Biol Philos (2017) 32:855–886
https://doi.org/10.1007/s10539-017-9594-y

Curiously the same: swapping tools between linguistics and evolutionary biology

Lindell Bromham[1]

Received: 22 March 2017 / Accepted: 13 September 2017 / Published online: 20 September 2017

Abstract One of the major benefits of interdisciplinary research is the chance to swap tools between fields, to save having to reinvent the wheel. The fields of language evolution and evolutionary biology have been swapping tools for centuries to the enrichment of both. Here I will discuss three categories of tool swapping: (1) conceptual tools, where analogies are drawn between hypotheses, patterns or processes, so that one field can take advantage of the path cut through the intellectual jungle by the other; (2) theoretical tools, where the machinery developed to process the data in one field is adapted to be applied to the data of the other; and (3) analytical tools, where common problems encountered in both fields can be solved using useful tricks developed by one or the other. I will argue that conceptual tools borrowed from linguistics contributed to the Darwinian revolution in biology; that theoretical tools of evolutionary change can in some cases be applied to both genetic and linguistic data without having to assume the underlying evolutionary processes are exactly the same; and that there are practical problems that have long been recognised in historical linguistics that may be solved by borrowing some useful analytical tools from evolutionary biology.

Keywords Language evolution · Historical linguistics · Biological evolution · Darwinism · Galton's problem · Phylogenetic non-independence · Spatial autocorrelation · Interdisciplinary

> The formation of different languages and of distinct species, and the proofs that both have been developed through a gradual process, are curiously the same. Darwin (1871) Vol 1, 1st edition.

✉ Lindell Bromham
Lindell.Bromham@anu.edu.au

[1] Macroevolution and Macroecology, Division of Ecology and Evolution, Research School of Biology, Australian National University, Canberra, ACT 0200, Australia

Language evolution and biological evolution are obviously very similar: both can be described by a process of descent with modification leading to hierarchical patterns of similarity between lineages, both involve the rise in frequency of individual differences until they form characteristic differences between populations, both generate diversity by splitting of lineages which may also suffer extinction. Language evolution and biological evolution are also obviously very different: where biological evolution acts by the changing frequency of traits inherited by offspring from their parents, language evolution involves changing patterns of intangible sounds, images and concepts, passed between both related and unrelated individuals of any age by any medium, which can be both unconsciously adopted and wilfully changed.

When do the similarities allow us to use the same tools to investigate both language and biological evolution? When do the differences make swapping tools dangerously misleading? In this paper, I will argue that as long as we are clear what task we are using the tools for, the fundamental differences between language evolution and biological evolution will in some cases be less important than they first appear. I will restrict the discussion to a comparison between evolutionary biology and historical linguistics (the study of language change over time), rather than considering linguistics as a whole, or the origin of language itself.

To be clear on the job we are doing with our tools, it might help to split them into several different kinds—conceptual tools, theoretical tools, and analytical tools—though the distinctions between theses categories may not be clear in all cases. Conceptual tools are analogies drawn between different processes, where the rhetorical device of comparison between similar examples is used to bolster the argument for the reasonableness of one case by reference to the other. Theoretical tools are models that can be used to investigate patterns and processes, using simplified and indirect representations of the systems to explore the influence of different factors. Analytical tools are practical applications, useful pieces of kit that can be profitably employed outside the field they were originally developed in.

These three kinds of tools seem to have had different patterns of swapping between historical linguistics and evolutionary biology. Conceptual tools have a long history of two-way exchange, with traffic back and forth between evolutionary biology and historical linguistics. Theoretical tools have thus far predominantly been borrowed in one direction, from biology to linguistics. And while some analytical tools have moved between fields, I will highlight some areas where there has been a puzzling lack of transfer, despite the development in biology of practical solutions to problems shared with historical linguistics.

Conceptual tools

Conceptual tools are analogies drawn between different fields that allow people to move along a path already cut through the intellectual jungle. In this sense, conceptual tools are rhetorical devices whose primary purpose is persuasion of the plausibility of a proposed process by demonstrating the operation of that process in another sphere. As analogies, conceptual tools do not rely on inferring homologous mechanisms, merely on making an idea more palatable by drawing a likeness to another more broadly accepted example (Croft 2008).

Borrowing conceptual tools from historical linguistics was a recurring element in the early promulgation of evolutionary biology. Darwin was not the first to suggest that species had evolved, and he was not the first to see the potential for natural selection to effect change in a population over time, but he was the first to make a detailed and logical case for connecting individual heritable variation to population change to the generation of biological diversity. Prosecuting the Darwinian revolution involved convincing people that the variation we can observe among individuals within populations is sufficient to drive the formation of immense differences between living species and to thus generate exuberant diversity over time. Studies of language change over time provided conceptual tools that evolutionary biology could borrow to make the case for Darwinism more convincing, by offering palatable evidence for both descent with modification and gradualism.

Evidence for descent with modification was effectively undeniable for languages. The study of philology had a long academic history and was well established by the late eighteenth Century, when comparison of Sanskrit to Latin and Greek led Sir William Jones, an English judge serving in Calcutta, to the conclusion that "no philologer could examine all three without concluding that they had all sprung from a common source" (Jones 1788). While much philology had been focused on the history of the Indo-European languages, as people travelled the globe, they not only brought back information on the distribution of species but also the similarities and differences between languages (Prichard 1832). For example, one of the books carried in the Beagle's library was John Hawkesworth's account of Cook's voyages, published in 1773, which included a comparison of basic word lists from languages such as Malay, Javanese and languages of the South Seas, concluding that they had all sprung from a common root, yet were also puzzlingly similar to the language used in far-distant Madagascar by speakers of apparently different race (Hawkesworth 1773)—a conundrum now solved by evidence that Madagascar was colonized by Austronesian speaking people who had travelled across the Indian Ocean (Kusuma et al. 2015).

Not only the historical patterns of descent but also the processes of change were examined in widely-read works on philology. For example, Jacob Grimm's work on German grammar provided one of the first case studies of systematic sound change, where a number of shifts in consonants could be reconstructed over time. Darwin refers in his notebooks to the ideas of philologist John Horne Tooke, who, publishing in the late eighteenth and early nineteenth centuries, endeavoured to trace the origin of words through changing historical use, such that, in Darwin's words, "most every language shows traces of anterior state" (Darwin 1838).[1] The process of descent with modification had been also been recognised in textual analysis. Stemmatic methods were used to infer lines of descent of texts by using

---

[1] Horne Tooke also used his philological work to provide a rather quirky legal defence in which he claims to have been made the victim in a court of law of two prepositions and a conjunction (Horne Tooke 1840).

occasional changes to reconstruct the chain of copies. In the early nineteenth century Karl Lachmann had developed a formalised approach to reconstructing ancient texts based on genealogical relationships between existing copies, which, for example, he used to reconstruct a single ancestral archetype of Lucretius' Epicurean poem *De rerum natura*.

There was then in the first half of the nineteenth century an intellectual milieu around the ideas of both species transformation and language genealogy. The development of neither field depended on the other: ideas of language evolution had preceded the formal introduction of evolutionary theory, and the evidence for biological evolution was drawn from geology and natural history not directly from studies of language. But the exchange of ideas stimulated both fields.

Here, I will focus only on arguments concerning gradual evolutionary change within languages, not whether the faculty of language itself was acquired by many small steps or a big leap (Berwick and Chomsky 2015). Of course, Darwin was compelled to argue for a gradual acquisition of language capability (Darwin 1838), and most of Darwin's early notes on language evolution are concerned with solving the puzzle of how thinking, speaking humans could be derived from dumb brutes by degrees (Alter 2007). But Darwin also showed an early interest in attempts to use gradualist explanations of language change to infer human history (Darwin 1837) as well as philological explanations of how language could be gradually modified over time to give "a radical diversity of tongues" (Darwin 1838). As noted by geologist Sir Charles Lyell, the process of evolutionary change in both species and languages could be studied without addressing the issue of origins: "we may enquire into the link of descent and hereditary transmission and the mode of variation and of growth without pretending to solve the problem of creation. Just as we may endeavour to investigate the rise and growth of languages and their success extinction without pretending to define or comprehend what may be the nature of that inventive power by which man gives origin to new terms and words" (Lyell 1860).

The interchange of ideas between evolutionary biology and linguistics is symbolised by the parallel adoption of the same iconography in both fields. August Schleicher, inspired by Darwin's Origin of Species, made the analogy between languages and species in great detail (Schleicher 1869). Darwin then drew on Schleicher's work in his Descent of Man (Darwin 1871). Darwin and Schleicher both used branching diagrams to display the process of descent, and Ernst Haeckel, inspired by Schleicher's language phylogenies, developed the tree iconography further, embedding the concept of an evolutionary tree at the heart of biology (List et al. 2016; Richards 2002). Schleicher pointed out that while Darwin's branching diagram was theoretical, his phylogeny of Indo-European languages was empirical (Schleicher 1869). This emphasizes the point that the evidence for demonstrable patterns of descent was considered to be much stronger for languages than the evidence for the theory of biological evolution, which was by many considered to represent the possible rather than the actual (Richards 2002).

In the famous quote given at the beginning of this paper, Darwin notes not only how the processes of formation of languages and species are curiously the same but also that the *proofs* that they are the products of gradual change are also the same. The empirically well-supported case for language change as a product of descent with modification was used to make the argument that species were likewise related by descent more convincing, providing a plausible analogy for gradual change driving evolution and divergence. For example, a key argument in the Origin of Species is that hierarchical taxonomic classification is not an arbitrary naming system but a natural reflection of descendant relationships between species. Darwin argued that all of the various classification schemes invented by naturalists have unconsciously reflected their genealogical history: "community of descent is the hidden bond which naturalists have been unconsciously seeking, and not some unknown plan of creation, or the enunciation of general propositions, and the mere putting together and separating object more or less alike" (Darwin 1859). Darwin draws an analogy to linguistics, particularly to make the point that relative amounts of change are subordinate to hierarchical relationships in such a classification: "It may be worth while to illustrate this view of classification, by taking the case of languages. If we possessed a perfect pedigree of mankind, a genealogical arrangement of the races of man would afford the best classification of the various languages now spoken throughout the world; and if all extinct languages, and all intermediate and slowly changing dialects, had to be included, such an arrangement would, I think, be the only possible one. … The various degrees of difference in the languages from the same stock, would have to be expressed by groups subordinate to groups; but the proper or even only possible arrangement would still be genealogical; and this would be strictly natural, as it would connect together all languages, extinct and modern, by the closest affinities, and would give the filiation and origin of each tongue" (Darwin 1859).

Language evolution not only provided undeniable evidence for the process of descent with modification, it also provided a clear case for gradualism, the central claim of Darwin's theory. For languages, a convincing case could be made for gradual accumulation of differences over time leading to distinct lineages. Indeed, Schleicher regarded the study of language evolution as providing proof of Darwin's theory of descent with modification by demonstrating that the principles he espoused could indeed generate diversity: "The kinship of the different languages may consequently serve, so to speak, as a paradigmatic illustration of the origin of species, for those fields of inquiry which lack, for the present at least, any similar opportunities of observation" (Schleicher 1869). The analogy with language change was presented to sweeten the pill of gradualism, by providing a case where gradual change could be actually witnessed in action. "Ancient and extinct forms of life often show combined or intermediate characters, like the words of a dead language with respect to its several offshoots or living tongues" (Darwin 1868).

