

Tools from evolutionary biology shed new light on the diversification of languages

Stephen C. Levinson^{1,2} and Russell D. Gray³

¹ Max Planck Institute for Psycholinguistics, Post Office Box 310, 6500 AH Nijmegen, The Netherlands

² Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Kapittelweg 29, 6525 EN Nijmegen, The Netherlands

³ Department of Psychology, University of Auckland, Auckland 1142, New Zealand

Computational methods have revolutionized evolutionary biology. In this paper we explore the impact these methods are now having on our understanding of the forces that both affect the diversification of human languages and shape human cognition. We show how these methods can illuminate problems ranging from the nature of constraints on linguistic variation to the role that social processes play in determining the rate of linguistic change. Throughout the paper we argue that the cognitive sciences should move away from an idealized model of human cognition, to a more biologically realistic model where variation is central.

Variation is the key

Evolutionary scientists study variation. A key element of the Darwinian revolution was the insight that variation within species was not some superficial noise that should be stripped away to reveal the underlying species essence [1]. Variation is the signal – over evolutionary time variation *within* species becomes variation *between* species. As a legacy of the 1950s Cognitive Science movement, cognitive scientists have often thought of language on a par with vision, olfaction or memory – a human faculty with a universal organization, subject only to minor variation. But compared to animal communication systems one of the most remarkable things about human language is that there are 7000 of them, and they are culturally variable at every level of their structure, from the sound system, to the grammar, to the semantics. In this review we explore why this is, what has driven the diversification of languages, and how these processes can be systematically studied. Rather than viewing this diversity as noise distorting underlying principles, we argue that this variation is a vital resource for understanding the crucial capacity that makes us human, and that tools derived from evolutionary biology give us powerful new ways of analyzing this variation [2–4].

Why do languages vary? Darwin [5] pondered this point, noting the curious parallels between languages and species, and indeed similar processes of speciation, drift, and adaptation can be observed in the language domain. Processes of group boundary formation account for change under demographic pressures, drift accounts for change

by geographic or social isolation, and adaptation for the changes that can be observed as languages reflect the cultural uses to which they are put (with e.g. color words reflecting the technology of dye and paint [6], kinship terms the systems of marriage, residence and inheritance [7], and highly embedded sentences the growth of literacy and its specialist genres [8]).

These observations are not new (although there is much recent literature providing new data and insight). What is new, however, is the recent development of highly sophisticated computational tools for exploring these processes of diversification. What these tools make possible is the meticulous inference of detailed past processes from the current diversity in all its complexity. This is time travel of a kind we could in the past only have dreamed of. The tools have their origin in evolutionary biology, and although originally designed for the inference of phylogenetic relationships from DNA or biological morphology, they allow the extraction of historical process from any kind of structural pattern that is subject to evolutionary processes, whether cultural or biological. They allow, for example, the inference of the evolution of stone tool assemblages [9], which in turn allow the inference of cognitive processes by our hominin ancestors [10].

What constrains linguistic variation?

The power of these new computational tools can be illustrated by a recent application to a central question in linguistics: To what extent is linguistic diversification and variation bounded by strong universal constraints? Generative linguists following Chomsky have assumed that the great unseen bulk of the structure of languages is universal, specified in Universal Grammar or UG, and variation may be circumscribed by a limited set of ‘parameters’ or binary switches, which in various combinations multiply out the observed variation [11,12]. Linguistic typologists follow Greenberg [13] in proposing that there are strong tendencies for specific features to clump together, so limiting variation. A classic example in both approaches is word order. Greenberg noted that in a sample of 30 languages the position of the verb vis-à-vis its object seems to control other word order features, especially the order of adpositions (prepositions if before the noun, postpositions if after), and other nominal elements like adjectives and determiners. A worldwide

Corresponding author: Levinson, S.C. (stephen.levinson@mpi.nl).

sample of about 1000 languages seems to support these constraints as relatively strong tendencies across language families and geographic areas [14], suggesting universal cognitive biases for ‘harmonic’ word orders within a language. In commenting on the “instructive and influential” work of Greenberg, Chomsky remarked, “These universals are probably descriptive generalizations that should be derived from principles of UG”[15], a view developed by Baker [11].

However, applying the new phylogenetic methods to the claims about universal functional dependencies in word-order suggests that at least part of the pattern is an artifact of historical process. As Galton pointed out in 1889 [16], sampling from a pool of languages or cultures is unlikely to yield statistically independent measures. Languages are linked through nested patterns of descent. If we are interested in functional linkage between features, we need to control for the possession of features shared due to inheritance rather than intrinsic linkage (see Figure 1). Traditionally, linguistic typologists try to do this by sampling *across* languages that are not closely related or geographically proximate, but they are hampered by the fact that a large proportion of the world’s languages are grouped within a few large families, while the relation between others is lost in time. If evolutionary biologists want to test whether, say, two genes are functionally linked, they use powerful Bayesian computational methods to infer the correlated gains and losses of genes *within* family trees (phylogenies) of organisms, thus controlling for shared inheritance [17].

