Copyright 2006, Living Tongues Institue for Endangered Languages.

Occasional Papers Series, no. 01

[DRAFT, PLEASE CHECK WITH AUTHORS BEFORE CITING VERBATIM]

Language Hotspots:
Linking language extinction, biodiversity & the human knowledge base.

Gregory D. S. Anderson (gdsa@livingtongues.org) and K. David Harrison (dharris2@swarthmore.edu)[1]

Human languages, like plant and animal species, are globally unevenly distributed and face uneven extinction risks. Adapting a model that has become prominent in conservation biology, we identify and map a set of global "language hotspots" for prioritization and further investigation. Hotspots are defined as geographic areas with the greatest linguistic endangerment and diversity; they are identified based on quantitative and qualitative criteria such as linguistic (genetic) diversity, degree of endangerment, and level of scientific documentation. Working within twenty-two designated hotspots that contain nearly half of the estimated 7,000 languages of the world, we will be able to formulate and address a number of hypotheses about language change, cultural dispersion, and interactions between language and environment. We take speech communities to be both a basic unit of analysis and the agent of change in processes of human dispersion, globalization, language extinction and ecosystem management. This research will significantly enhance our knowledge of how language—the prime example of a culturally transmitted cognitive system—can change over time and disperse across space.

Geography has clearly shaped human migration and settlement and therefore the distribution of linguistic diversity. Environmental and biogeographic factors such as latitude, elevation, terrain[2], and ecosystem type have also been proposed to correlate with language distribution, but few strong patterns have thus far been detected (Harmon 1996, Nichols 1990, 1992). Our detailed mapping of the world's most endangered language zones, coupled with recent advances in global satellite imagery and mapping technologies, is expected to reveal previously unquantified correlations.

Languages also affect the host ecosystems of their speakers, providing the most efficient conduit for technologies of resource management in traditional societies. Much (if not most) of humankind's accumulated knowledge of the natural world is encoded solely in languages that have never been written down or documented and are now facing extinction. With language

[1] We thank Robbie Hart, Chris Rainier, Allen Carroll and others for their feedback as we developed these dieas.

[2] for example, the Terrain Ruggedness Index (Riley, et al. 1999, Stepp et al 2005)

extinction comes an unprecedented erosion of human knowledge about the natural world. In our survey of endangered languages within the hotspots, we will collect data on how languages uniquely encode in their grammars and lexicons specific information about topography, endemic species and other environmental factors such as weather patterns and vegetation cycles. Linguistic encoding allows for efficient transfer of not only names for things, but also of complex and hierarchal taxonomic relations among species and other ecosystem elements. Most of this information is packaged in such a way that it cannot be directly translated, and thus erodes or dissipates when a community shifts over to speaking a global language (Nettle & Romaine 2000, Lizarralde 2001, Magga 2005, Harrison 2007).

1.0 Language hotspots

The concept of biodiversity hotspots has brought a new energy and perspective to the conservation movement. As the framework has evolved (Myers 1988, 1990, 1995, Mittermeier et al 1999, Meyers et al 2000, Mittermeier et al 2005 (Revisited)), it has helped focus research and set priorities in conserving endangered species and in identifying the estimated 83% of plant and animal species that remain unidentified and unclassified by Western science (Heywood 1995, Wilson 2003). By focusing on specific ecoregions with high levels of species endemism, species diversity and threat, biodiversity hotspots models help illuminate the rich interrelatedness of species and ecosystems, rather than construing individual species in isolation. These models also advance efforts to quantify and map the distribution and density of biodiversity, allowing us to get a better picture of extinction rates and trajectories.