The importance of the analogy with language evolution for the acceptance of Darwinian gradualism can be appreciated by considering the transformation of Sir Charles Lyell's views on biological evolution. It was upon Lyell's Principles of Geology that Darwin built, by analogy, his uniformitarian approach to biology. Like Lyell's geological explanations of the formation of the physical worlds, Darwin claimed that the features of the biological world and the changes of the past could and should be explained using only processes known to operate in the present day. Darwin had discussed his theory with Lyell by 1856, and Lyell strongly encouraged Darwin to publish (Wilson 1970). Although Lyell saw the value of Darwin's hypothesis, he was not, like Thomas Henry Huxley or Ernst Haeckel, an enthusiastic convert. But, appropriately for the leading uniformitarian of the day, Lyell came by gradual changes to eventually accept Darwinian evolution.

To read Lyell's notebooks is to witness someone struggling with the intellectual upheaval associated with accepting the new world view imposed by Darwinism. Lyell had read Lamarck in 1827, but, like most of his contemporaries, he rejected species transformation. As a geologist, Lyell had plenty of evidence of changing biota over time, but this did not require him to believe that species themselves could change. His rejection of species transformation was partly justified on empirical grounds, because he did not feel the imperfect record provided by fossils was able to support such a conclusion. But, privately, he was also uncomfortable with the logical conclusion of accepting that species had descended from other species would require one to "go the whole orang" and accept the origin of humans as descended from a primate progenitor, and that placing humans in the animal family tree would not only challenge "the position of Man in the Universe" but also diminish the dignity of the Creator (Bartholomew 1973; Wilson 1970). Instead, Lyell believed that species were so closely adapted to their particular environments that environmental change caused extinction and hence temporal species turnover.

Lyell's commitment to uniformitarian explanations balked at accepting a mode of change if he could not see the mechanism currently in operation. Thus Lyell's rejection of species transformation was also partly on mechanistic grounds, given that the "variety-making machinery" which fed the process was poorly understood. And here analogies to language change may have helped him to accept species change despite the lack of understanding of the mechanism of generation of variety. In 1860, Lyell wrote in his notebooks on the species question: "We may enquire into the link of descent and hereditary transmission and the mode of variation and of growth without pretending to solve the problem of creation. Just as we may endeavour to investigate the rise and growth of languages and their successive extinction without pretending to define or comprehend what may be the Nature of that inventive power by which Man gives origin to new terms and words" (Lyell 1860).

Though he was a reluctant transformationist, Lyell was a committed gradualist, and he argued against any theory for the origin and diversity of language that require special creation or unknown forces. Instead, in his notebooks, he explores how languages may diversify slowly over time: "The longer any particular type of language has existed, the greater will have been the number of its dialects and their divergence from parent stock, not only by desuetude but by new inventive additions, some of a higher, some of a simpler kind" (Lyell 1860). This brings Lyell to Darwin's point that a continual process of descent with modification can create a hierarchy of related forms with distinct differences between them: "Language may be historically and continuously connected, yet distinct" (Lyell 1860). He emphasises the similarity between the geographical evidence for language origin with the same form of evidence used to argue for species evolution: every language must have a single origin in time and space, and cannot be independently creating in two places, and once extinct cannot come back to life. This is clearly related to Alfred Russel Wallace's observation that species come into being connected in time and space with closely allied species (Wallace 1855), an insightful observation which apparently had nagged at Lyell since its publication in 1855 (Costa 2013; Wilson 1970). So Lyell's notebooks bear witness to him using analogy to language evolution to convince himself of the reasonableness of biological evolution: "As any language has been in some way connected with a preceding one so may every species. Languages may improve and may degenerate, may change with comparative rapidity, or be persistent for indefinite periods—and once extinct can never reappear. All this may be true of species" (Lyell 1860).

By the time Lyell published "Geological Evidences of the Antiquity of Man" in 1863, he had become an advocate for Darwinian evolution. In Huxley's words "Lyell, up to that time a pillar of the anti-transmutationists (who regarded him, ever afterwards, as Pallas Athene may have looked at Dian, after the Endymion affair), declared himself a Darwinian, though not without putting in a serious caveat. Nevertheless, he was a tower of strength, and his courageous stand for truth as against consistency, did him infinite honour" (Huxley 1887). Lyell devoted an entire chapter in the Antiquity of Man to the evolution of languages, dressing in linguistic cloth the same arguments that Darwin had rehearsed in the Origin of Species in order to link individual variation to the generation of diversity. Lyell points out that a natural historian arguing for species transformation might encounter the same objections to the theory as an imagined philologist explaining the theory of language transmutation to a naïve audience, such as the lack of intermediates between contemporary languages, the absence of records of past states, the disagreement as to the exact divisions between dialect and language (Lyell 1863). He then demonstrates how the philologist might use all of these observations to support his theory of transformation, as well as pointing to the genealogical relationships between related languages both living and dead, and the documented change in any given language over time. Thus the argument for descent with modification is of the same kind for both languages and species.

But Lyell goes further than merely pointing out that languages, like species, show evidence of descent with modification. He also uses analogy with language change to argue for the mechanism of Darwinian gradualism (which these days we would refer to as a microevolutionary explanation of macroevolutionary change). Language change provides an ideal conceptual tool because it appeals to an individual's every day experience of the adoption of small changes in language over their lifetime. "But another important question still remains to be considered, namely, whether the trifling changes which can alone be witnessed by a single generation, can possibly represent the working of that machinery which, in the course of many centuries, has given rise to such mighty revolutions in the forms of speech throughout the world. Everyone may have noticed in his own lifetime the stealing in of some slight alterations of accent, pronunciation or spelling, or the introduction of some words borrowed from a foreign language to express ideas of which no native term precisely conveyed the import. He may also remember hearing for the first time some cant terms or slang phrases, which have since forced their way into common use, in spite of the efforts of the purist. But he may still contend that "within the range of his experience", his language has continued unchanged, and he may believe in its immutability in spite of minor variations…. If the numerous words, idioms, and phrases, many of them of ephemeral duration, which are thus invented by the young and old in various classes of society, … could all be collected together and put on record, their number in one or two centuries might compare with the entire permanent vocabulary of the language”. Lyell’s detailed discussion of how observed features of language diversity can arise by the gradual accumulation of small changes, combined with the differing fates of populations and movements of languages in time and space, rewards reading today by its clarity in setting out an argument for macroevolutionary patterns arising from microevolutionary change.

So Charles Lyell, who had previously publicly rejected species transformation, was within a few years of the publication of Origin of Species using language evolution to make a cogent argument for descent with modification and Darwinian gradualism. Darwin himself then quoted Lyell’s arguments in The Descent of Man, though here the explanatory arrow flips direction again: Darwin uses the same arguments of analogy between language change and species transformation—homologies of form due to common descent, presence of rudiments, hierarchical classification, geographic patterns and extinction—to strengthen his claim for the gradual diversification of all languages from a common source rather than languages originating by a special act of creation (Darwin 1871).

In this manner, the conceptual tool of gradual accretion of change within populations leading to both difference and divergence was swapped back and forth between historical linguistics and evolutionary biology, borrowing the reasonableness of the process in one field to support the reasonableness of the same conclusion in the other. Note that this is an argument from analogy, not an appeal to homology. No one needs to claim that language evolution is in any way mechanistically the same as biological evolution: Darwin and Lyell point out that language change may be driven by fashion, social structure or educational policies (Darwin 1871; Lyell 1863). Instead, the appeal is to the plausibility of the explanation: just as we can see that languages change and diverge over time to create distinct varieties, so might we accept that species could likewise change over time. Both Lyell and Schleicher felt that a more convincing case for Darwinian evolution could be made with reference to evidence from historical linguistics, which should then in turn help to convince people of the reality of Darwinian evolution for the origin of species (Taub 1993). Darwin also used analogy of languages to species to make a palatable argument for gradualism for one of the most persistent challenges to his evolutionary theory, the origin of human intelligence and communication (Darwin 1871).

The borrowing of conceptual tools between historical linguistics and evolutionary continues to be useful (List et al. 2016; Whitfield 2008). For example, the appropriateness of representing historical change as a tree has been challenged by drawing analogies between lexical borrowing and horizontal gene transfer (Atkinson and Gray 2005; List et al. 2014). There are many processes in both language and biological evolution that do not result in a nested pattern of two-way splits, such as hybridization (the merger of previous separate species or languages to form a new biological lineage or contact language) and horizontal transfer (genes or words moving between populations). While the same terms may be used to describe these non-bifurcating processes, they are analogies and not homologies. For example, while Schleicher likened the influence of a foreign language to crossbreeding in biology (List et al. 2014), he cannot have considered that the adoption of foreign words was by the same mechanism as hybridization through sexual reproduction. But the implications of recognizing these very different processes are similar in both historical linguistics and evolutionary biology. In both fields, the fitting of data to a bifurcating model of descent has often led to removal of any data that show evidence of lateral transfer between lineages. Genes that do not show strictly vertical patterns of descent are typically excluded from phylogenetic studies, and loan words are removed from linguistic datasets before phylogenies are constructed (Greenhill and Gray 2012; Starostin 2000). But borrowings are not only legitimate historical signal, they may be as informative of the lineages' history and modes of change than those traits that track the main line of descent. So the value of the analogy is in highlighting patterns that may be inadequately reflected in classic tree-like depictions, rather than identifying homologous processes of exchange between lineages.

Linguists, unlike evolutionary biologists, have never been able to ignore horizontal transfer of language elements, and thus the bifurcating tree model has traditionally been less enthusiastically embraced in historical linguistics. Identification of the source of loan words has been used to trace language evolution (see examples in Durkin 2014; Evans 2011; Kirch 2001), and examples of language hybridization have helped to answer fundamental questions concerning language evolution (e.g. Blasi et al. 2017; DeGraff 2005; McWhorter 2001). By comparison, evolutionary biologists have been generally slower to embrace horizontal transfer as being anything more than a nuisance. Whereas the discombobulating effect of horizontal gene transfer was once seen as inhibiting the inference of the true history of the major branches of life, it is now regarded by some as being the main story rather than the distraction, leading to arguments that a bifurcating tree is a false representation of much of the evolutionary history of life on earth (Doolittle 1999; Martin 2011). Genes that moved to the nucleus from organelle genomes that had previously been expunged from phylogenetic analyses are now seen as important objects of study in their own right (Keeling and Palmer 2008). Discordance between phylogenies built from different genes, traditionally dismissed as noise rather than signal, has in some cases revealed interesting evolutionary histories, such as the transfer of genes between parasites and their hosts (Mower et al. 2004). Just as the analogy to language evolution was made in the early history of evolutionary biology to argue for descent with modification and gradualism, so the analogy between language borrowing and horizontal gene transfer is being explicitly drawn in order to make a case for a new approach to phylogenetics (Gray et al. 2010; List et al. 2014). In this way, a conceptual tool has paved the way for the adoption and sharing of new analytical tools.

Theoretical tools

Theoretical tools can be considered to be indirect representations of a problem in a framework that makes exploration of patterns and processes tractable and informative (Weisberg 2007). The form of theoretical tool that had been commonly swapped between biology and linguistics is an evolutionary model (Croft 2008). A model is a stripped down representation of a process, calling on a few stylised components and a small number of basic mechanisms that can be used to explore the behaviour of a system under particular conditions. It is the simplicity and abstraction of many models that allows theoretical tools developed to solve one particular problem to be repurposed to investigate different phenomena.

Many different evolutionary models have been applied to understanding language evolution, including processes of random change and selection (e.g. Baxter et al. 2006; Blythe 2012). The use of evolutionary models to study the process of language evolution has been questioned on the grounds that language evolution is insufficiently like genetic evolution to justify use of the same theoretical tools (Hauser et al. 2014; Marris 2008). But models do not have to be an accurate reflection of reality to be useful. Indeed, simple abstract models may be more useful tools than those that reflect the complex reality of biological or linguistic change (Smith 2014). For each theoretical tool we wish to use, the relevant question is whether the theoretical tool performs the job we ask of it for a particular data set.

