Evolutionary Linguistics

William Croft

Department of Linguistics, University of New Mexico, Albuquerque, New Mexico 87131;
email: wcroft@unm.edu

Key Words

replicator, selection, phylogeny, comparative method

Abstract

Both qualitative concepts and quantitative methods from evolutionary
biology have been applied to linguistics. Many linguists have noted the
similarity between biological evolution and language change, but usually
have employed only selective analogies or metaphors. The development
of generalized theories of evolutionary change (Dawkins and Hull) has
spawned models of language change on the basis of such generalized
theories. These models have led to the positing of new mechanisms of
language change and new types of selection that may not have biolog-
ical parallels. Quantitative methods have been applied to questions of
language phylogeny in the past decade. Research has focused on widely
accepted families with cognates already established by the comparative
method (Indo-European, Bantu, Austronesian). Increasingly sophisti-
cated phylogeny reconstruction models have been applied to these fam-
ilies to resolve questions of subgrouping, contact, and migration. Little
progress has been made so far in analyzing sound correspondences in
the cognates themselves.

INTRODUCTION

Evolutionary models have come to be employed in several areas of the study of language in the past two decades. The use of evolutionary models is naturally found in historical linguistics and also in the study of the origins of language. In the latter case, however, the employment of evolutionary models is handicapped by the absence of data regarding the transition from our primate ancestors to the emergence of modern human language, which is found in all societies. All that we can go by is the archaeological record and the comparison of the social-cognitive abilities and communication systems of humans and other animals, particularly nonhuman primates. Because the study of the origin of human language does not depend on linguistic data, it is not discussed in this article. Even so, the area under review is vast and growing, and therefore this review is restricted to research in which qualitative concepts and quantitative methods from evolutionary biology have been applied to the analysis of language, in particular language change and language phylogeny.

EVOLUTIONARY THEORY AND THEORIES OF LANGUAGE CHANGE

In historical linguistics, the parallels between biological and linguistic evolution have been observed since Darwin himself first took notice (for a historical survey, see Atkinson & Gray 2005). However, the differences in the domains of biology and language appear to have outweighed the similarities, and Darwinian evolutionary theory has developed over time. In the meantime, the advent of structuralism and generative grammar has led to the dominance of an ahistorical approach to the study of language (Croft 2002). As a consequence, linguistics has rarely used models from evolutionary biology. Nevertheless, the similarities between the two have led historical linguists to employ evolutionary analogies or metaphors. Analogies/metaphors indicate similarities between the two domains (biological evolution and language change) but do not imply an overarching generalized theory.

An isolated analogy from evolutionary biology that has proven to be useful in explaining language change is Lass's application of exaptation to historical linguistic phenomena (Lass 1990). Exaptation in biology is the employment of a phylogenetic trait for a function different from the one for which it was originally adapted; Lass slightly changes the definition to apply to linguistic structures that have lost their function but have come to be employed for another function.

A recent example of the employment of a biological metaphor is Blevins's theory of evolutionary phonology (Blevins 2004). Evolutionary phonology proposes to account for synchronic phonological patterns as the result of phonetically motivated changes in the transmission of sound systems from adult to child over time. It uses the notions of inheritance (via the child learning the adults' language), variation generated by "errors" in replication (mechanisms by which the listener alters what he hears from the speaker), and natural selection (certain sound changes are more/less likely in particular phonetic contexts). However, because of disanalogies between biological evolution and language change, Blevins explicitly rejects an evolutionary approach to sound change that is more than metaphorical.

Although analogies or metaphors between biological evolution and language change can be fruitful, one does not know which parallels between the two domains are legitimate to draw and which are not, or even more important, which parallel structures must be present for the analogy/metaphor to make sense. In particular, it is common to assume that the mechanisms that cause variation and selection in biological evolution must be the same in other domains such as language change, yet the mechanisms are domain specific. What is required is a generalized theory of evolutionary change that subsumes biological evolution, language change, and other phenomena of evolutionary change such as cultural evolution. Researchers have derived models of cultural evolution from

---

Phylogeny: a graph structure, usually a tree, representing the evolutionary history of a set of individuals (organisms, species, languages)

Selection: the process by which the interaction of the interactor with its environment causes differential replication of the relevant replicators

biological evolution (Boyd & Richerson 1985, Cavalli-Sforza & Feldman 1981, Durham 1991, Richerson & Boyd 2005). But the generalized theories of evolutionary change that have attracted the most attention in historical linguistics are those developed by Dawkins (1989, 1982) and Hull (1988, 2001).