Languages, though not precisely analogous to species, also have an ecology (Haugen 1972, Mülhausler 1996, Mufwene 2001). Like species, languages show uneven distribution, unequal threat of extinction, and they exist in a complex social and ecological matrix. The overwhelming majority of the world's languages have not yet been adequately described, digitally recorded, or scientifically documented to allow for their revitalization if they were to cease being spoken. Languages rely on both internal (socially adaptive) and external (political, demographic) factors for their survival. In the current picture of global language endangerment, there is a gap in understanding how interactions between humans and ecosystems are mediated and shaped by language in both directions. For example, many *Language Hotspots* are populated by traditional hunter-gatherer or other subsistence societies, and thus represent the locus of intense human interaction with the environment. Hotspots by definition have a high level of language diversity, and thus give a rich picture of the long-term co-existence and mutual influence of languages over long periods of time. Our project will explore the broad range of interactions among language, knowledge and environment, and thus provide insights into preserving language diversity.

We emphasize that biodiversity around the world has arisen slowly over hundreds of thousands if not millions of years. Language diversity on the other hand dates back at most 100,000 years. All previous attempts to correlate biodiversity and linguistic diversity have simply taken mechanical raw language numbers without consideration for the fact that in some areas, due to both geophysical and sociocultural factors may split at a more rapid or slow rate. It is only deep-level phyletic diversity that correlates in terms of time-depth to processes of

biodiversification. Among currently developed theories, only our Language Hotspots model (inspired by the ground breaking work of Nichols 1986) considers this.

2.0 Modeling language distribution and diversity

Three important factors must be known about a given language in order to situate it within the larger framework of global linguistic diversity. First, what language family/phylum does it belong to, and with what living languages does it share a common origin? Second, how well is the language known to science (e.g., documented, recorded, classified)? And thirdly, just how endangered or robust is it? A contiguous geographic area where we find a clustering of multiple languages with high genetic diversity, high degrees of endangerment, and about which there is a low level of scientific knowledge, is a good candidate for a *Language Hotspot*.

In order to reckon the genetic units necessary to calculate the GENETIC DIVERSITY INDEX of a hotspot, an enormous number of resources have been consulted. This includes the latest in comparative-historical research in a wide range of families and areas, e.g., Niger-Congo, Austronesian, Papuan, Khoisan, indigenous South American, Siberian, Nilo-Saharan, Australian, Mesoamerican and others (e.g., Pawley et al. 2005). A wide range of recent and classic published sources have been referenced for this preliminary component of the research, including works by McMahon and McMahon (2006), Greenberg (2005), Ruhlen (1991, 1994), Voegelin and Voegelin (1977). Human resources in the form of specialists in these domains have been consulted and made a significant contribution to this component of the model.[3] The very latest in the science of typological language sampling (Anderson 2006) also underlies the genetic index that forms one of the three principal numeric indices used to determine *Language Hotspots*.

Languages can be ranked according to the degree to which they are documented and therefore known to science, or what we term RESEARCH INDEX. Expanding a model first developed for languages of Vanuatu (Lynch and Crowley 2001), the basic criteria include in ascending order the existence of writing systems/publications, grammars, dictionaries, texts, linguistic analysis, and finally, accessibly archived audio and video materials.

The ENDANGERMENT INDEX is calculated according to a six-point scale, where 1 is *extinct* and 5 is *healthy*. Level 2 is *moribund*, and indicates that the youngest speakers are roughly 60 years old or older. Level 3 is *seriously endangered* and can be thought of as indicating that the youngest speakers are around 40. Level 4 is *endangered* and indicates that a language shift is underway, and that children are no longer learning the ancestral language as a first language. Level 5 is *threatened*, and indicates that the socio-cultural or demographic circumstances could result in endangerment at any time, perhaps even due to a single historical event or act of nature. Level 6 is *healthy* and indicates a language with a stable or growing population and used by speakers of all ages in most if not all social domains.

[3] Pre-eminent language-area specialists we consulted for this project include Malcolm Ross (La Trobe Univ.), Nick Evans (Australia National Univ.), Scott DeLancey (Univ. of Oregon), Søren Wichmann (Univ. of Copenhagen), Spike Gildea (Univ. of Oregon), Tom Güldemann (MPI-EVA, Leipzig), and Suzanne Romaine (Oxford Univ.) to name just a few.

