0.25 division); circles, broken stick model ($k = 1$ and $m = 1$); squares, generalized power fraction model ($k = 0$ and $m = 0$, i.e. size-independent selection and deterministic 0.5:0.5 division). The population sizes are \log_{10} transformed.

large k - and m -values consistently produced distributions with negative skewness and positive kurtosis (figure 3). It is therefore unlikely that any other combination of $k > 0$ and $m > 0$ can provide a good fit to the empirical data set. This was confirmed by numerous simulations for such k and m combinations which are not shown here.

The distribution of the population growth rates for the census districts was significantly different from the normal distribution (figure 4) ($p < 0.01$). Both skewness (1.65) and kurtosis (7.11) were significantly ($p < 0.01$) different from zero.

4. DISCUSSION

Resource division models, with their recurrent selection and division algorithm, can be used to test simple hypotheses about the mechanisms generating patterns of abundance, particularly in complex systems where experiments are not possible (cf. Nee *et al.* 1991; Gaston & Blackburn 1997; Gaston 1998). The present simulations demonstrate that, when the probability of segment selection for division is dependent on its size, the resulting distribution of segment sizes differs from lognormal and is negatively skewed after log transformation. The skewness is further exacerbated by any process of asymmetric

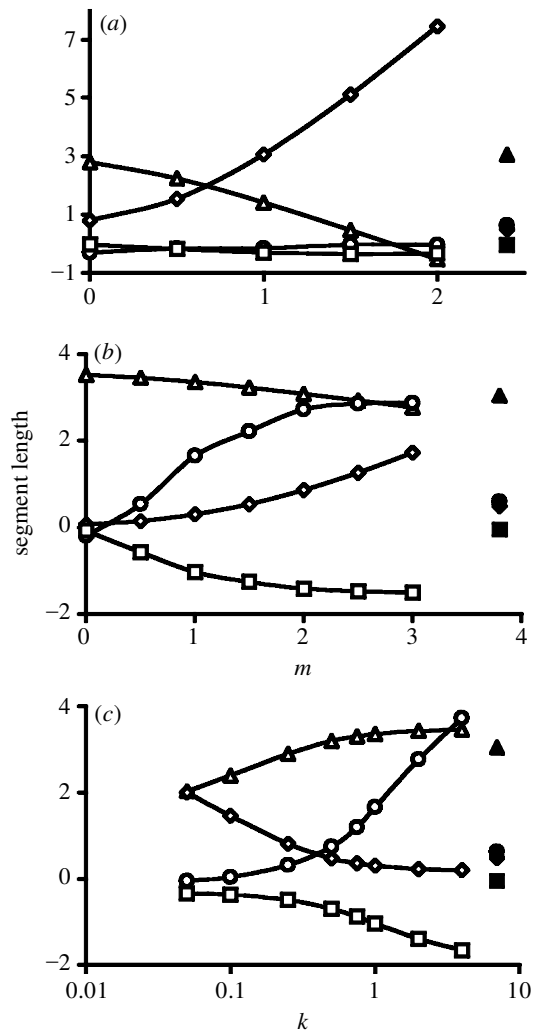


Figure 3. Distribution parameters for segment lengths generated by various niche division models. (a) $N = 816$ and $k = 0$, (b) $N = 816$ and $k = 1$ and (c) $N = 816$ and $m = 1$. The segment lengths are \log_{10} transformed and their mean (triangles), variance (diamonds), kurtosis (circles) and skewness (squares) are depicted. All distributions are averages from 100 replicates. The parameters of the language populations of PNG are given for comparison (solid markers).

division, producing segments of uneven length. In contrast, size-independent selection of segments produced a lognormal distribution for various segment division algorithms. Kolmogoroff (1941, cited by Sugihara 1980) showed that sequential division led to a lognormal size

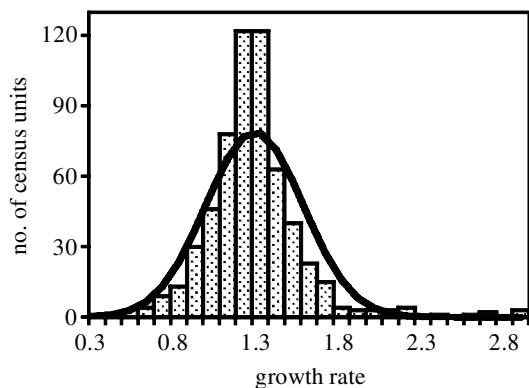


Figure 4. Distribution of the relative growth rates (1990/1980 population) for the census districts in PNG (bars). A normal distribution (line) is fitted to the data.

distribution, provided the magnitude and frequency of division were independent of segment size.