The most crucial feature of a generalized theory of evolutionary change is that evolutionary change is change by replication, a process by which some entity is copied in such a way that most or all of the structure of the replicate is the same as that of the original. The replication process is cumulative and iterative, leading to lineages. Second, evolutionary change is a two-step process: the generation of variation in the replication process, and the selection of variants via some mechanism. Dawkins' and Hull's models have these properties, as do the models of cultural evolution mentioned above. In the context of language change, Lass notes that these properties are necessary to understand languages as historical entities (Lass 1997, pp. 109–11), although he does not develop a detailed theory of language change on this basis (see also Nettle 1999, Wedel 2006). Evolutionary theory also rejects any notion of progress.

Dawkins' and Hull's models are related but differ in important respects. Dawkins generalizes the concept of a gene as a replicator. The replicator possesses a structure that can be copied. Variation occurs through the copying process. Dawkins also generalizes the concept of an organism as a vehicle. For Dawkins, a vehicle has been constructed by the replicator to aid in its replication. Selection is differential replication, but for Dawkins it is focused on the replicator (the selfish gene concept), whereas in more standard neo-Darwinian models, selection is a function of the organism/vehicle.

Dawkins proposed that there are units of culture with heritable structure that replicate; he called them memes and proposed a science of memetics to study them. Memes have been interpreted as cognitive entities in the minds of humans, instances of human behavior, or artifacts. Most memeticists, including Dawkins himself, assume memes to be cognitive enti-

ties. Hence concepts are replicators and minds or brains of individuals are vehicles. Memeticists generally use a parasite-host model for the relationship between memes/concepts and the mind or brain of the possessor: Memes are parasites that use the brain (the host) as their vehicle for replication.

The most extended analysis of language change in Dawkinsian memetic terms is by Ritt (2004). Ritt, a historical phonologist, focuses on phonological change. Following Dawkins, linguistic memes are concepts in the mind; specifically, they are some type of replicable brain structure. Ritt argues that phonemes, morphemes, phonotactic patterns, metrical feet, and phonological rules, or more precisely their conceptual representations, are memes. However, linguistic signs (form-meaning pairings) are not replicators because, in Ritt's view, they do not preserve enough structure in replication. Instead, signs are the result of an alliance of replicators.

In the Dawkinsian model, the linguistic behavior that a speaker produces on the basis of her conceptual memes exists for the purpose of replicating the memes, not for communication (Ritt 2004, p. 231); this is the selfish gene/meme theory. The replicators are replicated across speakers by an imitation process (see Blackmore 1999); variation is generated in imperfect imitation. Ritt proposes several selectional pressures for differential replication (i.e., selection), all operating in the mind/brain. He focuses his attention on meme coadaptation as a selectional pressure, using it to account for the interaction between foot structure and vowel changes in the history of English.

Hull adopts Dawkins' concept of replicator but generalizes the role of the organism to an interactor and defines an interactor as any entity that interacts with the environment so as to cause differential replication (that is, selection) of the relevant replicators. Thus, the interactor's interaction with the environment is the locus of selection; it is not a mere vehicle for the replication of the replicator, and Hull does not advocate the selfish gene/meme interpretation that negates the role of the interactor in

Replicator: an entity that is copied and preserves most of its structure in copying

Meme: a cultural replicator

Interactor: an entity that interacts with its environment in such a way as to cause differential replication of the relevant replicators

Lingueme: a linguistic replicator, that is, a token of linguistic structure produced in an utterance

evolution. Hull argues that interactors may exist at different levels in the biological hierarchy (gene, cell, organism, even a species), although the organism's interaction with its environment in natural selection is the canonical case. Hull's general analysis of selection is thus centered on the processes by which the interactor's interaction with its environment causes the differential replication of replicators.