Bayesian procedures are becoming increasingly familiar in the cognitive sciences where they have been used to model brain processes [18], reasoning [19], perception [20], language learning [21] and processing [22], but their application to historical inference may be less familiar to this readership. Recently, we applied Bayesian phylogenetic methods to detect systematic co-variation in eight word-order features within four of the world’s largest language families [23]. Our goal was to assess whether there is causal linkage between these features, so if one changes another will. Two stages of analysis are required. First, given that phylogeny really matters, we needed language family trees that objectively represent phylogeny based on a standard measure, namely changes in vocabulary. Ideally what one wants is not a single tree but rather a set of trees that reflect the uncertainty associated with the inferred branching patterns and the amount of evolution along each branch. In this way the inferences can be integrated over uncertainty in estimation of the trees. We used Bayesian methods to construct these trees from standard wordlists of basic vocabulary. The items of basic vocabulary were coded into sets of cognate (homologous) words, the performance of different models of cognate gain and loss evaluated, and then the best model used in a Markov Chain Monte Carlo search to infer the posterior distribution of the trees.

Given the trees, the second step was to understand how different branches of a language family might have evolved different word orders. This involves reconstructing the most likely word orders all the way up each of these trees, through all the ancestral nodes, given the observed word-order features in all the descendant languages. The key

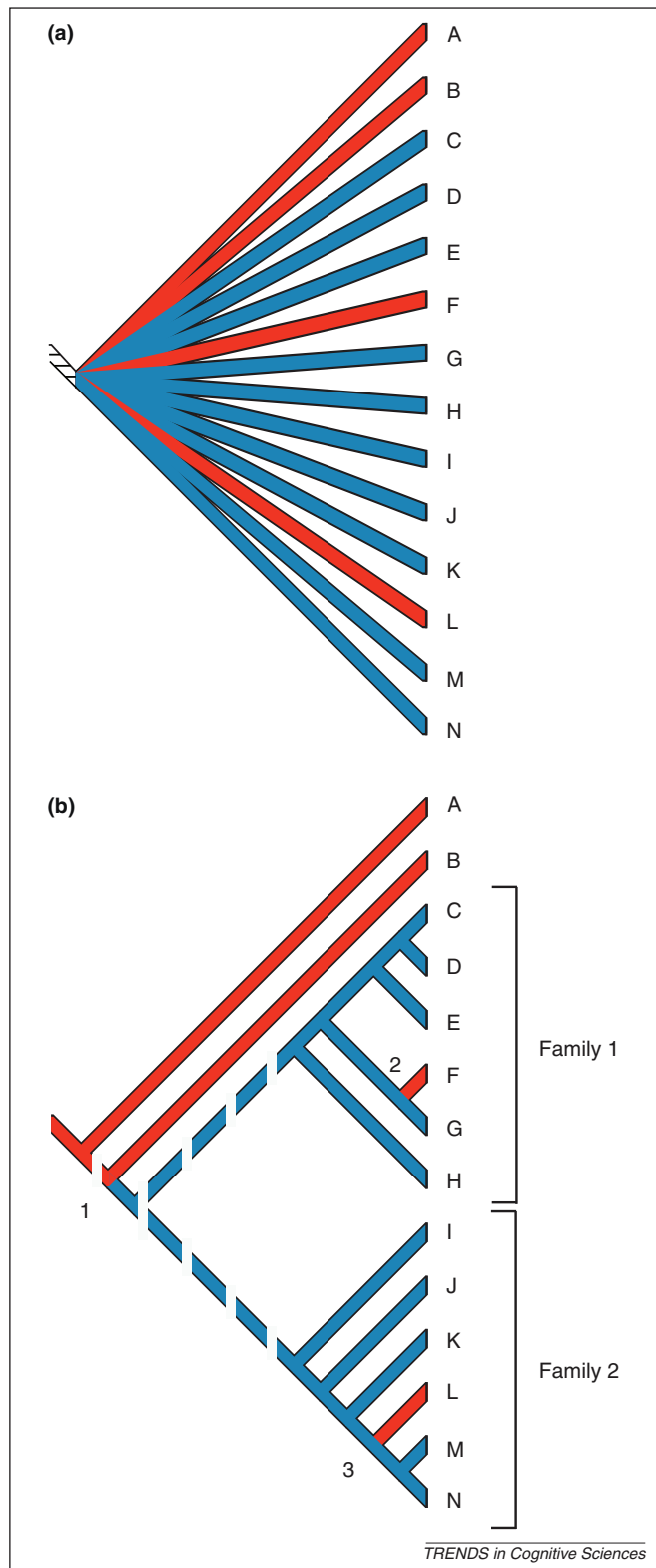


Figure 1. Why phylogeny matters. If languages were all equally related (a), then each language would provide a statistically independent sample. However, languages are linked through nested patterns of descent (b), some of which may be beyond our current ability to resolve (dashed lines). The phylogenetic dependencies mean that each language cannot be treated as an independent sample (Galton’s problem). Sub-sampling does not solve this problem. For example, sampling language E from language family one and language K from family two would not result in independent samples because the blue state is inherited from a common ancestor. Moreover, it would miss the changes in state to red that occurs in languages L and F, and consequently loose statistical power. The best solution to Galton’s problem is to infer the actual independent changes on the phylogeny (changes marked 1, 2 and 3).