As an example, consider the role of population genetic models in the development of evolutionary biology. Mathematical modeling of population genetics was dismissed by the biologist Ernst Mayr as “beanbag genetics”—depicting the complex processes of inheritance and selection as if they were as simple as drawing coloured beans from a bag. Mayr claimed that these simple beanbag models had contributed little in the way of conceptual progress (Dronamraju 2011; Rao and Nanjundiah 2011). This dismissal was unfair in several ways, not least of which is the failure to acknowledge that these models revolutionised evolutionary biology (Crow 2008): it also ignored the attention that the architects of the models had given to the complicating factors such as epistasis, linkage and dominance (Dronamraju 2011). But, perhaps more fundamentally, Mayr’s criticism mischaracterized the goal of models of evolution, which is not to mimic reality but to allow exploration of the possible outcomes of proposed processes (Crow 2001). Mayr stated that it was wrong to think that it was possible to arrive at the truth through mathematical formulae (see Rao and Nanjundiah 2011). But if modeling genes as beans in a bag give us the right answer—that it, it allows us to make useful predictions about an evolving system—then that theoretical tool is serving the purpose for which it was designed.

The debate over the utility of beanbag genetics is relevant to the consideration of the role of population genetic models to study language evolution. Clearly, language evolution differs from population genetics in many important ways, including mechanisms of generation of variation and modes of transmission and adoption of new forms. But, given that a model is a simplified and indirect representation of specific processes, we need to think clearly about what the purpose of the model is before we can decide whether differences between languages and species will impact on the utility of the model.

For example, randomness of mutation is often considered a pillar of evolutionary theory (Lenski and Mittler 1993). Genetic variation is generated by accidental changes to the genome, so new mutations arise without regard to need. Most changes are likely to be either without benefit or harmful, but, by chance, some mutations will improve an organism's chances of survival and reproduction. Natural selection filters these random changes, because mutations that reduce the chances of successful reproduction are less likely to be passed on to the next generation. In contrast, generation of variation in language evolution need not be random. While language change can be accidental and unconscious, it can also be deliberate and directed. Words may be invented in direct response to a communication need. Pronunciation may be deliberately modified to align with or distinguish from a particular group of speakers. Grammar may be changed by diktat. Given the potential for directed generation and adoption of variation in language evolution, how can we hope to use genetic models that are designed for systems with random, undirected mutation?

But to regard random mutation as a *sine qua non* of evolutionary models is to confuse an 'is' for an 'ought'. Population genetic models work on the relative fitness of mutations, regardless of their ultimate source. Natural selection filters out deleterious mutations, and promotes those with fitness advantages. Even if some mechanism evolved that produced proportionately fewer deleterious mutations, or more advantageous ones, the process by which the advantageous mutations rose in frequency until they were fixed in the population, or the deleterious ones reduced in frequency until they disappeared, would presumably be the same. The robustness of Darwinian evolutionary biology to the violation of the randomness of mutation is in some way illustrated by the fact that Darwin himself was not wedded to the idea of random generation of variation. While he considered most variation arose irrespective of need, he considered there was evidence for several additional non-random mutational processes—inheritance of acquired characteristics and environmentally-stimulated mutation (Darwin 1858, 1859).

Darwin's evidence on this point, drawn from a global network of correspondents, may no longer be considered reliable proof of non-random mutation. But claims of non-random generation of variation are not uncommon in more recent times. For example there was a flurry of excitement when empirical claims were made for directed mutation (Cairns et al. 1988), heritable changes that arise at greater frequency in response to pressing need, though further experimentation suggested that the pattern may be due to transient hypermutation under stress—still very interesting, but still random with respect to fitness (Rosenberg 1997). The distribution of fitness effects of new mutations is an empirical matter: that is, it is determined by observation rather than by direct theoretical prediction (Eyre-Walker and Keightley 2007). If we discovered that some lucky organism was able to skew this distribution by generating the right kinds of mutations just when they needed them, and thus to potentially speed the rate of adaptive evolution, this would obviously greatly impact on our understanding of evolutionary biology, but it would not necessarily require us to rewire our theoretical machinery (Foster 1999).

At this point in time, available evidence suggests that DNA sequence mutation is random with respect to fitness. As Theodosius Dobzhansky noted "An ideal situation would be if the organism were to respond to the challenge of the changing environment by producing only beneficial mutations where and when needed. But nature has not been kind enough to endow its creations with such a providential ability" (Dobzhansky 1937). But the models we use in evolutionary biology to understand the dynamics of population genetic change would still be useful if mutation was not random. So the potential for directed change in linguistics is an example of a fundamental and important difference between language evolution and biological evolution that does not necessarily impact on the utility of borrowing theoretical tools.

But there might be other differences between linguistic and genetic evolution that do impact on the applicability of evolutionary models. For example, in sexually reproducing populations of diploids, such as humans, where an individual inherits one copy of each chromosome from each parent, allele frequencies are prone to sampling error. Since each individual can only carry two variants of any gene, a small sample of individuals is unlikely to contain a fair representation of all the variation in the population. Because of this, rapid population genetic change is expected to occur in small founder populations due to the initial loss of unsampled alleles (the founder effect). This initial loss of genetic variation might then be compounded by ongoing loss of variation due to genetic drift, particularly while the population remains relatively small.

The founder effect has also been invoked to explain global patterns of language evolution. Decreasing phoneme inventories along chains of human migration has been attributed to serial founder effects (Atkinson 2011; Hajek 2004), and languages with smaller speaker populations have been reported to have smaller phoneme inventories (Hay and Bauer 2007). But it is not clear the extent to which serial founder events in human societies ought to have the same effect on languages as it does on genetic variation (Bowern 2011). If small numbers of individuals can know and use all the common language elements of the parent population, then the subsampling that inevitably occurs in genetic variation when a new population is started from a small number of founders might not occur for language features such as phonemes. A small founder population might be prone to subsequent loss of word variants through failure to use all variants consistently over generations (ongoing drift), though this might only occur if the small size is maintained after foundation, rather than in situations where the speaker population grows rapidly again to reach the same size as the parent population (Bromham et al. 2015). The fit of language data to a founder effect model is an ongoing area of investigation, tested with both simulations and empirical data (Fort and Pérez-Losada 2016; Hunley et al. 2012).

We should care about whether our models do the job we want them to do. Consider the application of population genetic models to language change. The Wright–Fisher model of population genetics provides a way of calculating the probability of gene frequencies in a population after a number of generations (Fisher 1930; Wright 1931). This model is typically described in terms of a population of organisms consisting of a constant number of diploid individuals (that is, organisms that have two, potentially non-identical, copies of every chromosome) with random mating (no biased mate selection), non-overlapping generations (the parent generation has died by the time their offspring reach maturity) and no selection (no individual more likely to reproduce than any other, therefore no gene more likely to be included in the next generation). The sampling process described is (usually) sexual reproduction: the individuals of one generation mate at random and have offspring that inherit one allele (variant) of each gene from each parent, selected at random. Given this process, and information on initial allele frequencies, mutation rate and population size, the model allows us to generate predictions of distributions of allele frequencies in future generations.

Despite the apparent lack of resemblance between language change and population genetics, variants of the Wright–Fisher model have been applied to language evolution to model the way that the language variants may be sampled unevenly from a pool of possible variants, leading to gradual change over time (Baxter et al. 2006, 2009; Bentley et al. 2011; Blythe 2012; Kirby et al. 2014; Reali and Griffiths 2010). But what is the equivalent of sexual reproduction for languages? Is it an individual acquiring language from their parents? Is it learning from other members of the same cohort? While an individual receives their genes "vertically" from their parents, language variants can be acquired both vertically (from parents) and horizontally (from unrelated individuals such as peers). Everyone, with few exceptions, carries the same set of genes, so each individual has exactly two copies of each gene (except for those on sex chromosomes) which they inherited from their parents. So we can define a population genetic model where each individual can carry two variants of each gene. But the population structure of a language is harder to define: individuals may differ in which parts of the language they know or regularly use, for example having a smaller or larger vocabulary, or knowing more or less variant forms of a particular lexical item. It is not even clear how to translate the parameters of the Wright–Fisher model to linguistics. In the usual formulation of the model, N is the number of individuals in biological population, so the number of alleles in the gene pool for a population of diploids is 2N. But in the case of language, is N the number of speakers (Bentley et al. 2004)? Or does N represent the number of occasions on which a given individual samples a given linguistic element (for example, hears a particular word associated with a particular meaning: Blythe 2012)? Or is N the number of different linguistic forms an individual is exposed to (Reali and Griffiths 2010)?

Empirical resemblance is unneeded as long as the model is doing the work we require of it. To use an entirely unrelated example, the hypercycle model for the origin of life describes a possible path way to the evolution of simple replicating evolvable entities. The hypercycle describes a hypothetical set of molecules that engage in both autocatalysis (stimulate the production of more molecules like themselves) and cross-catalysis (stimulate the production of another molecule: Maynard Smith 1979). If A makes more A and stimulates replication of B, and if B makes more B and stimulates production of C which makes copies of itself and stimulates production of A, then you have a hypercycle which might just have the potential for limited self-replication and evolution (Maynard Smith and Szathmáry 1995). The equations used to describe a hypercycle are mathematically equivalent to the Lotka-Volterra model used to describe predator–prey dynamics in ecology (Szathmáry 2013): if prey increase in number, predators increase in number, reducing the number of prey, reducing the number of predators, allowing prey to increase, and so on. There is no empirical resemblance between small autocatalytic molecules floating in solution and predators chasing prey over a landscape. But we can use the same form of model to describe and investigate both systems.

In the case of the application of Wright–Fisher models to linguistics, there may be empirical differences between language change and genetic change that we can wilfully ignore, as long as we are clear on exactly what we are using the model to do. We might consider that it doesn't make a practical difference whether the change in frequency of variants is by inheritance, learning or personal development. We might just be able to ignore whether N represents the number of bean plants, the number of different words for "bean", the number of times someone hears the word "bean", or actual beans in a bag. If we are using the model to generate a distribution of possible outcomes from an unbiased serial sampling process, the equations are the same. We can use the Wright–Fisher model to ask whether the representation of our different types of beans (plants, words, tokens, or objects) at one point has been randomly sampled from the available types in the previous time points.

To illustrate this approach, consider how the Wright–Fisher model has been applied to investigating the process of formation of a new contact language from two source languages (Meakins et al. submitted). Gurindji Kriol is a novel language that has emerged in the past 50 years in a community in Northern Australia, generated from the hybridization of Gurindji (from the Pama–Nyungan language family) and Kriol (a creole language derived primarily from English but with a distinct syntactic structure and grammar: McConvell and Meakins 2005; Meakins 2011, 2012). The formation of the new language has been so rapid that changes can be attested by recording language used by different age groups living in the same community. The grandparent generation speak both Gurindji and Kriol. Their children are no longer fluent in the two parent languages but mix elements from both source languages in their speech, although individuals may vary in which Kriol or Gurindji language elements they use. The grandchildren speak a single shared mixed language, using common language features that may have been derived from either Gurindji or Kriol. By comparing the usage of Gurindji and Kriol terms in each age group, and also the use of innovative new language features, we can investigate how the elements of the two parent languages have been sampled to make the mixed language. Have all language features from the two parent languages had the same chance of ending up in the mixed language spoken by the grandchildren? Or has the formation of the new language favoured adoption of elements from one or other parent language? Or has the mixed language preferentially included particular kinds of features, for example preferring simple over complex grammatical forms?