Calculating the RESEARCH INDEX for each language in a *Language Hotspot* involves extensive study of both legacy sources (Shorto et al. 1963, Key and Key 1967, Leva 1969, Murphy and Goff 1969, Der-Houssikian 1972, Agesthialingom & Sakthivel 1973, Yang 1974, Voorheeve 1975, Coppell 1981), more recent bibliographic sources (e.g. Huffman 1986, Jakupova 1988, Gieser 1989, Troike 1990, La Polla and Lowe 1994, Carrington 1996, Peiros 1998, Tobi 1998, Drolc et al. 1999, Carrington 1999), library collections written in numerous languages, and Internet resources. It also requires prolonged communication with specialists on languages, language families and language areas to gather the latest information on what actually exists and what is available.

The exact boundaries of a *Language Hotspot* depend on a range of topographic, cartographic, geo-political, and practical concerns. Some *Hotspots* largely follow well-known areal linguistic entities, e.g., Northern Pacific Coast-Western Plateau (NPP), East Africa (EAF) or Mesoamerica (MAM). Other *Language Hotspots* are centered in areas of high language density that also happen to be areas of high genetic linguistic diversity. Note that not all areas of high genetic linguistic diversity have large numbers of individual languages, e.g., the two Siberian or Southern Africa hotspots, but many of these languages happen to be endangered and be relatively poorly documented. Still other Language Hotspots emerge simply where there is an overlap of the geographic distribution of the languages of particular areas with respect to the three indices (genetic, endangerment, research). We will continue to refine our methods for drawing the boundaries of hotspots; areal linguistic and topographic and/or geophysical features will play a primary role in this determination.

Underlying the GENETIC INDEX component of the *Language Hotspots* model is over two centuries of scholarly work in linguistic taxonomy. Diversification and degree of relatedness among languages that putatively developed from a common ancestor language is often construed in terms of a phylogenetic, branching tree. The terms of biological taxonomy have been partly co-opted to describe genetic relatedness among languages, such that one talks of language families and phyla. A full taxonomy for the classification of languages narrows according to the following pattern (from largest to smallest units): macro-phylum > phylum > stock > family > language > dialect (> sub-dialect). English is a member of the Germanic family and the Indo-European stock.[4]

Some meta-methodological caveats need to be addressed briefly in this discussion. First, it is not actually possible to prove that two or more languages are actually related to one another. There is simply a preponderance of evidence weighting arguments in favor of one or another genetic relation. On occasion, there are languages —called isolates—that cannot be grouped with any taxonomic unit. Indeed, the higher the level the grouping is, in particular the phylum and

[4] Language "stock" is sometimes also confusingly called "family" in the scientific literature, and part of our agenda in this project is to contribute to standardization of terminology. The higher end taxonomic units (macro-phylum > stock) represent one useful legacy of lexicostatistics. Part of our concern in this project is to contribute to a standardization of terminology for genetic linguistics, for a word that means the level of both Proto-Germanic and Proto-Indo-European, or Proto-Afroasiatic or Proto-Trans-New-Guinea really is too broad.

macro-phylum level, by definition the more tenuous and the more controversial the proposed classification is likely to be.