Review

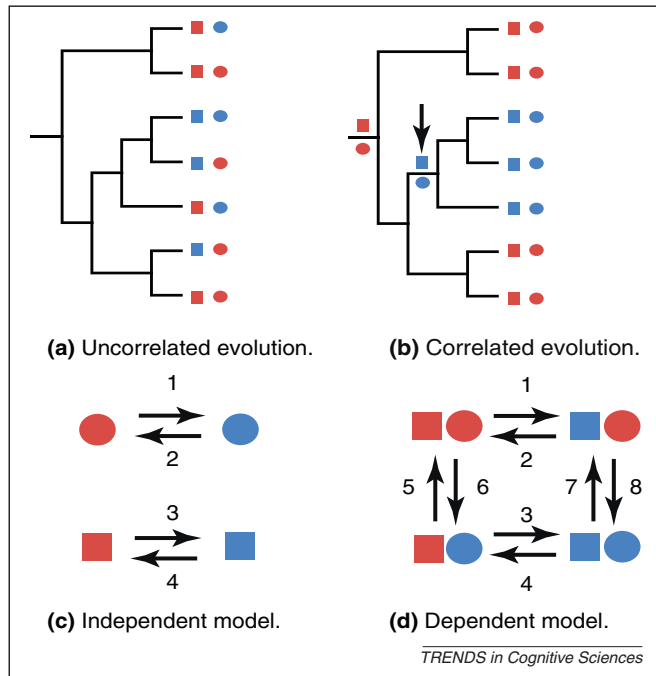


Figure 2. Analyzing correlated evolution on phylogenetic trees. (a) shows the evolution of two hypothetical traits (squares and circles) on a phylogenetic tree. The evolution of the states in these traits (red and blue) is not linked. (b) shows a phylogenetic tree with correlated character state changes. The changes are marked with an arrow pointing to the branch of the tree they are most likely to have occurred on. (c) shows the four transition rates that are estimated in the independent (uncorrelated) model of trait evolution. (d) shows the eight transition rates that are estimated in the dependent (correlated) model of trait evolution. If there is correlated evolution then the dependent model should fit better.

question was whether change in one word order (e.g. the position of the verb with respect to its object) would have knock-on effects on other word orders (yielding e.g. prepositions before nouns or postpositions after them), as would be expected if Greenberg's or Chomsky's views are correct. To answer this we fitted two models of evolution for pairs of word-order features. These models work by calculating the transition rates between word-order states along the phylogeny. The first model was an independent model, which assumes that there is no correlation between the pair of word-order features (the rates are estimated separately). The second model was a dependent model in which the transition rates between word-order states is jointly estimated (see Figure 2). If there is correlated evolution of word-order features then the second model should fit better.

We quantified the difference in fit between the models with Bayes factors and found little support for the second model, which would capture the causal linkage implied by universal constraints. There are indeed features that covary within language families, but they are not consistently the *same* features across the four language families. For example, only Austronesian showed strongly coupled changes between numeral–noun and genitive–noun orders. Even when we did find dependencies that were shared between language families, examination of the transition probabilities between linked states suggests that different historical patterns of change are responsible for the observed linkage in each language family. We drew the conclusion that, although languages are obviously hedged

in by what the human mind can learn and process [24], word-order variation is constrained by historical factors rather than by a set of faculty-specific mental constraints or parameters (but see [25,26] for alternative perspectives).

Global patterns in language diversity

Languages vary enormously in the number of phonemes (contrastive sounds) they possess. Some languages such as Rotokas in the Pacific have just a dozen contrastive sounds, whilst others such as the San languages in Africa have over 100. Recently, Quentin Atkinson [27] applied ideas and methods derived from the study of human genetic diversity to global patterns of phonological diversity. Human genetic diversity is greatest in Africa and declines with increasing distance from Africa. This pattern fits a 'serial founder effect' model in which small populations on the edge of an expansion progressively lose diversity as each hiving-off subgroup carries just a small proportion of the original population's diversity. In an analysis of 504 languages across the globe Atkinson found that the number of phonemes is also highest in Africa and tends to decrease with increasing distance from Africa. The fewest phonemes are found in South America and on tropical islands in the Pacific Ocean. In general, the areas of the globe that were most recently colonized show fewer phonemes in the local languages, whereas the areas that have hosted modern humans for millennia (particularly sub-Saharan Africa) still, on average, use the most phonemes. There are some obvious exceptions to the pattern Atkinson reports. Some of the world's largest phoneme inventories are found in the Caucasus and among Na Dene languages in North America, which was colonized relatively late. It is also not clear whether the serial founder effect model applies to phoneme inventories. Human populations are normally genetically diverse but speech communities typically have little variation in their phoneme inventories. In cases where there have been well attested serial founder events, such as the Austronesian expansion into the Pacific from Taiwan, the largest phoneme inventories are not located in Taiwan but rather in the languages of New Caledonia and the Loyalties [28]. Although the general pattern Atkinson found has been replicated using other data and methods [29–31], the findings are vulnerable to slight changes in assumptions [32], other gradients can be found and a distance-from-Africa gradient might well be due to other causes [30]. Regardless of the outcome of this debate, Atkinson's application of evolutionary thinking is likely to stimulate many more efforts to explain global patterns of linguistic diversity by drawing on the parallels between the processes of biological and linguistic diversification.