The Wright–Fisher model provides a theoretical tool for deriving the expected distribution of language variants in the grandchild generation if sampling of elements of the two source languages was entirely unbiased. By comparing this null expectation to the observed usage of Gurindji and Kriol forms, we could identify significant biases: Kriol forms were more likely to be included in the mixed language than Gurindji forms (Meakins et al. submitted). The analysis also provides a test of the hypothesis that the formation of contact languages is primarily driven by simplification. If this were true, then we would expect that the rate of adoption of simpler language variants into the contact language would be significantly higher than more complex forms. But complex Kriol features had the same rate of adoption as simple Kriol features, complex Kriol terms had a higher rate of adoption that simple Gurindji forms, and complex Gurindji forms had a greater rate of adoption that either simple or highly complex Gurindji forms. The lack of a signal of simplification in the formation of this particular contact language suggests that simplification is not a universal signature of contact language formation.

The Wright–Fisher model provides a tool for working out what we should expect to see if there was no preference for adopting one form of language feature over another, then we can compare the observed usage to the expectation in order to detect classes of language features that show patterns of adoption that can't be explained by random sampling. Comparing expectation to observation can be done blind to the actual mechanism whereby variants are sampled from one age group to the next. In this case, the population genetic model has provided us with a useful theoretical tool that allows us to ask whether the adoption of language elements from one generation to the next looks like random sampling or is biased toward certain sources or kinds of language features.

Analytical tools

Borrowing analytical tools is a practical way to solve similar problems encountered in both fields using tricks developed in one field or the other. Analytical tools will often be statistical techniques or computational methods that were developed in one field but can be adapted to be used on a different kind of data. For example, evolutionary biologists borrowed the analytical tools of game theory, typically applied in analysing economic behaviour, to investigate the evolution of behaviour in animal populations (Maynard Smith and Price 1973). The evolutionary game theory models have now been borrowed to explore cultural dynamics and language evolution (Jäger 2008; Nowak et al. 1999; O'Connor 2017). Game theory is a handy analytical tool that can be used to investigate a wide variety of systems where multiple strategies with different payoffs exist.

A clear case of borrowing analytical tools is the use of computer programs developed to produce phylogenies and networks. The enthusiastic adoption of the iconography of the evolutionary tree in biology, and the persistent but less universally accepted use of phylogenies in linguistics, has led in both fields to the widespread application of the computational phylogenetic methods, with programs originally developed for DNA sequence data being applied to comparative linguistic datasets (Bouckaert et al. 2012; Gray et al. 2009; Grollemund et al. 2015). Many of these methods assume that the data can be described by a bifurcating hierarchy, but analytical tools for representing non-bifurcating relationships, such as borrowing and hybridization, have also been horizontally transmitted between evolutionary biology and linguistics (Bryant et al. 2005; Gray et al. 2010; Hamed 2005; Willems et al. 2016). Stemmatic analyses have used phylogenetic methods from evolutionary biology to reconstruct ancient texts from surviving copies (Altschuler et al. 2013; Barbrook et al. 1998). Indeed, similar phylogenetic methods have been applied to comparisons between wording of legal documents in order to track the changes in anti-evolution legislation in the United States over time (Matzke 2016).

Given the amount of technological transfer between evolutionary biology and historical linguistics, it is perhaps surprising to consider cases of analytical tools that have failed to be swapped between the fields, despite a clear and acknowledged need. The case I would like to consider here is that of analytical tools for addressing for statistical non-independence due to patterns of descent.

The problem of non-independence of observations made on species or human societies has been recognized since the beginning of the development of evolutionary studies. As noted by Lyell (1863), Wallace (1855), Darwin (1859) and others, both languages and species show distinctly non-random patterns in space and time. The practical upshot of descent with modification that, on the whole, close relatives will tend to share many features (such as homologous traits or cognate words) and will also have a tendency to be located near each other and therefore share features of their environment. This complicates our ability to detect meaningful causal patterns by comparing observations from different languages or species. We can perhaps more easily appreciate the impact of the statistical non-independence due to descent if we consider three different ways that it impacts on cross-species and multi-language analyses (though there is not really a clear break between these three in practise, and in many cases they can be addressed using the same sets of tools). These are the problems of covariation, spatial autocorrelation, and phylogenetic non-independence.

The problem of covariation is so easy to demonstrate, it's actually fun. For example, much was made of the finding that there is a statistically significant correlation between the number of Nobel prize winners per country and the country's per capita chocolate consumption (Messerli 2012), with the widely reported implication that eating chocolate makes you smart.[2] Based on correlational evidence alone, we might just as easily conclude that flat pack furniture makes you smart because the number of Nobel prize winners is also related to the number of IKEA stores per country (Maurage et al. 2013). And chocolate consumption is also significantly associated with the number of serial killers per country (Roberts and Winters 2013).

Such relationships will naturally cause people to roll their eyes and wearily sigh "correlation doesn't prove causality". This is, of course, true—but only in the sense that correlation between any two variables does not imply a direct causal connection between the two. But it is fair to say that correlation does imply some form of causal relationship between variables (Shipley 2002). This is precisely what the formal statistical methods are aiming to establish: whether the covariation between two variables is beyond what you would expect from a chance association between the two. When chance association is rejected as improbable we conclude there is a high probability of some kind of causal link between the variables, even if we don't know what that link is. We might hazard a guess what the link actually is, or we might have no idea, but our ignorance has no impact on the conclusion that there is a strong probability that *something* is causing the variables to covary.

---

[2] Indeed evidence is growing for an effect of cocoa on cognitive functions (Sokolov et al. 2013), leading to calls for "cocoa interventions" for urban children to counter the mind-numbing effects of pollution (Calderon-Garciduenas et al. 2013). But these conclusions are based on empirical studies (such as the cutely named "Cocoa, Cognition, and Aging (CoCoA) Study": Mastroiacovo et al. 2015), rather than cross-cultural correlations.

Covariation is a problem when we want to explain anything that has a geographic pattern, because there is a fair bet that some other variable can be found that shares the same pattern. For example, there is a growing body of studies showing a link between parasite load and many different aspects of human behaviour and culture, including authoritarianism, religiosity, rates of scientific innovation, marriage patterns and spicy cuisine (Fincher and Thornhill 2012; Murray 2014; Murray et al. 2013; Schaller and Murray 2010). These studies have not been without their critics (e.g. Currie and Mace 2012; Pollet et al. 2014; Ross and Winterhalder 2016), but the flow of identified correlates with parasite load remains undiminished, at least partly because there is a very large pool of potential covariables. Human parasite load varies significantly with latitude (Cashdan and Steele 2013; Guernier et al. 2004), so anything else that varies with latitude will also vary with parasite prevalence. And there are a lot of things that vary with latitude.

For example, a recent paper correlated an index of traditionalism (derived from answers to survey questions from 30 countries) to parasite load, concluding that maintenance of traditional behaviours is driven by a need to reduce infection risk (Tybur et al. 2016). But a reanalysis shows that traditionalism also correlates significantly with latitude. In fact, latitude combined with country population size is better predictor of variation in traditionalism than parasite load (Schneemann et al. 2017). Because traditionalism correlates with latitude, it will correlate with anything else that correlates with latitude, including a wide range of climatic variables and many measures of biodiversity. This is interesting, but not very helpful for untangling causal links. For example, traditionalism also correlates significantly with bird species richness, yet we might be reluctant to conclude that living with a rich avifauna makes people more conservative (Schneemann et al. 2017).

Language diversity also scales with latitude (Collard and Foley 2002; Mace and Pagel 1995), so in explaining language diversity we need to be careful not to accidentally assign a causal relationship to variables that are linked through their common latitudinal gradients. For example, we should not be surprised to learn that language diversity can also be correlated to parasite load (Fincher and Thornhill 2008). More broadly, linguistic diversity and biodiversity show similar spatial patterns (Gorenflo et al. 2012): is this because they are causally linked with each or because they both scale with some other environmental factor (Gavin et al. 2013)? This brings us to the next form of statistical non-independence: spatial autocorrelation. As Lyell (1863) noted, related languages tend, by and large, to cluster together in space. This means that spatial samples do not represent independent observations about the relationship between language features and their environment.

To illustrate this point, consider one of the world's most intriguing hotspots of language diversity, the New Guinea region, where nearly a fifth of the world's recognized languages are jammed into less than 1% of the Earth's surface (Foley 2000). Spatial distribution of languages on the main island of New Guinea is significantly correlated with patterns of mammal species richness, leading to claims of a functional relationship between biodiversity and human cultural diversity (Turvey and Pettorelli 2014). However, despite calls for an integrated biocultural approach to conservation (Gavin et al. 2015), a negative association between endangerment of languages and mammals across New Guinea led to suggestions that languages and species face different threatening processes in this region (Turvey and Pettorelli 2014). But the interpretation of correlations between languages and biodiversity, and between threatened languages and threatened species, is complicated by spatial autocorrelation.

Maps of language and species distribution in New Guinea shows that some areas have relatively low diversity of both languages and mammal species, particularly in the southern lowlands (Fig. 1). From the spatial data alone, we don't know why this area has fewer languages and lower mammal diversity: whether it's just a coincidence that the low language diversity area happens to coincide with low mammal species diversity, or whether the same process has caused both mammals and languages to have low diversity in this region. But, whatever the cause, this clear spatial pattern creates a problem for our attempt to apply correlation analyses. If we divide New Guinea into grid cells and ask whether cells that have high biodiversity also tend have high language diversity, then every grid we sample from this lowland area will speak with one voice: low language richness, low mammal

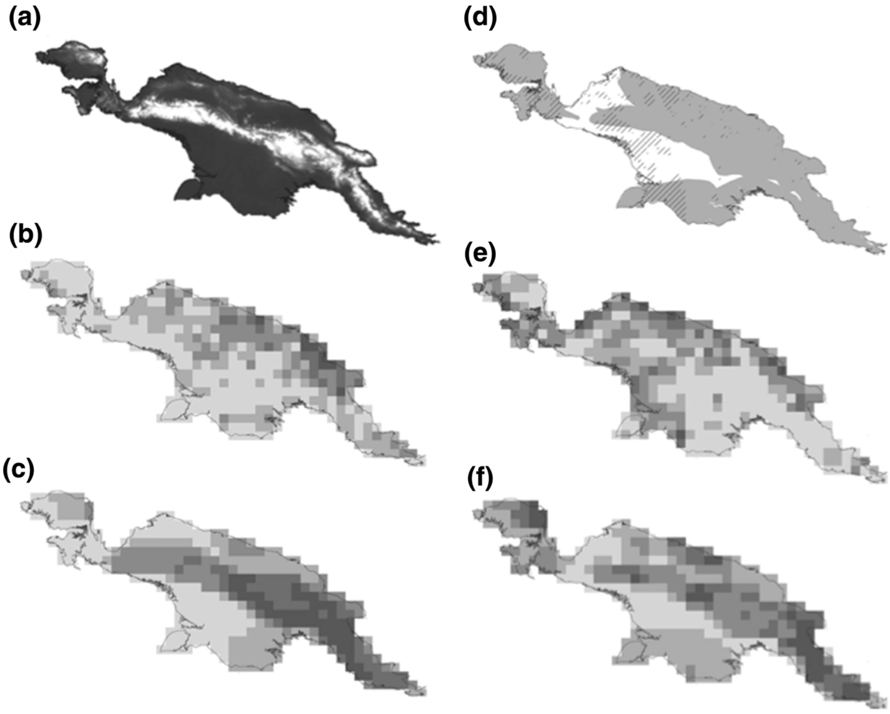

Fig. 1 Distribution of languages and mammal species in New Guinea. Patterns of language diversity (**b**) and mammal diversity (**c**) are non-random with respect to elevation (**a**). Patterns of threatened languages (**d**: hatched areas) and threatened mammals (**d**: shaded areas) also show distinct spatial clustering which is evident from the pattern of number of threatened languages (**e**) and threatened mammal species (**f**) per grid cell. **a** Elevation, **b** language richness, **c** mammal species richness, **d** threatened languages and mammals, **e** threatened languages and **f** threatened mammal species. Modified from Turvey and Pettorelli (2014), and reproduced under Creative Commons license CC-BY version 4.0

richness. The more samples we draw from this region, the louder the voice, and the stronger the apparently relationship between language and mammal diversity becomes. But multiple grid cells from this same region don't represent independent datapoints sampling the relationship between two variables. Instead they represent pseudoreplication, making the same empirical observation—this particular lowland area happens to have low language and mammal diversity—over and over again, yet counting it as different datapoints. We could make this same observation more times if we divide the area into increasingly small grid cells, but although that will increase the number of datapoints it won't change the amount of real evidence we have concerning the relationship between mammal species richness and language species richness. Including multiple grid cells representing the same spatial pattern is really just duplicating the same data point, so the grid cells do not satisfy the assumption of independence required by all statistical methods, where each datapoint has to be a fair and unbiased test of the relationship between one variable and another.