Hull (1988, 2001) presents a theory of conceptual change in science in which scientific concepts are replicators and the scientists are interactors. Hull's general analysis of selection does not assume any specific causal relationship from replicator to interactor, so there is no need to invoke a parasite-host model for the concept-scientist relation. However, scientists must be able to cause the differential replication of their ideas/concepts, which they do through publishing, lecturing, and teaching.

The most detailed application of Hull's general analysis of selection to language change is in Croft (2000, 2002, 2006). Croft argues for a model in which the linguistic replicators are behaviors, that is, tokens of linguistic structures in utterances produced by speakers. Croft coins the term lingueme to describe the linguistic replicator. The speakers themselves are interactors. The speaker replicates the replicators in speaking, generating variation in the production and comprehesion of utterances. In Croft's model, linguistic structures evolve via language use, not via language acquisition.

Croft's model, like Hull's, does not specify the mechanisms by which variation is generated. Croft, like all evolutionary biologists and most historical linguists, rejects teleological mechanisms. Croft allows for widely proposed intentional mechanisms, such as expressiveness and avoidance of misunderstanding. Croft also proposes nonintentional mechanisms inspired by theories of sound change, in which speakers or listeners attempt to conform to convention but fail to do so. Speakers are highly variable in the phonetic realization of phonemes (sound linguemes), as noted in much recent phonetic

research (e.g., Bybee 2001, Pierrehumbert 2001) and in the verbalization of meaning in grammatical structures (Croft 2008). Listeners are faced with the problem of analyzing the phonemes in an utterance from a complex acoustical signal, and they may reanalyze the mapping between the phonetic signal and the phonological structure, in processes Ohala (2003) calls hypocorrection and hypercorrection. Listeners are also faced with the problem of analyzing the semantic contribution of words, morphemes, and constructions from a complex communicative situation, and they may reanalyze the mapping between form and meaning in those units via different types of form-function reanalysis (Croft 2000, chapters 4–5).

The speaker as interactor is also the locus of selection: The speaker selects a variant to produce. In this respect, Croft's evolutionary model agrees with theories of the propagation of change (selection of variants) developed in sociohistorical linguistics. In the latter theories, investigators propose that various social factors associated with particular sociolinguistic variants in speech communities lead to the propagation (or extinction) of variants, although other factors including the social network structure and the frequency of exposure to variants also play a role. Croft argues that functional pressures operate only in the generation of variation, not in selection; others take the view that functional pressures operate in selection (e.g., Nettle 1999, pp. 30–35).

Some of these models have led to mathematical formalizations and simulations. Nettle (1999) simulates a model in which language change occurs via child language acquisition and argues that the rate of fixation is proportional to the number of speakers in the speech community. Wedel (2006) presents a usage/exemplar-based evolutionary model of sound change, including simulations of inheritance and selection. The work by Baxter et al. (2006) is a formalization of Croft's theory in a statistical physics framework (see also Blythe & McKane 2007).

The model developed in Baxter et al. (2006) shows that some types of selection mechanisms in language change (and other types of cultural evolution) may not exist as such in biological evolution. The classic selection mechanism is a type of replicator selection, that is, differential weighting of replicator variants (their fitness); in language change this is found in the differential social valuation of variants (compare Nettle 1999, pp. 29–30). In neutral evolution (genetic drift, not the same as linguistic drift), random fluctuations lead to change (Nettle 1999, pp. 16–17); these correspond to frequency effects in language change. In addition to these two mechanisms, there is neutral interactor selection, by which differences in the frequency of interaction with other interactors lead to differential replication. Neutral interactor selection corresponds to social network structure effects in language change (Milroy 1987); sexual selection may be an instance of neutral interactor selection in biological evolution. Finally, the model includes weighted interactor selection, in which differential weighting of the interactors with whom one interacts, independent of frequency of interaction, leads to differential replication. Weighted interactor selection is exemplified in the differential social valuation of different speakers; there is no obvious equivalent in biological evolution. Baxter et al. (2008) use their model and simulations based on it to argue that the New Zealand English variety could not have emerged by neutral evolution and neutral interactor selection alone.