Processes of language diversification

Darwin's "curious parallel" between speciation and language diversification receives modern support from the very close correspondence between biological diversity and cultural diversity on the surface of the planet (both seem to shadow the distribution of rainfall [33]). Although linguists, following Chomsky, have tended to emphasize abstract underlying language universals, new information about linguistic diversity suggests a different emphasis:

the fact is that languages differ fundamentally across all the layers of structure, from sound systems, syntax and morphology to semantics ([34,35]).

Phylogenetic methods and computational simulations of historical process open up many avenues for the investigation of the processes lying behind language diversification. A crucial question is why are there so many human languages, while nearly all other animals (barring the song birds) have fixed repertoires? Here the parallels to biological evolution are clearly relevant (see also [Box 1](#)):

- **Founder effects:** when a small sub-population hives off from a larger one, it carries only a small part of the larger population's variation, so further changes start from a different base.
- **Drift:** Isolated sub-populations lose mutual intelligibility over several hundred years due to the accumulation of 'mutations', 'copy errors' and other divergent changes (sound changes, semantic shifts, grammatical innovations)
- **Hybridization:** sub-populations come under the influences of different neighbours speaking their own languages or dialects ('language contact').
- **Adaptation:** sub-populations in new or changing ecological niches change behaviour in the direction of maximizing genetic fitness in the new environment; a parallel process in cultural evolution matches technological, cultural and linguistic behaviour to new niches (e.g. 'semantic shift' in kinship terms reflects changes in social structure [36]). While semantic and lexical

Box 1. Similarities and disparities between cultural and biological evolution

There are two questions that are recurrently asked about the nature of cultural evolution: Does the speed, and variable rate of change, put cultural evolution into a quite different bracket than genetic evolution? Secondly, does the degree of borrowing through contact with other cultures and languages constitute a quite different kind of process to that of biological evolution? The new methods from the biological sciences are once again proving a major source of insight here:

- **Rates of evolution:** Phylogenetic reconstructions of language history seem to be able to reach right back to the beginning of the Holocene or 10,000 years ago, before the development of agriculture, pottery or metallurgy [23,67], [68,69]. These deep reconstructions rely on the selection of features that change only slowly – such as the kin terms, some body parts, words for the sun, moon, fire, and the lower numerals. If structural features (such as word order, phoneme inventories, morphology) can be found that are equally resistant to change (see [46]), it might be possible that even deeper prehistory can be tapped [70].
- **Networks vs. trees:** It is true that horizontal transmission is perhaps less in evidence in the genetic evolution of higher vertebrates than it is in language, but it is the norm amongst the viruses and bacteria. New methods are being developed in evolutionary biology to model networks [71,72]. From the beginnings of historical linguistics in the nineteenth century, it has been controversial whether phylogenetic trees adequately capture the processes of language diversification. While phylogenetic inferences can be robust to moderate levels of undetected borrowing [73], and claims about the degree of borrowing are sometimes overstated [74], many commentators have still claimed that horizontal inheritance will overwhelm vertical inheritance. The new biological methods now give linguistics ways to quantify these two forces thus promising an empirical solution to this old controversy.

innovation and shift can be obvious responses to, e.g., new technology or new environments, the form side of language is also under adaptive change: Zipf's 'law of brevity' [37] describes the ineluctable reduction of word forms in response to frequency of usage, while dialect differences change to reflect social class or group formation [38].

But there may be other factors specific to language and cultural evolution. For example, human groups under pressure often exaggerate the differences to make ethnic barriers – a process Gregory Bateson dubbed 'schismogenesis' [39]. Whereas in other species drift may make mating signals unrecognizable, thereby hastening speciation, humans may artfully change the signals, thus engendering cultural sub-speciation as it were (see [40] for a classic study). Atkinson et al. [41] recently used phylogenetic methods to quantify the impact of schismogenesis. They used basic vocabulary data to construct phylogenies for the Austronesian, Bantu, Indo-European language families. Their results revealed that between 10 to 33% of the vocabulary differences in these families arose during rapid bursts of change associated with language-splitting events (see [Figure 3](#)).