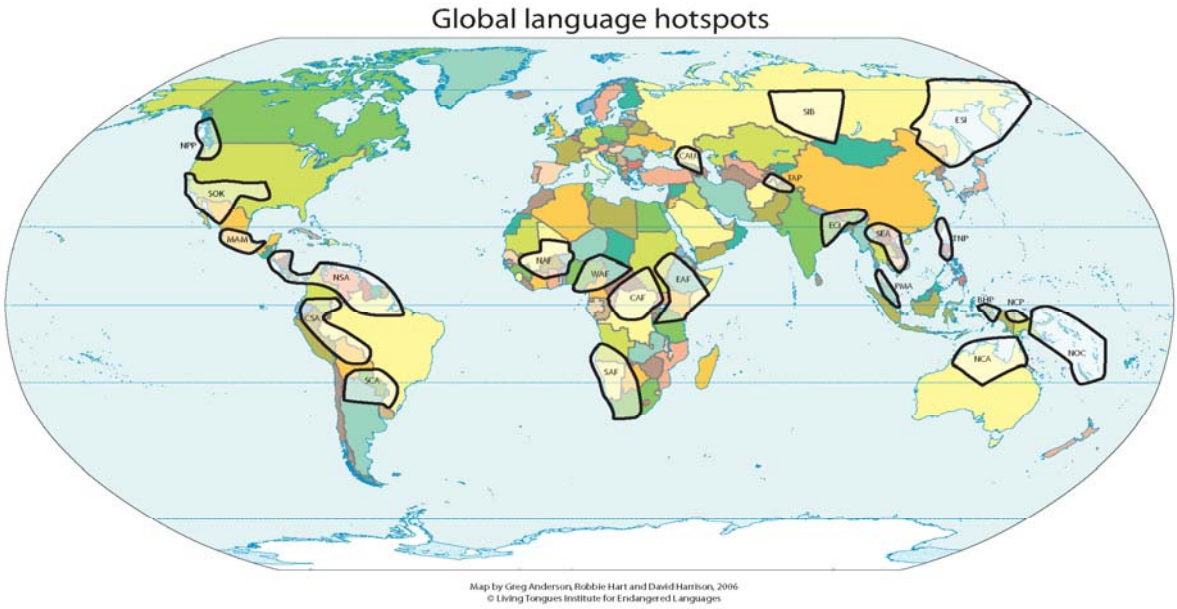
Furthermore, the genetic model is one that allows only one ancestral 'mother' language over time splitting or diverging and yielding one or more daughter languages, through 'natural' transmission between generations. However, in a small number of well-known cases, languages have two parents, or derive from numerous sources, as is the case with most Creole languages and the bizarre 'mixed' languages that have been reported in the literature, including Copper Island (or Mednyj) Aleut, a Russian-Aleut hybrid found in the EASTERN SIBERIA [ESI] hotspot (Golovko and Vakhtin 1990) and the Mbugu language of the EAST AFRICA [EAF] hotspot spoken in Tanzania, which blends Bantu and Cushitic elements (Mous 2004).

Another question that is far from insignificant or easy to overlook is how to accurately assess whether two or more languages are related, and what the internal relationships within a group of putatively related sister languages might be. One basic and widely used method, albeit one that has received a large amount of justifiable criticism, is known as Lexicostatistics, which calculates percentages of cognates for a basic word list across groups of languages, the higher the number, the more closely related the languages (Hymes 1960). However, it should be noted both that any two languages are likely to share chance sound-meaning correspondences of the type that yield the sets of cognates, or shared reflexes or retentions of ancestral items, that form the bases of linguistic families, and that virtually any word can be borrowed when viewing the totality of attested linguistic phenomena.

Researchers engaged in taxonomic investigation are generally aware of these issues and attempt to carefully sift through the totality of available evidence before offering linguistic classifications. These classificatory works have appeared in a number of different languages, and therefore a researcher in this domain must have passable literacy in at least the following languages: English, Russian, German, French, Spanish, and Portuguese. In addition, researchers might need access to materials written in Indonesian, Finnish, Mandarin, Italian, and Vietnamese. In our preliminary research on the Hotspots model we have surveyed thousands of published sources to come up with a state-of-the-art working list of language families that constitute the meaningful units used in the assessment of the overall genetic linguistic diversity of a set of languages, however defined geographically. These of course will be refined as our knowledge of the languages of the world proceeds. It is this working set of approximately 450 linguistic genetic units from which the genetic index in the Language Hotspots model derives.