Mufwene (2001, 2005) developed an evolutionary model for language change that is similar to Croft's. Mufwene focuses on two aspects of the evolutionary framework not discussed above. First, a language forms a population in the same way that a species does. Mufwene treats a language as a species, specifically a population of linguistic structures that exist in the minds of speakers in communities, because languages are variable, and their spread, extinction, and rates of change are dependent on speaker populations and not on the linguistic system per se. Mufwene follows the parasite-host model

for the relationship between a language and its speakers, but his focus is different from the memeticists. The memeticists argue that the linguistic concept (the parasite) uses the speaker (its host) to replicate itself, whereas Mufwene argues that the survival, spread, or extinction of a language is dependent on the survival, spread, or extinction of its host speakers.

The second aspect of evolution that Mufwene exploits is ecology. Languages, and the speakers on which they are dependent, are embedded in an environment, in particular a social environment but also the internal environment of the society and the linguistic varieties found in it. Mufwene uses his framework to analyze the development of creoles and the relationship between creoles and "normal" linguistic transmission. Mufwene argues that creoles emerge from the linguistic varieties available in the earliest stages of colonization, which in the case of European language–based creoles involve nonstandard varieties of the European language(s) and also nonnative European speakers (e.g., Irish nonnative speakers of English, Breton nonnative speakers of French) as well as nonnative speakers of African origin (as in American slave plantations) or indigenous nonnative speakers (as in the Pacific). The emerging creole is a result of natural selection of linguistic structures determined by the ecology of the social and economic situations of the early colonies. Mufwene argues that the emergence of creoles is not different in kind from "normal" language change, but only in degree—the same ecological model fits the development of the Romance languages from Vulgar Latin, for instance.

The emergence of theories of language change based on a generalized theory of evolutionary change is quite recent. Less systematic adaptations of evolutionary concepts in theories of language change continue to be made as well. Evolutionary, usage-based theories of language change are the strongest competitors to the innate-grammar, child-based theories of language change put forward by followers of Chomsky such as Lightfoot (2006).

PHYLOGENY RECONSTRUCTION AND GENETIC LINGUISTICS

A major potential application of more practical methods from evolutionary biology to historical linguistics is in the area of genetic linguistics, in particular the establishment of language families and their subgroupings. This task parallels what evolutionary biologists call phylogeny reconstruction, the reconstruction of the presumed historical branching of ancestral populations into contemporary populations, of either species or smaller biological groups. Phylogeny reconstruction is accomplished by comparing phenotypic traits of organisms, particularly proteins, or sequences of nucleotides in mitochondrial or nuclear DNA. Breakthroughs in DNA sequencing and the development of mathematical algorithms and computing power to execute those algorithms have led to an explosion of research in phylogeny reconstruction, of which the most famous result is the African Eve hypothesis: that all humans are believed to be descended from an ancestral population in Africa some 100,000–150,000 years ago (Cann et al. 1987, Vigilant et al. 1991).

The connection between phylogeny and language history was made in a paper by Cavalli-Sforza et al. (1988), which attracted a tremendous amount of attention but whose linguistic assumptions were rejected by most linguists. Cavalli-Sforza et al. produce a phylogeny of human populations and compare it with a phylogeny of linguistic populations, that is, the major language families in the world. They note a high degree of congruence of the two phylogenies, suggesting that language spread and diversification have occurred primarily via migration and splitting of speech communities, at least in prehistory. Much of the controversy among linguists is due to the authors' use of language families such as Amerind, Eurasiatic/Nostratic, Altaic, Na-Dene, Austric, and Indo-Pacific, which are not generally accepted among historical lingusts.

One crucial difference between the biological and linguistic data in Cavalli-Sforza et al. is that the biological data have been analyzed quantitatively whereas the linguistic data have not. Since that time an increasing number of researchers have attempted to apply quantitative techniques to the problem of language families and have made other attempts to relate language diversification to prehistoric demography. The remainder of this review focuses on quantitative methods from evolutionary biology applied to comparative historical (genetic) linguistics; owing to topic and length considerations, quantitative methods from other fields are not discussed nor is research on the relationship between prehistoric demography and language families derived by traditional linguistic method.