For modelling the origin of languages or species (leaving subsequent population changes aside), size-dependent probability of speciation and asymmetric division of the existing population are reasonable assumptions. A larger population has a greater chance of having some part of it geographically isolated than a smaller population. Large populations also survive longer, which further increases their chances of founding new populations. The pattern of division of a population into two isolates can be highly asymmetric when isolates arise at the periphery of the population's range. Large-scale isolation events can produce any degree of asymmetry and, thus, correspond to the random division algorithm (Gaston 1998). In contrast, there is no particular reason to expect division more regular than random, which would tend to split a population in half. It can therefore be concluded that a negatively skewed distribution (after log transformation) is the most likely expectation for language or species population sizes, although it is unclear how such a distribution can change as a result of subsequent changes in population sizes, including extinction (Gaston 1998).

The distribution of range sizes in various animal taxa is usually negatively skewed (Pagel *et al.* 1991; Gaston & Blackburn 1997; Gaston 1998). The distribution of population sizes of various species across large areas is difficult to evaluate since high-quality data, accurate even for rare species, are scarce. The best data set to date, on British birds, has a negative skewness, although not significantly different from zero (Nee *et al.* 1991; Gregory 1994). We are not aware of any adequate data on conspecific populations.

The present data on language population sizes and areas are lognormal or close to lognormal and their skewness after log transformation was not significantly different from zero. Although the accuracy of the present data is below average for human census data due to inadequate linguistic knowledge and the inaccessibility of many areas in New Guinea, it exceeds any population data available for other species. The similarity of the population size and area distributions is noteworthy. While the dynamics of many populations have been influenced by recent colonial and post-colonial social changes,

particularly increased mobility of the populace, the traditional territories of language populations have remained conserved from the early colonial administration (White 1972; Waiko 1993).

Models using a size-dependent probability of segment selection for division proved incapable of generating the observed distribution of population sizes or areas. The broken stick model belonged among these unsuccessful models as well. This model was originally derived by the random placement of $N-1$ points on a resource unit length (MacArthur 1960), which can be analogous to the random placement of $N-1$ barriers represented, for example, by features of geography isolating language populations.

Since the observed distribution is lognormal or close to lognormal, an alternative hypothesis, namely that the mechanism of language origin is relatively unimportant compared with subsequent population growth rates, seems plausible. The present numbers, distributions and sizes of the language populations are a composite result of language differentiation, extinction and population dynamics since the colonization of the island at least 40 000 years ago (Wurm & Hattori 1981; Wurm 1983; Groube *et al.* 1986). A lognormal distribution can be generated from identical population sizes provided that the relative growth rates are normally distributed. The observed growth rates were not normally distributed, but the large kurtosis represented the more important departure from normality than skewness. The distribution was close to symmetrical, except for several very large growth rates, which were mostly artefacts caused by recent migration to urban centres. The variation in the population growth rates could be generated by the variability in the environment and social factors found among the populations of New Guinea. For instance, the introduction of the sweet potato triggered an expansion of some language populations in the highlands a few centuries ago (Golson 1997); these populations remain among the largest in PNG. In the lowlands, malaria and subsistence on sago are important limiting factors keeping the overall population density low (Townsend 1971; Riley 1983).

Resource division models based on size-independent selection of segments for division generated lognormal distributions, but always with a higher variance and lower mean than the observed one. This was also the case for Sugihara's (1980) sequential breakage model, which was found to fit numerous data from animal communities (Nee *et al.* 1991). Lognormal distributions fitted to community data also had a higher variance than the present data set (Preston 1962; May 1975; Sugihara 1980). The only model that produced a lognormal distribution with comparably low variance and similar mean to the empirical data was the size-independent selection of segments and their deterministic, regular division ($k = 0$ and $m = 0$).