The problem is easier to appreciate when we consider the curious negative relationship between threatened languages and threatened mammal species (Turvey and Pettorelli 2014). Looking at the maps, we can see that most threatened languages are found in the western half of New Guinea, in the Indonesian state of Irian Jaya (Fig. 1d). There is a highest density of threatened mammal species in the highlands on the eastern side of the island, in the independent state of Papua New Guinea (PNG: Fig. 1f). It's possible that each of these observations has a single underlying cause, each effectively unconnected to the other: for example, it may be that indigenous languages have not fared well under Indonesian occupation, and mammals have not fared well under the dense settlement of the highlands. When the island is divided into grid cells and each cell contributes one datapoint to the analysis, these same two patterns are sampled again and again: cells in Irian Jaya tend to have higher language threat (but mostly have lower mammal threat than the PNG highlands), and cells in the PNG highlands have high mammal threat (but tend to have lower language threat than Irian Jaya). This generates a significant negative correlation, but the correlation arises from the similarity between neighbouring grid cells, not because mammal threat and language threat have, in this case, any clear causal relationship.

Because biologists frequently encounter this problem when analysing species diversity data, they use statistical techniques that correct for the similarity of grid cells based on their spatial proximity (Dormann et al. 2007). These methods typically use a matrix of distances between the observations to partition the variance between datapoints into that attributable to the spatial proximity of different samples, and the residual variation once proximity has been taken into account. Once you have taken the spatial relationship between grid cells for the New Guinea data into account, there is no additional correlation between language richness and species richness, and no correlation between threatened species and threatened languages (Cardillo et al. 2015). The apparent correlation is due to the spatial clustering of diversity, and so might not be a result of a functional link between biodiversity and language diversity.

We don't have to be using grid cells as datapoints to encounter problems of spatial autocorrelation. Any analysis that uses countries or cultures or languages as data points runs into the same problems. For example, the observed relationship between mean growing season and language diversity has been interpreted as evidence for a demographic model of language diversity, where year-round production of food allows smaller stable social groups to develop, leading to more languages in a given area (Nettle 1998). But a close look at the data reveals clustering of related cultures, for example the Melanesian countries all occur in the top right hand corner—long growing season, high spatial language diversity—and Middle Eastern countries in the lower left hand corner—shorter growing season, lower language diversity (Fig. 2). So even though there is a clear linear association, we can't tell if the relationship is caused by a causal link between growing season and language diversity, or if there is some other "lurking variable" (Naroll 1965) that tends to be similar between related cultures, and also scales with environment.

Similarly, plots of speaker population size against aspects of language structure also show spatially non-random patterns. Plots of the complexity of verb morphology against speaker population size have been used to argue that languages become simplified in large populations, potentially to facilitate learning by second-language speakers (Lupyan and Dale 2010). Looking at the data (Fig. 3), verb morphology tends to be high in the Americas and Australia (where average speaker population sizes are low) and low in Europe (where average speaker population sizes are large). While this fascinating pattern may demand explanation, we would need to be cautious about interpreting comparisons between complexity and population size, since the geographic patterns could be causing a correlation between the two whether they are functionally linked or not. If European languages happen to have low verb morphology, then it will correlate with all other traits that characterise European languages, such as higher average speaker population sizes.

The recognition of similar spatial patterns in biodiversity and language diversity has led to ecologists working on patterns of language diversity, providing a conduit

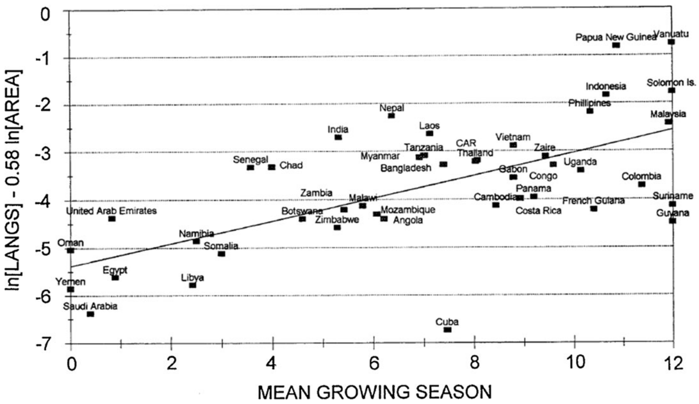

Fig. 2 Number of languages per country, accounting for differences in country size, against the mean growing season in months. Reproduced from Nettle (1998) with permission from Elsevier (License number 4046981123147)

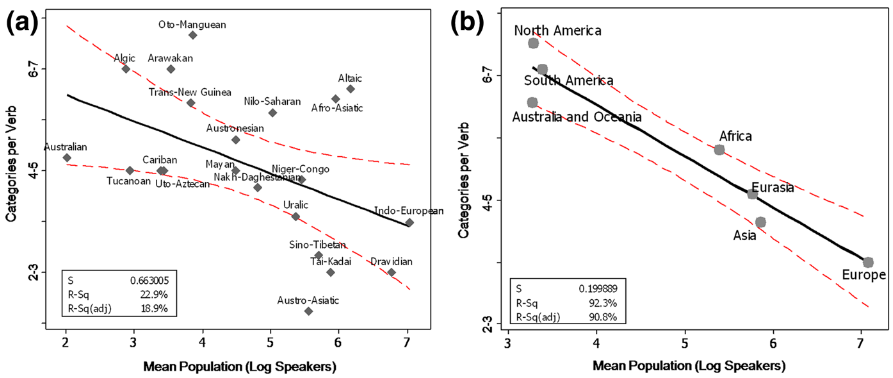

Fig. 3 Relationship between inflectional synthesis of verbs, based on data from over 2000 languages taken from the World Atlas of Language Structures (WALS), and speaker population size, averaged by continent. Reproduced from Lupyan and Dale (2010) under Creative Commons Attribution (CCBY) license

for analytical techniques developed in biogeography and macroecology to be applied to the study of languages (Amano et al. 2014; Gavin and Sibanda 2012; Gorenflo et al. 2012; Turvey and Pettorelli 2014). But there is a related area where there has been surprisingly little tool swapping, despite the clear recognition of shared problems: the problem of statistical non-independence due to descent, which in evolutionary biology is referred to as "phylogenetic non-independence" but in cultural studies and linguistics is known as "Galton's problem".

Galton's problem is the recognition that related cultures tend to share many attributes, so comparing cultures can lead to false inference of causal relationships that are actually due to co-inheritance of traits from a common ancestor. When Edward Burnett Tylor, one of the founders of cross-cultural analysis, presented a paper on marriage practices to the Anthropological Institute of Great Britain and Ireland in 1889, the President, Francis Galton, praised the comparative data on which the study was based, but queried the method of analysis, which did not take into account the interdependence of the observations. Galton noted that "full information should be given as to the degree in which the customs of the tribes and races which are compared together are independent. It might be, that some of the tribes had derived them from a common source, so that they were duplicate copies of the same original. Certainly, in such an investigation as this, each of the observations ought, in the language of statisticians, to be carefully 'weighted'" (see Tylor 1889).

As for spatial autocorrelation, observations do not count as independent tests of a hypothesis if they are effectively measuring the same observation again and again, in languages that have all inherited the same set of characters from their common ancestor. While the co-occurrence of particular linguistic traits has been interpreted as evidence that the presence of one trait demands the presence of the other, the tight coupling of traits may be due to their shared presence in a single ancestral language, and their retention in many of its descendants (Bickel 2009; Evans 2013). The traits may be associated by phylogenetic inertia rather than because one trait necessitates the other. For example, analysis of hundreds of languages led to the hypothesis that word order universals constrained languages to predictable sets of grammatical structures (for example that object–verb ordering predicts postpositions, relative clauses and genitives before the noun), but reanalysis in a phylogenetic context showed that most of these patterns are lineage-specific (Dunn et al. 2011). Counting each language as an independent observation of the association between word order and grammatical structure inflates the appearance of consistent relationships.

Galton's problem is widely recognized in the study of language evolution (Bickel 2009; Everett et al. 2015; Levinson and Gray 2012; Mesoudi et al. 2006; Moran et al. 2012; Roberts and Winters 2013). Yet few linguistic studies that make comparisons between languages employ effective countermeasures. Adding language family as a factor in the analysis does not remove the problem, because phylogenetic non-independence occurs at all level of the genealogical hierarchy: closely related languages will tend to be more similar to each other than to their more distant relatives within the same family, and more closely related language families will tend to share more in common with each other than more distant families (Dediu and Levinson 2012; Eff 2004; Pagel et al. 2013). For the same reason, balanced sampling (choosing one representative of each group) will also not make the problem of non-independence go away: there will still tend to be more similarities between representatives from closely related groups than to more distant groups. The paucity of practical solutions to Galton's problem in linguistics is odd given the amount of energy expended on solving the same problem in evolutionary biology.

The problem of non-independence due to descent has been appreciated in evolutionary biology for as long as it has in cultural studies. Darwin (1859) drew attention to the problem of spurious correlations caused by descent: "We may often falsely attribute to correlation of growth, structures which are common to whole groups of species, and which in truth are simply due to inheritance; for an ancient progenitor may have acquired through natural selection some one modification in structure, and, after thousands of generations, some other and independent modification; and these two modifications, having been transmitted to a whole group of descendants with diverse habits, would naturally be thought to be correlated in some necessary manner".

Long and bloody battles have been fought in evolutionary biology over the need for, and means of, correcting for phylogenetic non-independence in analyses that look for patterns in the co-occurrence of traits in different species (e.g. Harvey et al. 1995; Pagel and Harvey 1992; Westoby et al. 1995). While it was initially claimed that some fields could ignore the problems—such as ecology—the recognition of the fundamental nature of the problem of covariation by descent has gradually seeped through biology. All statistical analyses assume independence of datapoints, therefore all statistical tests are violated when the features of the samples (such as species or languages) are more likely to be similar between close relatives. Many methods have been proposed to deal with this, ranging from choosing datapoints that satisfy the assumptions of independence (e.g. the method of independent comparisons) to modelling the covariation due to descent (e.g. phylogenetic least squares regression), though all these methods may be essentially the same under the skin even if bearing little surface resemblance (Blomberg et al. 2012; Felsenstein 2004; Pagel and Harvey 1992).