The Comparative Method and Phylogeny Reconstruction

The comparative method is excellently summarized in the following passage from Ross & Durie (1996, p. 7):

The comparative method (in its strict sense) can be summarized as a set of instructions:

1. Determine on the strength of diagnostic evidence that a set of languages are genetically related, that is, that they constitute a "family."
2. Collect putative cognate sets for the family (both morphological paradigms and lexical items).
3. Work out the sound correspondences from the cognate sets, putting "irregular" cognate sets on one side.
4. Reconstruct the protolanguage of the family as follows:
   a. Reconstruct the protophonology from the sound correspondences worked out in [step] 3, using conventional wisdom regarding the directions of sound changes.
   b. Reconstruct protomorphemes (both morphological paradigms and lexical items) from the cognate sets collected in [step] 2, using the protophonology reconstructed in [step] 4a.
5. Establish innovations (phonological, lexical, semantic, morphological,

morphosyntactic) shared by groups of languages within the family relative to the reconstructed protolanguage.

6. Tabulate the innovations established in [step] 5 to arrive at an internal classification of the family, a "family tree."

7. Construct an etymological dictionary, tracing borrowings, semantic change, and so forth, for the lexicon of the family (or of one language of the family).

Ross & Durie note that the steps in this process are iterated because they are clearly interrelated. For example, what counts as an innovation in sound change in a daughter language or languages depends on what is reconstructed as a protophoneme.

Major similarities exist between the comparative method and phylogeny reconstruction (Greenberg 1992 gives a detailed comparison; see also Lass 1997, chapters 3–4). First, the differentiation of languages and the differentiation of species are assumed to be largely treelike. Nevertheless, both biologists and linguists allow for the possibility of reticulation in phylogenetic trees, certainly for closely related species (hybridization) and languages (dialect mixture). Second, both biologists and historical linguists recognize the important diagnostic value of shared innovations (see steps 5–6 above). These fundamental similarities mean that the tree-building algorithms developed for biological phylogenies are built on the same principles as is historical linguistics and therefore should be applicable to reconstructing linguistic family trees, or language phylogenies.

A major difference between the two domains, however, is that in language, sound change leads the divergence of the phonological forms of words even though they are cognate. Cognate forms are not identical. In genetic comparison, "cognate" sequences of nucleotides are identical. One consequence of this difference is that almost all applications of phylogeny reconstruction algorithms from biology to historical linguistics have used as the input data (character traits) to the phylogeny reconstruction algorithm only lists of established cognates in accepted language families because cognates can be treated as identical. That is, the family tree (the subgrouping) is reconstructed on the basis of the presence versus absence of cognate forms for particular meanings in particular daughter languages. This approach represents a signficant loss of information, but many important issues in historical linguistics can nevertheless be placed in an evolutionary perspective even in this approach.

Some techniques from evolutionary biology are beginning to the applied to the problem of sound correspondence. In particular, techniques from DNA sequencing have been applied to the alignment problem. This is a problem rarely touched on in textbooks on the comparative method: aligning two forms to identify the corresponding sounds. This technique is not straightforward because phonemes may be inserted (epenthesis), deleted, merged, or transposed (metathesis); also, many words contain fossilized affixes that do not correspond with anything in their cognates in the other languages. Computational linguists have addressed the alignment problem for historical comparison, using techniques also found in biology (Covington 1996, 1998; Kessler 1995; Kondrak 2002, 2003; Nerbonne & Heeringa 1997; Oakes 2000). This research is based on minimizing edit distance (Levenshtein distance) between the strings being aligned. Kondrak notes that biologists are beginning to use probabilistic models such as Hidden Markov models and suggests this as a possible new technique to use in historical linguistics (Kondrak 2002, p. 23). Nevertheless, mathematical formalization of the identification of sound correspondences is presently in its infancy.

Mathematical Techniques from Biology

A wide array of techniques from phylogeny reconstruction in biology have been applied to historical linguistics, even given the limitation to cognate judgments that exists only in

established language families. This review can give only a broad qualitative description of some of the methods; the mathematics of the methods is described in various references (see also McMahon & McMahon 2003, 2006 for discussion of the methods from a comparative historical linguistics perspective).