Some processes maintaining a low variance around the modal population size seem to be likely in view of the fact that half of the 1053 populations of New Guinea fall within a narrow range of sizes from 350 to 3000 individuals and that the coefficient of variance is lower for the population size than population area distributions. It has been suggested that there may be an upper limit of around 500 individuals on the size of human groups

which can act as a basic, self-defining, sociodemographic unit (Birdsell 1953; Hunn 1994; cf. Sillitoe 1977). Although the populations in New Guinea are structured to tribes and clans, which usually correspond better to such units than the whole language population, it is clear that such a social structure would make the fragmentation of large populations more likely; it may also be a result of active choices by people (cf. Boehm 1997). The highly dissected terrain of New Guinea also contributes to the fragmentation of large populations. On the other hand, very small populations are always in danger of extinction due to stochastic factors. The role of environmental conditions in the differentiation of languages is indicated, for instance, by similar patterns of mammal-specific and human linguistic diversity (Mace & Pagel 1995; cf. Meggers 1975).

Interestingly, an effective population size of 500 has been suggested as the minimum size sufficient to maintain genetic variation for adaptation to a changing environment (Lande 1988). A similar number of categories also represents an upper limit on the size of traditional, orally transmitted classification systems, possibly reflecting cognitive limits on the size of such systems (Berlin 1992). The low variance in the population sizes encountered in New Guinea can thus have social and environmental as well as genetic explanations. As with any complex system, the dissection of individual factors is difficult, but further study of the 'magic number 500' (Hunn 1994) as the socially maximum and genetically minimum size of a socially coherent and reproductively (semi-) isolated population can be particularly enlightening.

We thank our PNG collaborators and friends J. Auga, W. Boen, C. Dal, S. Hiuk, B. Isua, M. Kasbal, R. Kutil, K. Molem and M. Manumbor, whose collective command of eight languages inspired this paper. G. Weiblen commented on the manuscript and Hana Drozdova and Ron Englund provided technical help. This research was funded by grants from the Fulbright Commission (to V.N.), the US National Science Foundation (DEB-97-07928), Czech Grant Agency (206/99/1115) and the Czech Ministry of Education (ES041).

REFERENCES

- Barbujani, G. 1991 What do languages tell us about human microevolution? *Trends Ecol. Evol.* **6**, 151–156.
- Berlin, B. 1992 *Ethnobiological classification. Principles of categorization of plants and animals in traditional societies*. Princeton University Press.
- Birdsell, J. B. 1953 Some environmental and cultural factors influencing the structuring of Australian aboriginal populations. *Am. Nat.* **87**, 171–207.
- Boehm, C. 1997 Impact of the human egalitarian syndrome on Darwinian selection mechanics. *Am. Nat.* **150**, S100–S121.
- Cavalli-Sforza, L. L., Piazza, A., Menozzi, P. & Mountain, J. 1988 Reconstruction of human evolution: bringing together genetic, archaeological, and linguistic data. *Proc. Natl Acad. Sci. USA* **85**, 6002–6006.
- Diamond, J. 1997 The language steamrollers. *Nature* **389**, 544–546.
- Droz, P. & Novotny, V. 1999 *PowerNiche: niche division models for community analysis*, v. 1. Published and available at <http://www.entu.cas.cz/png/powerniche/index.html>.
- Durham, W. H. 1991 *Coevolution. Genes, culture, and human diversity*. Stanford University Press.
- Gadgil, M. & Malhotra, K. C. 1983 Adaptive significance of the Indian caste system: an ecological perspective. *A. Hum. Biol.* **10**, 465–478.
- Gaston, K. 1998 Species-range size distributions: products of speciation, extinction and transformation. *Phil. Trans. R. Soc. Lond. B* **353**, 219–230.
- Gaston, K. & Blackburn, T. M. 1997 Age, area and avian diversification. *Biol. J. Linn. Soc.* **62**, 239–253.
- Golson, J. 1997 From horticulture to agriculture in the New Guinea highlands. A case study of people and their environments. In *Historical ecology in the Pacific Islands: prehistoric environmental and landscape change* (ed. P. V. Kirch & T. L. Hunt), pp. 39–50. Yale University Press.
- Gregory, R. D. 1994 Species abundance patterns of British birds. *Proc. R. Soc. Lond. B* **257**, 299–301.
- Grimes, B. F. (ed.) 1996 *Ethnologue*, 13th edn. Dallas, TX: Summer Institute of Linguistics.
- Groube, L., Chappell, J., Muke, J. & Price, D. 1986 A 40,000 year-old human occupation site at Huon Peninsula, Papua New Guinea. *Nature* **324**, 453–455.
- Hughes, J. B., Daily, G. C. & Ehrlich, P. R. 1997 Population diversity: its extent and extinction. *Science* **278**, 689–692.
- Hunn, E. 1994 Place-names, population density, and the magic number 500. *Curr. Anthropol.* **35**, 81–85.
- Kolmogoroff, A. N. 1941 Über das logarithmisch normale Verteilungsgesetz der Dimensionen der Teilchen bei Zerstückelung. *CR (Doklady) Acad. Sci. URSS* **31**, 99–101.
- Lande, R. 1988 Genetics and demography in biological conservation. *Science* **241**, 1455–1460.
- MacArthur, R. H. 1960 On the relative abundance of species. *Am. Nat.* **94**, 25–36.
- Mace, R. & Pagel, M. 1994 The comparative method in anthropology. *Curr. Anthropol.* **35**, 549–564.
- 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.
- Matthiessen, P. 1962 *Under the mountain wall. A chronicle of two seasons in the Stone Age*. New York: Viking Press.
- May, R. M. 1975 Patterns of species abundance and diversity. In *Ecology and evolution of communities* (ed. M. L. Cody & J. M. Diamond), pp. 81–120. Cambridge, MA: Belknap.
- Meggers, B. J. 1975 Application of the biological model of diversification to cultural distributions in tropical lowland South America. *Biotropica* **7**, 141–161.
- Nee, S., Harvey, P. H. & May, R. M. 1991 Lifting the veil on abundance patterns. *Proc. R. Soc. Lond. B* **243**, 161–163.
- Nettle, D. 1999 Linguistic diversity of the Americas can be reconciled with a recent colonization. *Proc. Natl Acad. Sci. USA* **96**, 3325–3329.
- Pagel, M. D., May, R. M. & Collie, A. R. 1991 Ecological aspects of the geographical distribution and diversity of mammalian species. *Am. Nat.* **137**, 791–815.
- Papua New Guinea National Statistical Office 1988 *1980 national population census. Final figures: provincial summary*, vols 1–19. Port Moresby, Papua New Guinea: National Statistical Office.
- Papua New Guinea National Statistical Office 1993 *1990 national population census. Final figures: census unit populations*, vols 1–18. Port Moresby, Papua New Guinea: National Statistical Office.
- Preston, F. W. 1962 The canonical distribution of commonness and rarity. *Ecology* **43**, 410–432.
- Riley, I. D. 1983 Population change and distribution in Papua New Guinea: an epidemiological approach. *J. Hum. Evol.* **12**, 125–132.
- Sillitoe, P. 1977 Land shortage and war in New Guinea. *Ethnology* **16**, 71–81.
- Silzer, P. J. & Clouse, H. H. 1991 *Index of Irian Jaya languages. A special publication on Irian, bulletin of Irian Jaya*, 2nd edn. Dallas, TX: Summer Institute of Linguistics.