With some notable exceptions (e.g. Dunn et al. 2011; Haynie and Bowern 2016), why have analytical tools developed in evolutionary biology to deal with the pervasive problem of phylogenetic non-independence not been adopted in studies of linguistic diversity to solve the widely-recognized Galton's problem? Perhaps the reasons are social, such as a lack of effective communication between the fields—but then recent developments in phylogenetic methodology have flowed freely from evolutionary biology to linguistics so it seems odd that other evolutionary methods should not likewise travel between the two disciplines. It may be cultural, perhaps due to emphasis on different kinds of statistical analyses, but then there are plenty of big-data approaches to understanding linguistic diversity so it seems unlikely to be an allergic reaction to fancy statistical analyses. There may even be labelling problems: evolutionary biologists generally don't recognize the phrase "Galton's problem", just as linguists rarely refer to "phylogenetic non-independence"—while this may seem trivial it has great practical importance because searching under one name of the problem may not reveal solutions developed under the other. And, unhelpfully, biologists usually refer to the body of techniques developed to counter these problems as "the comparative method" (generating datapoints that represent statistically independent observations: Harvey and Pagel 1991), a phrase which unfortunately has a different meaning to linguists (identification of features shared by descent in order to reconstruct historical relationships: Hoenigswald 1963).

But a more practical reason for the low uptake of phylogenetic comparative methods in linguistics is the understandable yet misguided belief that you can't do anything about phylogenetic non-independence unless you haves a fully resolved phylogeny. The same problem slowed the introduction of phylogenetic comparative method in evolutionary biology, which got bogged down in arguments about which phylogeny we should use and how do we know what the true tree is anyway and what we should do when we don't have a phylogeny (Miles and Dunham 1993). There is also the problem that, while most species phylogenies are constructed from DNA sequences that are typically not directly related to the traits being tested for evolutionary patterns, language phylogenies may be based on the same corpus of data which is being examined for patterns and correlations, introducing a potential circularity.

Not having a handy solution to the problem does not make the problem go away: as Mark Pagel and Paul Harvey, two of the founders of the phylogenetic comparative method in biology, said: "wishing does not make it so" (Pagel and Harvey 1992). Ignoring Galton's problem leaves researchers vulnerable to over-interpreting inherited patterns as meaningful causal relationships. So what can linguists do? Some might place "build more phylogenies" at the top of the wish list, but there are solutions that do not require a fully-resolved, bifurcating language tree. Any information on relatedness is better than none at all. For example, a language taxonomy might not be a perfect reflection of phylogenetic relationships, but it will do a better job of ameliorating covariation due to descent than using no phylogeny at all (Schneemann et al. 2017). If there are multiple equally plausible hypotheses about relationships between languages, then the analysis can be conducted on all of them to ask how much difference the uncertainty in genealogical relationships make to the conclusions we can drawn. When seeking evidence of true causal relationships, selecting fewer safer datapoints is likely to be a better strategy than including lots of datapoints of dubious value (as biologists navigating the bold new landscape of big-data phylogenomics are slowly coming to realise).

To illustrate that it is possible to apply the principles of the phylogenetic comparative method without actually using a phylogeny in the analysis, consider the influence of population size on language evolution. If related languages are more likely to be similar in population size than they are to unrelated languages, and if the feature of interest is also more similar among relatives, then there is a risk of being distracted by incidental correlation. If Australian languages have a tendency to complex verb morphology, and given that Australian speaker populations are comparatively small, then including these languages in a global analysis of complexity and population size will tend to promote a negative relationship between population size and verb morphology, whether or not the two are causally connected (Lupyan and Dale 2010: Fig. 3). In this sense the phylogenetic nestedness of observations is, like spatial autocorrelation, a “nuisance factor”, something that is preventing us from seeing the relationship we are really interested in. A phylogenetic comparative test is a way of looking for relationships between variables above and beyond the relationship that comes for free from patterns of descent (Pennell and Harmon 2013).

There are several ways to address this problem. One is to take a phylogenetic approach, modelling the evolution of both traits along a phylogeny and asking if change in one tends to be accompanied by a change in the other (Levinson and Gray 2012). But if you don’t have a phylogeny, you might still be able to compare traits in related languages and ask whether the traits of interest are correlated with each other. Consider a pair of related languages: as long as you know that the languages you are considering sprang from a common ancestor, then you also know that at some point in the past they both started with the same population size and the same trait values. Any difference in trait and population size has evolved since then, so you can ask if the one with the large population also has the simpler morphology, or the faster rate of change, or whatever interests you. Of course, one such comparison does not tell you much, as the association could be due to chance or coincidence. But if you can gather together many different independent pairs of related languages, you can begin to generalise.

Sister pairs are pairs of close relatives that share a more recent common ancestor with each other than they do with any other language in your analysis. In other words, the pairs represent unique trajectories of evolutionary history. Since sister languages begin from the same starting point, any difference between them must have evolved since they last shared a common ancestor. And that difference has evolved independently of any differences between any of the other non-overlapping pairs, so each pair represents an independent observation of the association between population size and the trait of interest. This means you can do statistical tests on the observed patterns—which might be as simple as asking whether you get a positive association between trait and population size more often than you’d expect by chance, or it might be a more sophisticated analysis using models of evolutionary change to compare observation to expectation (Bromham et al. 2015).

Comparing phylogenetically-independent pairs is like running a historical experiment again and again (Bromham 2016)—what happens if I take a language, split it into two speaker populations, and let one evolve larger population size, will it also evolve faster rates of change or less complex morphology (or any other trait of interest)? Selection of pairs is easiest when you have a fully resolved phylogeny that you trust is a fair representation of the history of the languages. But, even in the absence of a phylogeny, it's a fair bet that for any group of languages, a knowledgeable expert will be able to identify some comparisons that represent each others' close relatives. As long as you can identify pairs of languages that you are confident share a more recent common ancestor with each other than they do with any of the other pairs, you can conduct a phylogenetic comparative analysis (Bromham et al. 2015). Selection of well-attested sister pairs has been unpopular due to the feelings of loss associated with "throwing the data away" (not including all available observations in an analysis). But lots of data does not help you if it doesn't satisfy the assumptions of the test: it's better to have fewer reliable datapoints. The analytical tools of the phylogenetic comparative method help us to chip away the confounding factors to reveal the underlying causal relationships.

Conclusion

Interdisciplinary exchange can have significant impacts on research agendas. Often, this exchange is discussed in terms of workers with different areas of expertise coming together to solve a multidisciplinary problem that cannot be profitably tackled from a narrow perspective, thus requires the integration of previously separate strands of research (Repko 2007). But the tool-swapping discussed in this paper is a different kind of interdisciplinary exchange, where ideas, models and methods from one field are borrowed to aid progress in another field (see Luszki 1958). Sometimes this is achieved through the co-operation of multidisciplinary teams, but sometimes it is by workers in one field adapting tools borrowed from a different field to suit their own particular tasks (Bromham et al. 2016).

The fields of historical linguistics and evolutionary biology have been intertwined at the edges, creating conduits for the swapping of tools between disciplines. Some of these tools can help solve shared problems, such as methods for dealing with statistical non-independence. Some tools provide new ways of looking at existing data, and might stimulate the collection of new kinds of data, such as the application of population genetic models to language change. Other tools provide a kind of intellectual lubricant, allowing ideas and evidence from one field to aid the uptake of similar ideas in the other, as observations of language change over time did in the establishment of Darwinian gradualism.

It could be argued that ideas could flow more freely between fields of study in the nineteenth century before discipline boundaries codified and researchers became more specialised. For example, the great naturalist Alfred Russel Wallace collected word lists as well as species on his travels, compiling tables of one hundred and seventeen basic vocabulary items for dozens of languages in the Malay archipelago, “more than half of which I believe are quite unknown to philologists” (Wallace 1872). However, fuzzier discipline boundaries do not provide a general explanation of the vibrant exchange of ideas in the nineteenth century between historical linguistics and evolutionary biology. Darwin and Lyell used language as an analogy to make a rhetorical point, not because they were engaged in philological studies. While Schleicher maintained a fruitful relationship with the great developmental biologist Ernst Haeckel, Schleicher’s works were entirely in the sphere of linguistics, not biology. The exchange of ideas between linguistics and biology was, then as now, predominantly through experts in one field taking an academic interest in the other. This tradition continues today, despite the hardening of discipline specializations.

Much of the discussion around the interdisciplinary transfer between linguistics and biology has focussed on interrogating whether the underlying evolutionary processes are the same or different (Croft 2008; List et al. 2016; Van Wyhe 2005; Whitfield 2008). Lack of shared mechanism has been a point of criticism of analogies drawn between language evolution and biological evolution (List et al. 2016), and, conversely, similar patterns have been used to argue for similar processes (Whitfield 2008). There are many ways in which language evolution and biological evolution are fundamentally different. Some of these will make us think twice about borrowing tools designed in one field for use in the other—for example, it is not yet clear if founder effect models built on diploid genetic inheritance are a useful way to describe loss of language variants. Other differences may seem overwhelming yet might not end up making much difference to the models, for example it may be that deliberate language change will not be an impediment to using evolutionary models that were designed for a system that has random generation of heritable variation. I have argued here that homology of process is not always critical for profitable tool swapping. Analogies between the two fields have in the past been useful explanatory devices, whether or not those analogies reflect the same underlying mechanisms. Identifying useful analogies can also lead to the adoption of potentially useful methods from one field to the other, such as drawing analogies between horizontal gene transfer and hybridization in biology and borrowing and contact-induced change in linguistics.

Despite the many fundamental differences, there are a number of similarities between language change and biological evolution that influence the way both biologists and linguists go about the task of understanding the generation and maintenance of diversity. One is that the closest that either historical linguistics or evolutionary biology can come to a universal statement is that there are probably no universal patterns of change and every rule has an exception (except perhaps this one: Evans 2013; Welch 2016). And the other is that extinction is nearly always irreversible, and the loss of either a language or a species represents a tragedy to those who wish to understand evolutionary processes. Each species or language represents the accumulation of unique features and novel combinations of structures, and as such each attests to the relationships between traits, representing the outcome of an evolutionary process under novel circumstances with a unique history, shaped by both chance and directional change. Just as the loss of a language is the loss of a unique set of words and structures, so the loss of a species is the loss of a unique set of genetic variants that worked together as an integrated whole. Every species or language lost is a thread lost from the global tapestry made over a vast number of lifetimes. The more threads we lose, the less we can understand about the processes and patterns that shaped that diversity.

Acknowledgements With gratitude to Kim Sterelny and Simon Greenhill for inviting me to contribute to their workshop on "Evolution of Language" held at the Australian National University in August 2016, and to two anonymous referees for their detailed and helpful comments on the manuscript.