Given a set of data such as cognate judgments among a set of genetically related languages, two common ways are used to construct phylogenies. Distance-based methods use an overall pairwise distance measurement between two languages. Historical linguistics has already used a distance-based method to measure the proportion of shared cognates: the lexicostatistical method. (Lexicostatistics has also been used when similar word forms are not known to be cognate; here we discuss only assumed cognacy.) Distance-based methods have the advantage of being computationally very fast. However, this method suffers from a number of defects (Atkinson & Gray 2005, p. 520; Gray & Atkinson 2003, p. 436). The distance measure loses the information about which word forms are cognate. Languages change at different rates (Bergsland & Vogt 1962, Blust 2000, Gray & Atkinson 2003). Individual word meanings also change at different rates (Greenberg 2005, pp. 108–11; Joos 1964; Kruskal et al. 1973; Pagel 2000; Pagel et al. 2007; Pagel & Meade 2006). Finally, borrowing can distort the distance between languages in ways that reflect contact, not common ancestry. However, all these problems can be overcome to some extent with newer phylogenetic techniques, briefly described in this and the following sections.

The loss of information in distance-based methods can be addressed by using character-based methods. In character-based methods, particular cognate forms are used to compute the tree. Character-based methods are most closely associated with cladistics, which distinguishes shared innovations from shared retentions and uses only the former for constructing the tree (subgrouping). Two approaches have been used for determining the best fit of the tree to the data. In maximum parsimony, the algo-

rithm minimizes the number of character state changes (in the case of cognate lists, minimizes the number of replacements of word forms). An alternative approach is the compatibility criterion, which minimizes the number of characters that must be assumed to have been innovated more than once. Indo-European trees have been constructed using both maximum parsimony (Rexová et al. 2003) and compatibility (Ringe et al. 2002).

Owing to the complexity of the data, any phylogeny reconstruction algorithm will produce large numbers of trees, measured according to the criteria used (e.g., maximum parsimony). In fact, the space of possible trees is often so large that heuristics must be used to identify the trees to be used in analysis. The traditional approach in both distance-based and character-based methods is to base the result on a set of optimal trees. A consensus tree is produced, typically using a majority rules strategy, which posits the nodes in the consensus tree that are found in at least half the input trees. The robustness of the nodes in the tree relative to the data is commonly tested with a bootstrapping technique (e.g., Holden et al. 2005, p. 60). In bootstrapping, the original data set is sampled with replacement (i.e., one is always sampling from the full original data set) until a new data set of the same size is produced. A pseudoreplicate tree is constructed for the new data set, and the process is repeated many times. The pseudoreplicate trees are compared with the original (consensus) tree, and the robustness of each node in the original tree is the percent of the pseudoreplicate trees that contain the node in question.

More recently, Bayesian methods have been applied to phylogeny reconstruction (Atkinson & Gray 2005, p. 521). In a Bayesian approach, one samples not just the most optimal trees produced, but all possible trees in proportion to their likelihood (Holden et al. 2005, pp. 60–62). Again, a consensus tree is produced from the sample. The proportion of trees with a particular node in the sample is equivalent to the Bayesian posterior probability of that node. A common method used to construct

the sample to model the posterior probability distribution of trees is a Markov chain Monte Carlo method (see, for example, Gray & Atkinson 2003, Pagel & Meade 2006). The Bantu languages have been classified using a distance-based method (Bastin et al. 1999), maximum parsimony (Holden 2002, Rexová et al. 2006), and Bayesian methods (Holden et al. 2005, p. 60; Rexová et al. 2006).

Two other facts about phylogeny reconstruction must be noted here. First, trees produced by the algorithms are unrooted: They show groupings but do not indicate which group is the most distant. Rooting is generally established by using an outgroup, a group that is agreed to be most distant from all other members of the data set. Second, the algorithms attempt to construct binary trees. The only way to identify multiple branching is through short binary branches and/or failure to construct statistically robust binary branching in part of the tree.

Differentiating Chance Cognation and Borrowing from Cognates

Step 1 in the comparative method, the identification of a set of languages as forming a language family, poses a basic problem. This selection cannot be done randomly because even with a small number of languages, the number of ways in which they can be classified quickly becomes astronomical (see, e.g., Greenberg 2005, p. 43). Yet most introductions to comparative linguistics assume that the investigator begins with a set of languages that are already related.