In general, in cultural and linguistic evolution demographic factors play an important but complex role. Recent findings suggest that technological and linguistic evolution respond differently to demographic pressure: Small populations tend to be technologically simple [42] but linguistically complex especially in morphology [43], while large ones tend to be technologically complex and linguistically simple (a linguistic leveling supporting communication across widespread sub-populations) – a sort of "paradox of cultural evolution". But small language groups seem to react differently according to the number and demographic size of their neighbours: small populations under threat by numerous neighbours are probably more likely to employ 'schismogenesis' (or 'esoterogeny') [43,38]. To add to the puzzle, different levels of linguistic structure seem to respond differently to population increase, with morphology simplifying [43], but inventories of phonemes increasing [44].

The relation between cultural and biological evolution has often been thought about in terms of biology evolving a platform for cultural take-off, with cultural and technological development then free to fly its own course. But this is probably the wrong model. Instead, variable population genetics continuously affords, constrains and probably channels cultural evolution [45]. It seems likely for example that population variation in audition and vocal tract anatomy seed cultural variation in phoneme inventories – certainly a good case has been made for a robust correlation between genes and the contrastive use of tone, as in tone languages [46,47]. Very slight genetic differences in distinct populations would be sufficient to bias the probabilities of cultural and linguistic solutions of different sorts, as shown by both modeling and experimental work on cultural transmission [48,49].

Communication systems may have driven cognitive development in various species [50]. There is growing consensus that there is a substantial nonlinguistic cognitive infrastructure to human language, consisting of the

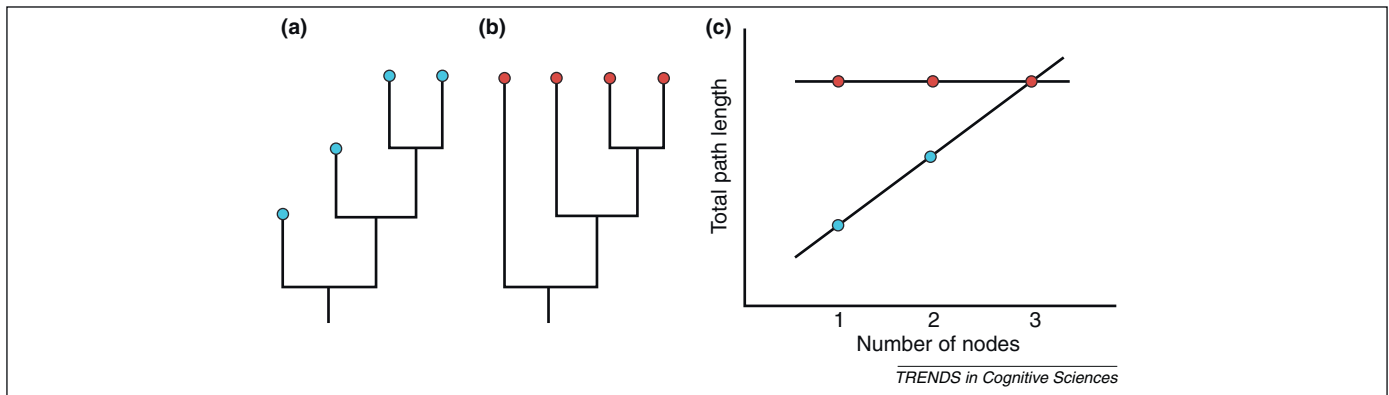


Figure 3. How might social processes affect historical patterns of language change? If speech communities exaggerate differences at the time when they are drifting apart (schismogenesis), then lineages that have been through more splitting events will undergo more change. The phylogenetic tree in (a) shows the 'punctuational' pattern this process would produce – the branch lengths are longer in lineages that have been through more splitting events. In contrast, if the rates of change are not affected by the number of splitting events, then all the tips of all the branches will be equal irrespective of the number of splitting events the lineage has been through (b). The size of any schismogenic effect can be quantified by plotting the path length from the root of the tree to each of its tips against the number of nodes (splitting events) the path goes through (c). The slope of the resulting graph estimates the magnitude of the effect [41].

motivational structure, the cooperative basis, the theory of mind and universal interactional ethology [51,52,53]. Understanding how and when this evolved can best be approached by phylogenetic means, involving the testing of specific cognitive capacities of the existing species within our family, and then inferring the changes on ancestral nodes [54]. Thus even though we are hampered by having no good animal models for language due to the loss of our sister taxa (the extinct hominids), we can still make excellent use of the existing variation to aid our reconstruction of the evolution of language (see Box 2). In the meantime, the cracking of ancient DNA, together with increasingly sophisticated interpretation of fossils, makes it look ever more likely that the antiquity of language stretches much further back than has been thought.