References

Alter SG (2007) Darwin and the linguists: the coevolution of mind and language, part 1. Problematic friends. Stud Hist Philos Sci C Stud Hist Philos Biol Biomed Sci 38(3):573–584

Altschuler EL, Calude AS, Meade A, Pagel M (2013) Linguistic evidence supports date for homeric epics. BioEssays 35(5):417–420

Amano T, Sandel B, Eager H, Bulteau E, Svenning J-C, Dalsgaard B, Rahbek C, Davies RG, Sutherland WJ (2014) Global distribution and drivers of language extinction risk. Proc R Soc B Biol Sci 281(1793):20141574

Atkinson QD (2011) Phonemic diversity supports a serial founder effect model of language expansion from Africa. Science 332(6027):346–349

Atkinson QD, Gray RD (2005) Curious parallels and curious connections—phylogenetic thinking in biology and historical linguistics. Syst Biol 54(4):513–526

Barbrook AC, Howe CJ, Blake N, Robinson P (1998) The phylogeny of the Canterbury Tales. Nature 394:839

Bartholomew M (1973) Lyell and evolution: an account of Lyell's response to the prospect of an evolutionary ancestry for man. Br J Hist Sci 6(3):261–303

Baxter GJ, Blythe RA, Croft W, McKane AJ (2006) Utterance selection model of language change. Phys Rev E 73:046118

Baxter GJ, Blythe RA, Croft W, McKane AJ (2009) Modeling language change: an evaluation of Trudgill's theory of the emergence of New Zealand english. Lang Var Change 21:257–296

Bentley RA, Hahn MW, Shennan SJ (2004) Random drift and culture change. Proc R Soc Lond B Biol Sci 271(1547):1443–1450

Bentley RA, Ormerod P, Batty M (2011) Evolving social influence in large populations. Behav Ecol Sociobiol 65:537–546

Berwick RC, Chomsky N (2015) Why only us: language and evolution. MIT Press, Cambridge

Bickel B (2009) A refined sampling procedure for genealogical control. STUF Lang Typol Univers Sprachtypol Univers 61(3):221–233

Blasi DE, Michaelis SM, Haspelmath M (2017) Grammars are robustly transmitted even during the emergence of creole languages. Nat Hum Behav. https://doi.org/10.1038/s41562-017-0192-4

Blomberg SP, Lefevre JG, Wells JA, Waterhouse M (2012) Independent contrasts and PGLS regression estimators are equivalent. Syst Biol 61(3):382–391

Blythe RA (2012) Neutral evolution: a null model for language dynamics. Adv Complex Syst 15(03n04):1150015

Bouckaert R, Lemey P, Dunn M, Greenhill SJ, Alekseyenko AV, Drummond AJ, Gray RD, Suchard MA, Atkinson QD (2012) Mapping the origins and expansion of the Indo-European language family. Science 337(6097):957–960

Bowern C (2011) Out of Africa? The logic of phoneme inventories and founder effects. Linguist Typol 15(2):207–216

Bromham L (2016) Testing hypotheses in macroevolution. Stud Hist Philos Sci A 55:47–59

Bromham L, Hua X, Fitzpatrick TG, Greenhill SJ (2015) Rate of language evolution is affected by population size. Proc Nat Acad Sci USA 112:2097–2102

Bromham L, Dinnage R, Hua X (2016) Interdisciplinary research has consistently lower funding success. Nature 534(7609):684–687

Bryant D, Filimon F, Gray R (2005) Untangling our past: languages, trees, splits and networks. In: Mace R, Holden S, Shennan S (eds) The evolution of cultural diversity: a phylogenetic approach. Left Coast Press, Walnut Creek

Cairns J, Overbaugh J, Miller S (1988) The origin of mutants. Nature 335:142–145

Calderon-Garciduenas L, Mora-Tiscareño A, Franco-Lira M, Cross J, Engle R, Gómez-Garza G, Jewells V, Weili L, Medina-Cortina H, Solorio E, C-k Chao, Zhu H, Mukherjee P, Ferreira-Azevedo L, Torres-Jardón R, D'Angiulli A (2013) Flavonol-rich dark cocoa significantly decreases plasma endothelin-1 and improves cognition in urban children. Front Pharmacol 4:104. https://doi.org/10.3389/fphar.2013.00104

Cardillo M, Bromham L, Greenhill SJ (2015) Links between language diversity and species richness can be confounded by spatial autocorrelation. Proc R Soc B Biol Sci 282:20142986

Cashdan E, Steele M (2013) Pathogen prevalence, group bias, and collectivism in the standard cross-cultural sample. Hum Nat 24(1):59–75

Collard IF, Foley RA (2002) Latitudinal patterns and environmental determinants of recent human cultural diversity: do humans follow biogeographical rules? Evol Ecol Res 4(3):371–383

Costa JT (2013) Engaging with Lyell: Alfred Russel Wallace's Sarawak law and ternate papers as reactions to Charles Lyell's principles of geology. Theory Biosci 132(4):225–237. https://doi.org/10.1007/s12064-013-0188-1

Croft W (2008) Evolutionary linguistics. Annu Rev Anthropol 37:219–234

Crow JF (2001) The beanbag lives on. Nature 409(6822):771

Crow JF (2008) Haldane and beanbag genetics. Int J Epidemiol 37(3):442–445

Currie TE, Mace R (2012) Analyses do not support the parasite-stress theory of human sociality. Behav Brain Sci 35(02):83–85

Darwin C (1837) Letter to Caroline Darwin. Darwin Correspondence Project. http://www.darwinproject.ac.uk/DCP-LETT-346

Darwin C (1838) Notebook n: [metaphysics and expression (1838–1839)]. Cul-dar126.- transcribed by Kees Rookmaaker. In: P Barrett (ed). (Darwin Online http://darwin-online.Org.Uk/ Barrett P (ed.))

Darwin C (1858) On the perpetuation of varieties and species by natural means of selection. J Proc Linn Soc Lond Zool 3:45–50

Darwin C (1859) The origin of species by means of natural selection: or the preservation of favoured races in the struggle for life. John Murray, London

Darwin C (1868) The variation of animals and plants under domestication. John Murray, London

Darwin C (1871) The descent of man, and selection in relation to sex. John Murray, London

Dediu D, Levinson SC (2012) Abstract profiles of structural stability point to universal tendencies, family-specific factors, and ancient connections between languages. PLoS ONE 7(9):e45198

DeGraff M (2005) Linguists' most dangerous myth: the fallacy of creole exceptionalism. Lang Soc 34:533–591

Dobzhansky TG (1937) Genetics and the origin of species. Columbia University Press, New York

Doolittle WF (1999) Phylogenetic classification and the universal tree. Science 284(5423):2124–2128

Dormann CF, McPherson JM, Araújo MB, Bivand R, Bolliger J, Carl G, Davies RG, Hirzel A, Jetz W, Kissling WD (2007) Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography 30:609–628

Dronamraju K (2011) Haldane, Mayr, and beanbag genetics. Oxford University Press, Oxford

Dunn M, Greenhill SJ, Levinson SC, Gray RD (2011) Evolved structure of language shows lineage-specific trends in word-order universals. Nature 473(7345):79–82

Durkin P (2014) Borrowed words: a history of loanwords in English. Oxford University Press, Oxford

Eff EA (2004) Does Mr. Galton still have a problem? autocorrelation in the standard cross-cultural sample. World Cult 15(2):153–170

Evans N (2011) Dying words: endangered languages and what they have to tell us. Wiley, New York

Evans N (2013) Language diversity as a resource for understanding cultural evolution. In: Richerson PJ, Christiansen MH (eds) Cultural evolution: society, technology, language, and religion. MIT Press, Cambridge

Everett C, Blasi DE, Roberts SG (2015) Climate, vocal folds, and tonal languages: connecting the physiological and geographic dots. Proc Nat Acad Sci USA 112(5):1322–1327

Eyre-Walker A, Keightley PD (2007) The distribution of fitness effects of new mutations. Nat Rev Genet 8(8):610–618

Felsenstein J (2004) Inferring phylogenies. Sinauer, Sunderland

Fincher CL, Thornhill R (2008) A parasite-driven wedge: infectious diseases may explain language and other biodiversity. Oikos 117(9):1289–1297

Fincher CL, Thornhill R (2012) Parasite-stress promotes in-group assortative sociality: the cases of strong family ties and heightened religiosity. Behav Brain Sci 35(02):61–79

Fisher RA (1930) The genetical theory of natural selection. Clarendon Press, Oxford

Foley WA (2000) The languages of new guinea. Annu Rev Anthropol 29(1):357–404

Fort J, Pérez-Losada J (2016) Can a linguistic serial founder effect originating in Africa explain the worldwide phonemic cline? J R Soc Interface 13(117):20160185

Foster PL (1999) Mechanisms of stationary phase mutation: a decade of adaptive mutation. Annu Rev Genet 33(1):57–88

Gavin MC, Sibanda N (2012) The island biogeography of languages. Glob Ecol Biogeogr 21(10):958–967

Gavin MC, Botero CA, Bowern C, Colwell RK, Dunn M, Dunn RR, Gray RD, Kirby KR, McCarter J, Powell A (2013) Toward a mechanistic understanding of linguistic diversity. Bioscience 63(7):524–535

Gavin MC, McCarter J, Mead A, Berkes F, Stepp JR, Peterson D, Tang R (2015) Defining biocultural approaches to conservation. Trends Ecol Evol 30(3):140–145

Gorenflo LJ, Romaine S, Mittermeier RA, Walker-Painemilla K (2012) Co-occurrence of linguistic and biological diversity in biodiversity hotspots and high biodiversity wilderness areas. Proc Nat Acad Sci USA 109(21):8032–8037

Gray RD, Drummond AJ, Greenhill SJ (2009) Language phylogenies reveal expansion pulses and pauses in pacific settlement. Science 323(5913):479–483

Gray RD, Bryant D, Greenhill SJ (2010) On the shape and fabric of human history. Philos Trans R Soc Lond B Biol Sci 365(1559):3923–3933

Greenhill SJ, Gray RD (2012) Basic vocabulary and Bayesian phylolinguistics: issues of understanding and representation. Diachronica 29(4):523–537

Grollemund R, Branford S, Bostoen K, Meade A, Venditti C, Pagel M (2015) Bantu expansion shows that habitat alters the route and pace of human dispersals. Proc Natl Acad Sci 112(43):13296–13301. https://doi.org/10.1073/pnas.1503793112

Guernier V, Hochberg ME, Guégan J-F (2004) Ecology drives the worldwide distribution of human diseases. PLoS Biol 2(6):e141

Hajek J (2004) Consonant inventories as an areal feature of the new guinea-pacific region: testing Trudgill’s hypotheses. Linguist Typol 8(3):343–350

Hamed MB (2005) Neighbour-nets portray the Chinese dialect continuum and the linguistic legacy of China’s demic history. Proc R Soc Lond B Biol Sci 272(1567):1015–1022

Harvey PH, Pagel M (1991) The comparative method in evolutionary biology. Oxford University Press, Oxford

Harvey PH, Pagel M, Read A, Nee S (1995) Further remarks on the role of phylogeny in comparative ecology. J Ecol 140:421–446

Hauser MD, Yang C, Berwick RC, Tattersall I, Ryan MJ, Watumull J, Chomsky N, Lewontin RC (2014) The mystery of language evolution. Front Psychol 5:401. https://doi.org/10.3389/fpsyg.2014.00401

Hawkesworth J (1773) An account of the voyages undertaken… For making discoveries in the southern hemisphere, and successively performed by Commodore Byron, Captain Wallis, Captain Carteret, and Captain Cook… Drawn up from the journals which were kept by the several commanders, and from the papers of Joseph Banks, esq; by John Hawkesworth, ll. D. Illustrated with cuts, and… Charts and maps. For W. Strahan, and T. Cadell

Hay J, Bauer L (2007) Phoneme inventory size and population size. Language 83(2):388–400

Haynie HJ, Bowern C (2016) Phylogenetic approach to the evolution of color term systems. Proc Natl Acad Sci USA 113(48):13666–13671

Hoenigswald HM (1963) On the history of the comparative method. Anthropol Linguist 5(1):1–11

Horne Tooke J (1840) Epea pteroenta, or, the diversions of Purley. Thomas Tegg, London

Hunley K, Bowern C, Healy M (2012) Rejection of a serial founder effects model of genetic and linguistic coevolution. Proc R Soc Lond B Biol Sci 279(1736):2281–2288