Our current proposed list of language hotspots—which set the research agenda for this project—includes the following 22 regions of extreme diversity that encompass nearly half the world's 7,000 languages:

<u>Americas</u>			
North Pacific Coast & Western Plateau	NPP	Southwest and Oklahoma	SOK
Mesoamerica	MAM	Northern South & Central America	NSA
Central South America	CSA	South Central South America	SCA
<u>Africa</u>			
South Africa	SAF	East Africa	EAF
Central Africa	CAF	West Africa	WAF
Northwest Africa	NAF		
<u>Eurasia</u>			
Central Siberia	CSI	Eastern Siberia	ESI
Tajikistan, Afghanistan, Pakistan	TAP	Eastern and Central India	ECI
Peninsular Malaysia	PMA	Southeast Asia	SEA
Taiwan, Northern Philippines	TNP		
<u>Australia and Oceania</u>			
Bird's Head Peninsula	BHP	North Central Papua	NCP
Northern and Central Australia	NCA	Near Oceania	NOC



Map courtesy of Living Tongues Institute for Endangered Languages.

The *Language Hotspots* list itself will be modified as new information from field expeditions is incorporated. This will include refinement of the status of individual languages within a proposed hotspot; for example, over the past few years, field expeditions identified a few remaining speakers of languages that had been reported extinct. Some of these languages, upon closer study, proved to be genetic units unto themselves, e.g., *Leco* of Bolivia and *Kusunda* of Nepal, both prematurely pronounced dead. Now is truly the last opportunity to document these languages, as a language once thought dead is clearly not a thriving one.

Not only will languages be definitively confirmed as alive or dead as a result of our expeditions, but the demographic and sociolinguistic information on many dozens of languages will be updated and improved. The current most widely-used collection of language population data, *The Ethnologue* (Gordon 2006), relies largely on tipsters to send in information, and many of its numbers are found to be obsolete when scientists actually visit the speech community. The active collection and surveying techniques we detail herein (see section 2.0) will yield more accurate numbers for many languages, thus allowing us to more accurately assess the threats that hotspot languages now face.

An enormous range of scientific discoveries pertaining to the languages themselves are likely to be made as a direct result of the *Language Hotspots* model. The materials that do exist on many endangered languages are extremely limited, perhaps a single wordlist of 100 or 200 words or just a few sentences. These languages are storehouses of traditional knowledge and a key window into the pre-history of the vast majority of the peoples of the world, and as such, countless invaluable discoveries are awaiting concerned researchers.

Further, we expect to considerably expand our knowledge of the complexity and diversity of human language based on the scientific linguistic discoveries that will accompany a better knowledge of these languages, and with them a refinement of the limits and nature of human cognition. As is well known in the linguistics community, it is precisely these endangered languages that often offer stimulating and quirky data that challenges established linguistic theory. For example, it was once thought that a language could not utilize a basic sentence structure that placed the object first, verb second and subject last. Only after the discovery of this structure in a number of unrelated Amazonian languages (e.g., *Urarina*, 3,000 speakers) did it become clear that such a structure was both cognitively possible and linguistically attested (Nettle and Romaine 2000, Harrison 2007). Countless similar new discoveries are likely to be made while exploring the *Language Hotspots*.

3.0 Improving language mapping

In contemporary linguistics, we have maps of language distribution that are either points or polygons, which we can overlay onto political maps (e.g., Gordon 2005). While useful as a first approximation, such maps omit crucial information like terrain, climate, elevation, water, rainfall, and other the factors that underlie biodiversity and shape linguistic ecologies.

The hotspots approach to linking these three domains is made possible by recent advances in GIS, satellite imaging and cartography. We are not proposing to duplicate commercial whole-world mapping efforts such those of the Ethnologue (www.ethnologue.org) or Global Mapping International (www.gmi.org). By focusing on our identified hotspots we can realistically improve

on the mappings of these well-known databases, producing a more thorough and detailed mapping of the most vulnerable languages, along with their demographic and geo-topographic distribution.

Although the *Language Hotspots* project will be valuable in many ways, it will be particularly important in terms of geographic analysis and visualization. Spatial data on language distribution compiled during the project can and will be analyzed against other thematic data sets. Comparison with Conservation International's biodiversity hotspots, for instance, may reveal correlations between cultural and biological diversity. Shuttle Radar Topography Mission (SRTM) data on terrain and elevation can be juxtaposed with language information to test the hypothesis that physical barriers such as rivers and ridgelines tend to lead to high concentrations of indigenous languages.