Why should cognitive scientists, concerned mostly with timescales on the millisecond level, have an interest in

Box 2. Outstanding questions

- Can phylogenetic methods lift the current 10,000 year ceiling on inferences about deep language links, either by improving vocabulary methods or by using structural features of languages?
- Can the new ways of analyzing lateral gene transfer and hybridization developed in biology be usefully employed to model cross-linguistic and cross-cultural diffusion?
- If sociolinguistics is to historical linguistics what microevolutionary studies are to studies of macroevolution, can large-scale historical linguistic patterns be modeled from our knowledge of sociolinguistic processes?
- Can phylogenetic methods be applied to primate cognition to map the evolution of human cognition?
- Can we quantify the relative roles of factors governing rates of linguistic and cultural evolution – e.g. demography and competition for resources, divergence events (e.g. migration, cultural contact), societal complexity, subsistence mode?
- Can computational methods derived from evolutionary biology be used to model sound change and automate cognate judgments [75,76]?
- Can the methods of population genetics be imported wholesale into the study of linguistic, cultural and psychological variation?
- More generally, how should the whole of the cognitive sciences be brought within an evolutionary framework, where variation stands central?

these issues of historical process measured often in centuries and millennia? The reasons are numerous:

- We can't compare human language to other animal models without understanding the cultural processes that play a part in language. The crucial role of vocal imitation in human development is without parallel among the primates, although interesting parallels can be found among the oscine birds ([55]).
- Cultural plasticity, with its potential for fast change, limits the extent to which the language capacity can be fixed in the genotype, [56,49].
- Evolutionary processes draw on synchronic variation in the population: no variation, no evolution. If the cognitive sciences are to engage with evolutionary thinking, they will have to move away from the emphasis on idealized capacities towards a focus on variation in cognitive performance [57]. We will need to stop talking about human memory, language processing, or reasoning as if it were a single uniform system, and really engage with the study of human expertise and individual differences. Fundamental differences in language processing and brain structure correlate with gender or familial handedness, for example [58,59]. These developments are likely to be pushed by the increasing successes of imaging genomics, where population differences in genetics can be correlated with variation in brain structure and function (see [60]).
- Language variation breeds cognitive diversity. Different languages require different processing algorithms, so that there are systematic differences in brain activation according to the language being spoken, e.g. Chinese involves more bilateral activation in word recognition [61]. Different spatial language correlates with different spatial memory strategies [62]. Given brain plasticity, speakers of different languages are likely to have slightly different brains [63]. Literacy, for example, substantially rewires the brain [64,65].
- Language diversification offers fundamental insight into the wide spectrum of human capacities: a natural laboratory for expertise effects in the communicative domain – compare sign language to tone languages,

languages with no nasal phonemes (like m or n) to ones with 50 different nasal phonemes, languages with morphology (like Latin or Finnish) to those with little or none (like Chinese) [34].

Language stands as proxy for other aspects of cultural diversification in cognition. As a compelling recent commentary [66] reminds us, the Western student subject pool that forms the basis for nearly all modern psychological findings is in many respects untypical and culturally specialized. The current loss of linguistic and cultural diversity threatens not only the scientific study of language but also our understanding of human cognitive potential. It is time for the cognitive sciences to take serious stock of the cognitive variation that fascinated earlier generations of psychologists, from Luria and Vygotsky to Cole and Scribner, from W. H. Rivers to Segall and Jahoda.

Thus, although the study of the cultural evolution of language may seem far removed from studies of cognition, this research delivers an important perspective on human performance from a God's eye view. For history is nothing but the study of behavior as largely shaped by myriads of human decisions. Powerful new techniques for reconstructing historical process allow us to data-mine the vast encyclopedia of current human behavior patterns to extract the cognitive biases and evolutionary processes that have served over millennia to shape the present.