Huxley T (1887) On the reception of the ‘origin of species’. In: Darwin F (ed) Life and letters of Charles Darwin. John Murray, London

Jäger G (2008) Applications of game theory in linguistics. Lang Linguist Compass 2(3):406–421

Jones W (1788) The third anniversary discourse on the Hindus. Asiat Res 1:422–423

Keeling PJ, Palmer JD (2008) Horizontal gene transfer in eukaryotic evolution. Nat Rev Genet 9(8):605–618

Kirby S, Griffiths T, Smith K (2014) Iterated learning and the evolution of language. Curr Opin Neurobiol 28:108–114

Kirch PV (2001) On the road of the winds: An archaeological history of the pacific islands before european contact. University of California Press, Berkeley

Kusuma P, Cox MP, Pierron D, Razafindrazaka H, Brucato N, Tonasso L, Suryadi HL, Letellier T, Sudoyo H, Ricaut F-X (2015) Mitochondrial DNA and the y chromosome suggest the settlement of madagascar by indonesian sea nomad populations. BMC Genom 16(1):191

Lenski RE, Mittler J (1993) The directed mutation controversy. Science 259:188–194

Levinson SC, Gray RD (2012) Tools from evolutionary biology shed new light on the diversification of languages. Trends Cogn Sci 16(3):167–173

List J-M, Nelson-Sathi S, Geisler H, Martin W (2014) Networks of lexical borrowing and lateral gene transfer in language and genome evolution. BioEssays 36(2):141–150

List J-M, Pathmanathan JS, Lopez P, Bapteste E (2016) Unity and disunity in evolutionary sciences: process-based analogies open common research avenues for biology and linguistics. Biol Direct 11:39

Lupyan G, Dale R (2010) Language structure is partly determined by social structure. PLoS ONE 5(1):e8559

Luszki MB (1958) Interdisciplinary team research: methods and problems. New York University Press, New York

Lyell C (1860) Scientific journal no vi. In: Wilson LG (ed) Sir Charles Lyell's scientific journals on the species question. Yale University Press, New Haven

Lyell C (1863) The geological evidence of the antiquity of man. John Murray, London

Mace R, Pagel M (1995) A latitudinal gradient in the density of human languages in North America. Proc R Soc Lond B 261(1360):117–121

Marris E (2008) The language barrier: some researchers think that the evolution of languages can be understood by treating them like genomes–but many linguists don't want to hear about it. Nature 453(7194):446–449

Martin WF (2011) Early evolution without a tree of life. Biol Direct 6(1):36

Mastroiacovo D, Kwik-Uribe C, Grassi D, Necozione S, Raffaele A, Pistacchio L, Righetti R, Bocale R, Lechiara MC, Marini C, Ferri C, Desideri G (2015) Cocoa flavanol consumption improves cognitive function, blood pressure control, and metabolic profile in elderly subjects: the cocoa, cognition, and aging (CoCoA) study—a randomized controlled trial. Am J Clin Nutr 101(3):538–548. https://doi.org/10.3945/ajcn.114.092189

Matzke NJ (2016) The evolution of antievolution policies after Kitzmiller versus Dover. Science 351(6268):28–30

Maurage P, Heeren A, Pesenti M (2013) Does chocolate consumption really boost Nobel award chances? The peril of over-interpreting correlations in health studies. J Nutr 143(6):931–933

Maynard Smith J (1979) Hypercycles and the origin of life. Nature 280(5722):445–446

Maynard Smith J, Price GR (1973) The logic of animal conflict. Nature 246(5427):15–18

Maynard Smith J, Szathmáry E (1995) The major transitions in evolution. W. H. Freeman, Oxford

McConvell P, Meakins F (2005) Gurindji Kriol: a mixed language emerges from code-switching. Aust J Linguist 25(1):9–30

McWhorter J (2001) The world's simplest grammars are creole grammars. Linguist Typol 5(2/3):125–166

Meakins F (2011) Case marking in contact: the development and function of case morphology in Gurindji Kriol. John Benjamins, Amsterdam

Meakins F (2012) Which mix?—code-switching or a mixed language—Gurindji Kriol. J Pidgin Creole Lang 27(1):105–140

Meakins F, Hua X, Algy C, Bromham L (submitted) Birth of a new contact language does not favour simplification. Language Submitted April 2017

Mesoudi A, Whiten A, Laland K (2006) Towards a unified science of cultural evolution. Behav Brain Sci 29(4):329

Messerli FH (2012) Chocolate consumption, cognitive function, and Nobel laureates. N Engl J Med 367(16):1562–1564. https://doi.org/10.1056/NEJMon1211064

Miles DB, Dunham AE (1993) The use of phylogenetic comparative analyses. Annu Rev Ecol Syst 24:587–619

Moran S, McCloy D, Wright R (2012) Revisiting population size vs. phoneme inventory size. Language 88(4):877–893

Mower JP, Stefanović S, Young GJ, Palmer JD (2004) Plant genetics: gene transfer from parasitic to host plants. Nature 432(7014):165–166

Murray DR (2014) Direct and indirect implications of pathogen prevalence for scientific and technological innovation. J Cross Cult Psychol 45(6):971–985

Murray DR, Schaller M, Suedfeld P (2013) Pathogens and politics: further evidence that parasite prevalence predicts authoritarianism. PLoS ONE 8(5):e62275

Naroll R (1965) Galton's problem: the logic of cross-cultural analysis. Soc Res 32(4):428–451

Nettle D (1998) Explaining global patterns of language diversity. J Anthropol Archaeol 17(4):354–374

Nowak MA, Plotkin JB, Krakauer DC (1999) The evolutionary language game. J Theor Biol 200(2):147–162

O'Connor C (2017) Games and kinds. Br J Philos Sci (in press). Available at http://philsci-archive.pitt.edu/12606/

Pagel MD, Harvey PH (1992) On solving the correct problem: wishing does not make it so. J Theor Biol 156(4):425–430

Pagel M, Atkinson QD, Calude S, Meade A (2013) Ultraconserved words point to deep language ancestry across Eurasia. Proc Natl Acad Sci USA 110(21):8471–8476

Pennell MW, Harmon LJ (2013) An integrative view of phylogenetic comparative methods: connections to population genetics, community ecology, and paleobiology. Ann N Y Acad Sci 1289(1):90–105

Pollet TV, Tybur JM, Frankenhuis WE, Rickard IJ (2014) What can cross-cultural correlations teach us about human nature? Hum Nat 25(3):410–429

Prichard JC (1832) Abstract of a comparative review of philological and physical researches as applied to the history of the human species. Br Assoc Adv Sci Rep 2:529–544

Rao V, Nanjundiah V (2011) J. B. S. Haldane, Ernst Mayr and the beanbag genetics dispute. J Hist Biol 44(2):233–281

Reali F, Griffiths TL (2010) Words as alleles: connecting language evolution with bayesian learners to models of genetic drift. Proc R Soc B Biol Sci 277(1680):429–436

Repko A (2007) Integrating interdisciplinarity: how the theories of common ground and cognitive interdisciplinarity are informing the debate on interdisciplinary integration. Issues Interdiscip Stud 25:1–31

Richards RJ (2002) The linguistic creation of man: Charles Darwin, August Schleicher, Ernst Haeckel, and the missing link in Nineteenth-Century evolutionary theory. In: Dörries M (ed) Experimenting in tongues: studies in science and language. Standford University Press, Stanford, pp 21–48

Roberts S, Winters J (2013) Linguistic diversity and traffic accidents: lessons from statistical studies of cultural traits. PLoS ONE 8(8):e70902

Rosenberg SM (1997) Mutation for survival. Curr Opin Genet Dev 7(6):829

Ross CT, Winterhalder B (2016) A hierarchical bayesian analysis of parasite prevalence and sociocultural outcomes: the role of structural racism and sanitation infrastructure. Am J Hum Biol 28(1):74–89

Schaller M, Murray DR (2010) Infectious diseases and the evolution of cross-cultural differences. Evol Cult Hum Mind 51:243–256

Schleicher A (1869) Darwinism tested by the science of language (trans Bikkers VW). John Camden Hotten

Schneemann H, Hua X, Greenhill SJ, Bromham L (2017) Parasites and politics: no evidence for a direct causal effect. In preparation

Shipley B (2002) Cause and correlation in biology: a user's guide to path analysis, structural equations and causal inference. Cambridge University Press, Cambridge

Smith ADM (2014) Models of language evolution and change. Wiley Interdiscip Rev Cogn Sci 5(3):281–293

Sokolov AN, Pavlova MA, Klosterhalfen S, Enck P (2013) Chocolate and the brain: neurobiological impact of cocoa flavanols on cognition and behavior. Neurosci Biobehav Rev 37(10):2445–2453. https://doi.org/10.1016/j.neubiorev.2013.06.013

Starostin S (2000) Comparative-historical linguistics and lexicostatistics. Time Depth Hist Linguist 1:223–265

Szathmáry E (2013) On the propagation of a conceptual error concerning hypercycles and cooperation. J Syst Chem 4(1):1

Taub L (1993) Evolutionary ideas and 'empirical' methods: the analogy between language and species in works by Lyell and Schleicher. Br J Hist Sci 26(02):171–193

Turvey ST, Pettorelli N (2014) Spatial congruence in language and species richness but not threat in the world's top linguistic hotspot. Proc R Soc B Biol Sci 281(1796):20141644

Tybur JM, Inbar Y, Aarøe L, Barclay P, Barlow FK, de Barra M, Becker DV, Borovoi L, Choi I, Choi JA, Consedine NS, Conway A, Conway JR, Conway P, Adoric VC, Demirci DE, Fernández AM, Ferreira DCS, Ishii K, Jakšić I, Ji T, van Leeuwen F, Lewis DMG, Li NP, McIntyre JC, Mukherjee S, Park JH, Pawlowski B, Petersen MB, Pizarro D, Prodromitis G, Prokop P, Rantala MJ, Reynolds LM, Sandin B, Sevi B, De Smet D, Srinivasan N, Tewari S, Wilson C, Yong JC, Žeželj I (2016) Parasite stress and pathogen avoidance relate to distinct dimensions of political ideology across 30 nations. Proc Natl Acad Sci USA 113(44):12408–12413. https://doi.org/10.1073/pnas.1607398113

Tylor EB (1889) On a method of investigating the development of institutions; applied to laws of marriage and descent. J Anthropol Inst G B Irel 18:245–272

Van Wyhe J (2005) The descent of words: evolutionary thinking 1780–1880. Endeavour 29(3):94–100

Wallace AR (1855) On the law which has regulated the introduction of species. Ann Mag Nat Hist 16:184–196

Wallace AR (1872) The Malay Archipelago: the land of the orang-utan and the bird of paradise, a narrative of travel, with studies of man and nature. Macmillan and Co, London

Weisberg M (2007) Who is a modeler? Br J Philos Sci 58(2):207–233

Welch JJ (2016) What's wrong with evolutionary biology? Biol Philos 32(2):263–279

Westoby M, Leishman MR, Lord JM (1995) On misinterpreting the phylogenetic correction. J Ecol 83(3):531–534

Whitfield J (2008) Across the curious parallel of language and species evolution. PLoS Biol 6(7):e186

Willems M, Lord E, Laforest L, Labelle G, Lapointe F-J, Di Sciullo AM, Makarenkov V (2016) Using hybridization networks to retrace the evolution of Indo-European languages. BMC Evol Biol 16(1):180. https://doi.org/10.1186/s12862-016-0745-6

Wilson LG (1970) Sir Charles Lyell's scientific journals on the species question. Yale University Press, New Haven

Wright S (1931) Evolution in Mendelian populations. Genetics 16:97–159