Organizing language information by geography also enables it to be visualized in ways that will be compelling to scientists and laypeople alike. Complex interrelationships become easily discernable; both printed and on-screen maps can become the basis for telling the rich and vital story of indigenous languages, and the threatened cultural diversity they represent.[5]

11. Literature cited

Agesthialingom, S. and S. Sakthivel. 1973. *A bibliography of Dravidian linguistics*. Annamalainagar: Annamalai University

Anderson, Gregory D. S. 2006. *Auxiliary Verb Constructions*. Oxford: OUP.

Anderson, Gregory D. S. 2006-ms. Proto-Munda and Proto-Austroasiatic: Problems in diachronic linguistics. Presented at *Workshop on Historical Linguistics and Hunter-Gatherer Populations in Global Perspectives*. Leipzig, August, 2006.

Anderson, Gregory D. S. and K. David Harrison. 2006. Towards a comprehensive documentation of Ös (Middle and Upper Chulym dialects). *Turkic Languages* 10.

Carrington, L. 1996. *A linguistic bibliography of the New Guinea area*. Canberra: ANU.

Carrington, Lois and Geraldine Triffitt. 1999. *OZBIB: A linguistic bibliography of Aboriginal Australia and the Torres Strait Islands*. Canberra: Australian National University.

Conklin, H. C. 1954. The relation of Hanunóo culture to the plant world. PhD. Diss. Yale Univ.

_____ 1980. *Folk Classification: A Topically arranged bibliography of contemporary and background references through 1971*. New Haven, CT: Yale University.

Coppell, W.G. 1981. *Austronesian and other languages of the Pacific and South-East Asia : an annotated catalogue of theses and dissertations*. Canberra; ANU.

Der-Houssikian, Haig. 1972. *A bibliography of African linguistics*. Edmonton/Champaign: Linguistic Research, Inc.

Drolc, Uschi, Caroline Frank and Franz Rottland (compilers), w/ the assistance of Livingstone Walusimbi. 1999. *A linguistic bibliography of Uganda*. Köln: Rüdiger Köppe Verlag.

Fisher, James F. 1990. *Sherpas: Reflections on Change in Himalayan Nepal*. Berkeley: University of California Press.

---

[5] We acknowledge Chief Cartographer Allen Carroll (National Geographic Maps (NGMaps), as the source of the ideas and formulations in the last two paragraphs.

Foale, Simon. 1998. What's in a name? An analysis of the West Nggela (Solomon Islands) fish taxonomy. *SPC Traditional Marine Resource Management and Knowledge Information Bulletin* 9: 3-19.

Froese, R. and D. Pauly (eds.) 2006. *FishBase*. World Wide Web electronic publication. www.fishbase.org, version (06/2006).

Gieser, C. Richard. 1989. *Bibliography of the Summer Institute of Linguistics, Philippines, 1953-1988*. Manila : Summer Institute of Linguistics.

Golovko, Evgenij V. and Nikolaj B. Vakhtin 1990. Aleut in Contact: The CIA Enigma. *Acta Linguistica Hafniensia* 22: 97-125.

Gordon, Raymond G. (ed.) 2005. *Ethnologue: Languages of the World*. Dallas: SIL.

Greenberg, Joseph. 2005. *Genetic Linguistics*. Oxford: Oxford University Press.

Grimes, Barbara F. (ed.). 2000. *Ethnologue: Languages of the World* 14th edition. Dallas: Summer Institute of Linguistics. [Also online at www.ethnologue.com].

Harrison, K. David. 2007. *When Languages Die: The Extinction of the World's Languages and the Erosion of Human Knowledge*. Oxford: Oxford Univ. Press.

Haugen, E. 1972. *The ecology of language*. Stanford: Stanford University Press.