References

- 1 Lewontin, R.C. (1983) Darwin's revolution. *New York Review Books* 30, 21–27.
- 2 Felsenstein, J. (2004) *Inferring phylogenies*, Sinauer
- 3 Lemey, P. et al., eds (2009) *The Phylogenetic Handbook: A Practical Approach to Phylogenetic Analysis and Hypothesis Testing*, Cambridge University Press
- 4 McMahon, A. and McMahon, R. (2006) *Language classification by numbers.*, Oxford University Press
- 5 Darwin, C. (1871) *The descent of man*, John Murray
- 6 Levinson, S. (2000) Yéli Dnye and the theory of basic color terms. *J. Linguist. Anthropol.* 10, 3–55
- 7 Friedrich, P. (1964) Semantic structure and social structure: an instance from Russian. In *Explorations in cultural anthropology* (Goodenough, W., ed.), McGraw-Hill
- 8 Karlsson, F. (2007) Constraints on multiple center-embedding of clauses. *J. Linguist.* 43, 365–392
- 9 Lycett, S.J. (2011) 'Most beautiful and most wonderful': Those endless stone tool forms. *J. Evol. Biol.* 9, 143–171
- 10 Stout, D. (2011) Stone tool making and the evolution of human culture and cognition. *Philos. Trans. R. Soc. Lond. Ser. B: Biol. Sci.* 366, 1050–1059
- 11 Baker, M.C. (2001) *Atoms of language: The mind's hidden rules of grammar*, Basic Books
- 12 Di Sciullo, A. and Boeckx, C. (2011) *The biolinguistic enterprise: New perspectives on the evolution and nature of the human language faculty*, Oxford University Press
- 13 Greenberg, J. (1963) *Universals of Language*, Cambridge University Press
- 14 Dryer, M.S. (2011) The evidence for word order correlations. *Language Typology* 15, 335–380
- 15 Chomsky, N. (1998) Minimalist Inquiries: The Framework. *MIT Occasional Papers in Linguistics* 15
- 16 Naroll, R. (1961) Two solutions to Galton's Problem. *Philos. Sci.* 28, 15–29
- 17 Barker, D. et al. (2007) Constrained models of evolution lead to improved prediction of functional linkage from correlated gain and loss of genes. *Bioinformatics* 23, 14–20
- 18 Friston, K. (2010) The free-energy principle: A unified brain theory? *Natl. Rev. Neurosci.* 11, 127–138
- 19 Tenenbaum, J.B. et al. (2006) Theory-based Bayesian models of inductive learning and reasoning. *Trends Cogn. Sci.* 10, 309–318
- 20 Battaglia, P.W. et al. (2003) Bayesian integration of visual and auditory signals for spatial localization. *J. Opt. Soc. Am.* 20, 1391–1397
- 21 Bannard, C. et al. (2009) Modelling children's early grammatical knowledge. *PNAS* 106, 17284–17289
- 22 Chater, N. and Manning, C. (2006) Probabilistic models of language processing and acquisition. *Trends Cogn. Sci.* 10, 287–291
- 23 Dunn, M. et al. (2011) Evolved structure of language shows lineage-specific trends in word-order universals. *Nature* 473, 79–82
- 24 Christiansen, M. and Chater, N. (2008) Language as shaped by the brain. *Behav. Brain Sci.* 31, 489–558
- 25 Cysouw, M. (2011) Understanding transition probabilities. *Linguist. Typol.* 15, 415–432
- 26 Croft, W. et al. (2011) Greenbergian universals, diachrony, and statistical analyses. *Linguist. Typol.* 15, 433–454
- 27 Atkinson, Q.D. (2011) Phonemic diversity supports a serial founder effect model of language expansion from Africa. *Science* 332, 346
- 28 Blust, R. (2009) *The Austronesian languages*, Australian National University
- 29 Jaeger, T.F. et al. (2011) Mixed effect models for genetic and areal dependencies in linguistic typology. *Linguist. Typol.* 15, 281–320
- 30 Maddieson, I. et al. (2011) Geographical distribution of phonological complexity. *Linguist. Typol.* 15, 267–280
- 31 Wichmann, S. et al. (2011) Phonological diversity, word length, and population sizes across languages: The ASJP evidence. *Linguist. Typol.* 15, 177–198
- 32 Cysouw, M. et al. (in press). Still no evidence for an ancient language expansion from Africa. *Science*
- 33 Nettle, D. (1999) *Linguistic Diversity*, Oxford University Press
- 34 Evans, N. and Levinson, S. (2009) The myth of language universals: Language diversity and its importance for cognitive Science. *Behav. Brain Sci.* 32, 429–492
- 35 Levinson, S. and Evans, N. (2010) Time for a sea-change in linguistics. Response to comments on 'The myth of language universals'. *Lingua* 120, 2733–2758
- 36 Friedrich, P. (1966) The linguistic reflex of social change: From Tsarist to Soviet Russian kinship. *Sociol. Inquiry* 36, 159–185
- 37 Zipf, G.K. (1936) *The psycho-biology of language: An introduction to dynamic philology*, Routledge
- 38 Trudgill, P. (2011) *Sociolinguistic Typology: Social Determinants of Linguistic Complexity*, Oxford University Press
- 39 Bateson, G. (1935) Culture contact and Schismogenesis. *Man* 35, 178–183
- 40 Gumperz, J.J. and Wilson, R. (1971) Convergence and creolization: A case from the Indo Aryan/Dravidian border in India. In *Pidginization and creolization* (Hymes, D., ed.), pp. 151–167, Cambridge University Press
- 41 Atkinson, Q.D. et al. (2008) Languages evolve in punctational bursts. *Science* 319, 588
- 42 Henrich, J. (2004) Demography and cultural evolution: How adaptive cultural processes can produce maladaptive losses: The Tasmanian case. *American Antiquity* 69, 197–214
- 43 Lupyan, G. and Dale, R. (2010) Language structure is partly determined by social structure. *PLoS ONE* 5, e8559
- 44 Hay, J. and Bauer, L. (2007) Phoneme inventory size and population size. *Language* 83, 388–400
- 45 Dediu, D. (2011) Are languages really independent from genes? If not, what would a genetic bias affecting language diversity look like? *Hum. Biol.* 83, 279–296
- 46 Dediu, D. (2011) A Bayesian phylogenetic approach to estimating the stability of linguistic features and the genetic biasing of tone. *Proc. R. Soc. Lond. B: Biol. Sci.* 278, 474–479
- 47 Dediu, D. and Ladd, R. (2007) Linguistic tone is related to the population frequency of the adaptive haplogroups of two brain size genes, ASPM and Microcephalin. *PNAS* 104, 10944–10949
- 48 Kirby, S. et al. (2008) Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proc. Natl. Acad. Sci. U.S.A.* 105, 10681–10686
- 49 Smith, K. (2011) Learning bias, cultural evolution of language and the biological evolution of the language faculty. *Hum. Biol.* 83, 261–278