- Smith, T., Bhatia, K., Prasad, M., Koki, G. & Alpers, M. 1994 Altitude, language, and class I HLA allele frequencies in Papua New Guinea. *Am. J. Phys. Anthropol.* **95**, 155–168.
- Sokal, R. S., Oden, N. L. & Thompson, B. A. 1988 Genetic changes across language boundaries in Europe. *Am. J. Phys. Anthropol.* **76**, 337–361.
- Sugihara, G. 1980 Minimal community structure: an explanation of species abundance patterns. *Am. Nat.* **116**, 770–787.
- Tokeshi, M. 1993 Species abundance patterns and community structure. *Adv. Ecol. Res.* **24**, 111–186.
- Tokeshi, M. 1996 Power fraction: a new explanation of relative abundance patterns in species-rich assemblages. *Oikos* **75**, 543–550.
- Tokeshi, M. 1997 Species coexistence and abundance: patterns and processes. In *Biodiversity. An ecological perspective* (ed. T. Abe, S. A. Levin & M. Higashi), pp. 35–58. New York: Springer.
- Townsend, P. K. 1971 New Guinea sago gatherers: a study of demography in relation to subsistence. *J. Ecol. Food Nutr.* **1**, 19–24.
- Waiko, J. D. 1993 *A short history of Papua New Guinea*. Melbourne, Australia: Oxford University Press.
- White, O. 1972 *Parliament of a thousand tribes, Papua New Guinea: the story of an emerging nation*. Melbourne, Australia: Wren Publishing.
- Wurm, S. A. 1983 Linguistic prehistory in the New Guinea area. *J. Hum. Evol.* **12**, 25–35.
- Wurm, S. A. & Hattori, S. 1981 *Language atlas of the Pacific area*. Canberra: Australian Academy of the Humanities.
- Zgraggen, J. A. 1975 *The languages of the Madang district, Papua New Guinea*. Canberra: The Australian National University.