Heywood, V.H (ed.) 1995. *Global Biodiversity Assessment*. Cambridge: CUP.

Huffman, F. E. 1986. *Bibliography and index of mainland Southeast Asian languages and linguistics*. New Haven: Yale University Press.

Hymes, Dell H. 1960. Lexicostatistics so far. *Current Anthropology* 1: 3-44.

Hyslop, Catriona. 1999. The linguistics of inhabiting space: spatial reference in the North-East Ambae language. *Oceania* 70: 25-42.

Jakupova, G. K. 1988. *Bibliografija po tatarskomu jazykoznaniju: 1778-1980*. Kazan': TKI.

Key, Harold and Mary Key 1967. *Bolivian Indian tribes; classification, bibliography, and map of present language distribution*. Norman, OK: Summer Institute of Linguistics of the University of Oklahoma

Kroeber, A.L. 1925. *Handbook of the Indians of California*. Washington: Government Printing Office. [Bulletin of the Smithsonian Institution, Bureau of American Ethnology 78].

La Polla, Randy J. and John B. Lowe. 1994. *Bibliography of the International Conferences on Sino-Tibetan Languages and Linguistics I-XXV*. Berkeley : Sino-Tibetan Etymological Dictionary and Thesaurus Project, Center for Southeast Asia Studies, Univ. of California.

Le Page du Pratz, Antoine-Simon. 1975 [1774]. *The History of Louisiana*. Joseph G. Tregle, Jr. (ed.) Baton Rouge: Louisiana State University Press. [Facs. of 1774 London edition].

Leva, Antonio Enrico. 1969. *Il Contributo italiano alla conoscenza delle lingue parlate in Africa*. Rome : Istituto poligrafico dello Stato.

Lizarralde, Manuel. 2001. Biodiversity and loss of indigenous languages and knowledge in South America. In Luisa Maffi (ed.) 2001, pp. 265-281. *On Biocultural Diversity*. Washington, D.C.: Smithsonian.

Lynch, John and Terry Crowley. 2001. *Languages of Vanuatu: A new survey and bibliography*. Canberra: Pacific Linguistics.

Magga, Ole Henrik. 2005. Indigenous Knowledge Systems—The True Roots of Humanism. Talk delivered at the World Library and Information Congress, Oslo, Norway, August 2005.

Marine Biological Laboratory. 2004. Universal Biological Indexer and Organizer. Online at http://www.ubio.org/. Accessed January 2006.

McMahon, A. and R. McMahon. 2006. *Language Classification by Numbers*. Oxford: OUP.

Mittermeier, R. A., Myers, N., Mittermeier, C. G, Gil, P. R., Fonseca, G.A.B., Konstant, W. R., Mast, R. B., Thomsen, J. B., Bowles, I. A., Olivieri, S., Ayres, J. M. & Hannah, L. 1999. Hotspots and Global Biodiversity Conservation. In: Mittermeier, R. A., Myers, N., Gil, P. R. & Mittermeier, C. G. (eds.), *Hotspots: Earth's Biologically Richest and Most Endangered Terrestrial Ecoregions*. pp. 21-67. Mexico: CEMEX.

Mittermeier, Russell A., Patricio Robles Gil, Michael Hoffman, John Pilgrim, Thomas Brooks, Cristina Goettsch Mittermeier, John Lamoreux, and Gustavo A.B. da Fonseca. Preface by Peter A. Seligmann. 2005. *Hotspots Revisited*. Washington, DC: Conservation Int'l.

Mous, Maarten. 2004. *The making of a mixed language: The case of Ma'a/Mbugu*. Amsterdam: John Benjamins.

Mufwene, Salikoko. 2001. *The Ecology of Language Evolution*. Cambridge: CUP.

Müllhäusler, E. 1996. *Linguistic ecology*. New York: Routledge.

Murphy, John D. and Harry Goff 1969. *A bibliography of African languages and linguistics*. Washington: Catholic University of America Press.