- 50 Shultz, S. and Dunbar, R.I.M. (2010) Encephalization is not a universal macroevolutionary phenomenon in mammals but is associated with sociality. *PNAS* 107, 21582–21586
- 51 Enfield, N.J. and Levinson, S., eds (2006) *Roots of human sociality - culture, cognition and interaction*, Berg
- 52 Levinson, S. (2006) On the human 'interactional engine'. In *Roots of human sociality. Culture, cognition and human interaction* (Enfield, N.J. and Levinson, S., eds), pp. 39–69, Berg
- 53 Call, J. and Tomasello, M. (2008) Does the chimpanzee have a theory of mind? 30 years later. *Trends Cogn. Sci.* 12, 187–192
- 54 Haun, D.B.M. *et al.* (2006) Cognitive cladistics and cultural override in Hominid spatial cognition. *PNAS* 103, 17568–17573
- 55 Ohms, V.R. *et al.* (2010) Zebra finches exhibit speaker-independent phonetic perception of human speech. *Proc. R. Soc. Lond. B: Biol. Sci.* 277, 1003–1009
- 56 Chater, N. and Christiansen, M. (2010) Language acquisition meets language evolution. *Cogn. Sci.* 34, 1131–1157
- 57 Levinson, S. (in press) The original sin of cognitive science. *Top. Cogn. Sci.*
- 58 Catani, M. *et al.* (2007) Symmetries in human brain pathways predict verbal recall. *PNAS* 104, 17163–17168
- 59 Townsend, D.J. *et al.* (2001) Familial handedness and access to words, meaning and syntax during sentence comprehension. *Brain Lang.* 78, 308–331
- 60 Papassotiropoulos, A. and Stephan, D.A. *et al.* (2006) Common Kibra Alleles Are Associated with Human Memory Performance. *Science* 314, 475–478
- 61 Valaki, C.E. *et al.* (2004) Cortical organization for receptive language functions in Chinese, English and Spanish: A cross-linguistic MEG study. *Neuropsychologia* 42, 967–979
- 62 Haun, D.B.M. *et al.* (2011) Plasticity of human spatial cognition: Spatial language and cognition covary across cultures. *Cognition* 119, 70–80
- 63 Allen, J. *et al.* (2008) Morphology of the Insula in Relation to Hearing Status and Sign Language Experience. *J. Neurosci.* 28, 11900–11905
- 64 Petersson, K.M. *et al.* (2007) Literacy: A cultural influence on functional left-right differences in the inferior parietal cortex. *Eur. J. Neurosci.* 26, 791–799
- 65 Carreiras, M. *et al.* (2009) An anatomical signature for literacy. *Nature* 461, 983–986
- 66 Henrich, J. *et al.* (2010) The weirdest people in the world? *Behav. Brain Sci.* 33 (2–3), 61–83
- 67 Gray, R.D. and Atkinson, Q.D. (2003) Language-tree divergence times support the Anatolian theory of Indo-European origin. *Nature* 426, 435–439
- 68 Pagel, M. *et al.* (2007) Frequency of word-use predicts rates of lexical evolution throughout Indo-European history. *Nature* 449, 717–720
- 69 Pagel, M. (2009) Human language as a culturally transmitted replicator. *Nat. Rev. Genet.* 10, 405–415
- 70 Dunn, M. *et al.* (2005) Structural phylogenetics and the reconstruction of ancient language history. *Science* 309, 2072–2075
- 71 Collins, J. *et al.* (2011) Quantifying hybridization in realistic time. *J. Comput. Biol.* 18, 1305–1318
- 72 Nelson-Sathi, S. *et al.* (2011) Networks uncover hidden lexical borrowing in Indo-European language evolution. *Proc. R. Soc. Lond. B: Biol. Sci.* 278, 1794–1803
- 73 Greenhill, S.J. *et al.* (2009) Does horizontal transmission invalidate cultural phylogenies? *Proc. R. Soc. Lond. B: Biol. Sci.* 276, 2299–2306
- 74 Bowern, C. *et al.* (2011) Does lateral transmission obscure inheritance in hunter-gatherer languages? *PLoS ONE* 6, e25195
- 75 Boucharde-Côté, A. *et al.* (2009) Improved reconstruction of protolanguage word forms, In *NAACL09*, 65–73
- 76 Lydia, S. *et al.* (2011) A pipeline for computational historical linguistics. *Lang. Dyn. Change* 1, 89–127