Nichols (1996) argues that a single diagnostic trait, or a small set of traits, is sufficient to identify a valid linguistic family, but there are serious statistical problems with this approach (see Kessler 2001, p. 32). It is difficult to find any other technique to identify language families apart from compelling similarities in form and meaning distributed across a set of words among a subset of languages under comparison. This technique is essentially Greenberg's method of multilateral comparison (see the

papers collected in Greenberg 2005). Multilateral comparison has often been misunderstood in the linguistic literature as if Greenberg intended it to replace the comparative method. In fact, he intended it only to be an approach to steps 1 and 2 of the comparative method (Greenberg 2001, p. 127). Step 1 is achieved only jointly with step 2: A set of languages forms a family because of the presence of similar form-meaning pairings that are concluded to be (putative) cognates.

The crucial problem, then, is to differentiate putative cognate forms—those that are phonetically and semantically similar most likely because of common ancestry—from forms in the data that are similar because of contact (borrowing), sound symbolism, or chance. Greenberg's primary contention is that the source of form-meaning similarities in a matrix of word meanings across languages can be fairly accurately determined from the distribution pattern of the form-meaning similarities in the matrix, as well as from phonological patterns in the form-meaning similarities (Greenberg 1957, p. 69; see also Greenberg 2005, chapter 2). Techniques from biology exploiting these distributional patterns can be applied to the problems of chance cognation and borrowing. (Sound symbolism is a fourth source of similarity, but the likelihood of sound symbolism is generally minimized by discounting or excluding meanings that are likely to be sound symbolic—i.e., by exploiting a characteristic distribution pattern across meanings.) This approach differs from the standard assumption among users of the comparative method, namely that only regular sound correspondences can be used to differentiate cognates from borrowing and chance resemblance; therefore, until the establishment of sound correspondences can be formalized, other approaches remain only tentative. Nevertheless, it appears that a considerable amount of information can be extracted from distributional patterns of form-meaning similarities, at least in the relatively shallow language families to which these biological phylogeny reconstruction techniques have been applied.

The likelihood of similarity being due to chance is one that has attracted much attention in historical linguistics, but little can be concluded from the literature on it. Ringe (1992, 1993) and Greenberg & Ruhlen (1992) have proposed methods to evaluate chance cognation, but they are mathematically flawed (Baxter & Manaster Ramer 1996, Greenberg 1993, Kessler 2001, Manaster Ramer & Hitchcock 1996). A more promising approach for evaluating chance cognation is the permutation method, which is not specifically drawn from evolutionary biology (Baxter 1995; Baxter & Manaster Ramer 2000; Justeson & Stephens 1980; Kessler 2001; Oswalt 1970, 1991). The permutation method compares the proportion of form-meaning similarities in the observed data set to random permutations of the forms across the meanings to derive a probability that the observed form-meaning similarities could occur by chance. However, a result indicating that form-meaning similarities in the observed data set are not due to chance gives information only about the data set as a whole. It does not provide information about the validity of particular nodes in the tree. Bootstrapping or Bayesian posterior probabilities, however, can be used to address the chance cognation problem because both give a probability of the validity of each node (genetic grouping) in the tree. Of course, some form-meaning similarities may still be due to chance. The use of character-based methods will provide hypotheses as to which characters (words) are not contributing to the building of the tree and therefore may be chance similarities.

The problem of differentiating borrowing from common ancestry has been addressed in two general ways. The first is to use techniques that assume the data to be treelike and treat anomalous similarities as derived from borrowing. Minett & Wang (2003) examine a distance-based method that compares branch lengths of the tree and lexical distances, but they found that it does not differentiate borrowing from cognates. However, a character-based method using maximum parsimony did allow inference of a likelihood of borrowing for characters (words) among the Chinese languages. Nevertheless, Wang & Minett (2005) formalize Hinnebusch's (1999) proposal that skewing in lexicostatistical percentages may indicate borrowing and suggest that this distance-based approach might be useful.

The second is to use techniques that allow reticulation, that is, they allow branches to rejoin, representing contact relations and creating phylogenetic networks instead of trees. Bryant et al. (2005) use these techniques to represent the conflicting signals of languages that have undergone significant contact. These techniques allow one to ask how treelike the data are, rather than assuming the data are treelike. Bryant et al. show that Indo-European is quite treelike, at least in its basic vocabulary; the same result is reached by Warnow et al. (2006; see also Holden & Gray 2006, discussed below). Ben Hamed (2005) and Ben Hamed & Wang (2006) show that network techniques applied to the complex relationships among dialect differentation, dialect continua, dialect contact, and diglossia in Chinese tend to correlate with known geographical, linguistic, and population history.