Myers, N., R. Mittermeier, N. Mittermeier, da Fonseca, and Kent (2000). Biodiversity hotspots for conservation priorities. *Nature* 403: 853-858.

Nabhan, Gary Paul. 2003. *Singing the Turtles to Sea: The Comcáac (Seri) Art and Science of Reptiles*. Berkeley, CA: Univ. of California Press.

Nettle Daniel. 1998. Explaining global patterns of linguistic diversity. *J. Anthropol. Archaeol.* 17:354–74A.

_____ 1999. *Linguistic Diversity*. Oxford: Oxford University Press.

Nettle, Daniel and Suzanne Romaine 2000. *Vanishing Voices: The Extinction of the World's Languages*. Oxford: Oxford Univ. Press.

Pawley, Andrew, R. Attenborough, J. Golston and R. Hide (eds.) 2005. *Papuan Pasts, cultural, linguistic and biological histories if Papuan-speaking peoples*. Canberra: ANU.

Peiros, Ilia 1998. *Comparative Linguistics in Southeast Asia*. Canberra: ANU.

Riley, S. J., S. D. DeGloria and R. Elliot. 1999. A terrain ruggedness index that quantifies topographic heterogeneity, *Intermountain Journal of Sciences* 5: 1-4.

Ruhlen, Merritt. 1991. *Guide to the World's Languages, vol. 1: Classification*. Stanford: Stanford University Press.

_____ 1994. *On the Origin of Languages: Studies in Linguistic Taxonomy*. Stanford: Stanford University Press.

Shorto, H. L., Judith M. Jacob and E. H. S. Simonds (compilers). 1963. *Bibliographies of Mon-Khmer and Tai linguistics*. London/New York: Oxford Press.

Smith, H. H. 1932. *Ethnobotany of the Ojibwe*. Bulletin of the Public Museum of the City of Milwaukee, Milwaukee, Wisconsin. 4:(3)327-525.

Stepp, J. R., S. Cervone, H. Castaneda, A. Lasseter, G. Stocks and Y. Gichon. 2004. Development of a GIS for Global Biocultural Diversity. *Policy Matters*. 13 Nov. 2004. 267-272.

Stepp, J. R., S. Cervone, and H. Castaneda. 2005. Mountains and Biocultural Diversity. *Mountain Research and Development*. 25(3)223-227.

Sutherland, William J. 2003. Parallel extinction risk and global distribution of languages and species. *Nature* 423: 276-279.

Teshome, A., B. R. Baum, L. Fahrig, J. K. Torrance, T. J. Arnason, and J. D. Lambert. 1997. Sorghum [*Sorghum bicolor* (L.) Moench] landrace variation and classification in North Shew and South Welo, Ethiopia. *Euphytica* 97: 255-263.

Tobi, Sueyoshi. 1998. *A bibliography of Nepalese languages and linguistics*. Kathmandu: Central Department of Linguistics, Tribhuvan University.

Troike, Rudolph C. 1990. *Bibliography of bibliographies of the languages of the world*. Amsterdam: John Benjamins.

United Nations Environment Program. 2001. Globalization threat to world's cultural, linguistic and biological diversity. News Release 1/18/2001. Online at < http://www.unep.org/Documents.multilingual/Default.asp?DocumentID=192&ArticleID =2765. Accessed August 2005.

Voegelin, C. F. and F. M. Voegelin. 1977. *Classification and Index of the World's Languages*. New York: Elsevier.

Voorheeve, C. L. 1975. *Languages of Irian Jaya: checklist, preliminary classification, language maps and wordlists*. Canberra: ANU.

Wichmann, Søren and Arpiar Saunders (forthcoming) How to use typological databases in historical linguistic research. *Diachronica*.

Wilson, E. O. 2003. The encyclopaedia of life. *Trends in Ecological Evolution* (18): 77–80.

Yang, Paul Fumian. 1974. *Chinese linguistics: a selected and classified bibliography = Zhongguo yu yan xue fen lei can kao shu mu*. Hong Kong: Chinese Univ. of Hong Kong.