One problem with network techniques is that most historical linguists believe that there are few if any instances of true reticulation: One can identify the parent versus the source of contact, and even for so-called mixed languages, the contribution of the two is asymmetric in systematic ways (Croft 2003). Borrowing can be differentiated from common ancestry by the distribution of borrowings and their forms: For instance, some semantic categories are less likely to be borrowed, and borrowing often links the borrowing language to a single source language (e.g., Greenberg 1957, p. 71; 2005, p. 38). McMahon et al. (2005) use different weightings of more versus less stable vocabulary to evaluate the form-meaning resemblances between Quechuan and Aymaran.

Another important route to distinguishing cognates from borrowing is to model the process of word birth, cognate formation, homoplasy (independent convergence, e.g., by chance), borrowing, and word death and

compare the results of this model to that of actual language families. Two examples of models including all these processes are Warnow et al. (2006) and Nicholls & Gray (2006; see McMahon & McMahon 2003 for an earlier model that does not allow for homoplasy). Both of these models were used to refine a phylogeny of Indo-European (see below). Bryant (2006) constructs a model comparing radiation (and subsequent isolation) versus "network breaking" (dialect continua followed by gradual divergence) and compares it with linguistic data on Polynesian.

Language Phylogeny and Human Prehistory

The use of techniques from evolutionary biology in historical linguistics is quite new, and current results must be taken as tentative. The review concludes by mentioning some of the more interesting results that have emerged from these techniques. The most intensive work has been done on Indo-European, the most intensively studied language family, and Bantu and Austronesian, the largest present-day language families, which are nevertheless quite shallow.

Ringe et al. (2002) use not only lexical cognates but also morphological and phonological traits from 24 Indo-European languages (the most ancient members of their branches) and a character-based compatability algorithm to derive a phylogeny for Indo-European. Their best tree is rooted with Anatolian as the outgroup, on the (not universally accepted) assumption that Anatolian is the most distant branch. Their phylogeny provides evidence on the following controversial issues in Indo-European phylogeny: that Tocharian is the most distant branch after Anatolian, and that Balto-Slavic, Italo-Celtic, and Greco-Armenian are valid subgroups. The position of Albanian is unclear, and Germanic is problematic; Ringe et al. argue that an eastern Indo-European group later came into contact with western Indo-European. Nakhleh et al. (2005) use a model that includes the possibility of borrowing (see Warnow

et al. 2006) and argue that Germanic was likely in contact with Italic, Balto-Slavic, and possibly Greco-Armenian. Rexová et al. (2003) use maximum parsimony on the 200-word lists for 84 Indo-European languages from Dyen et al. (1992). They conclude that Germanic forms a clade (valid subgroup) with Italic, possibly Celtic, and very possibly Albanian and that there is an eastern satem group consisting of Balto-Slavic and Indo-Iranian. Gray & Atkinson (2003) used Bayesian methods allowing for unequal rates of change (inter alia) to construct a phylogeny and attribute a date and location for proto-Indo-European, specifically Anatolia around 9000 years ago (see also Atkinson & Gray 2006). Their tree treats Tocharian as the most distant branch after Anatolian and provides strong evidence for Balto-Slavic and weaker evidence for Germanic-Italic-Celtic. Atkinson et al. (2005) replicate their result using the data set from Ringe et al. (2002) (see above) and a stochastic-Dollo model of vocabulary evolution, which include the possibility of borrowing (see Nicholls & Gray 2006). Pagel & Meade (2005) use a Bayesian method on a smaller sample of Indo-European languages and use the tree to argue that the ancestral culture likely had monogamy and a dowry system, and shifts away from this system were first to polygyny and then to the absence of a dowry or presence of a bride-price.

Holden et al. (2005) compare a Bayesian phylogeny reconstruction of Bantu with Holden's (2002) maximum parsimony analysis. Both analyses agree that the northwest Bantu groups are most divergent, with the Bayesian analysis suggesting they are paraphyletic (i.e., they do not together form a valid taxon). The east, southeast, and southwest languages form a single clade, within which there is a clear East Bantu group. Holden et al. argue that the migration of the Bantu peoples follows the spread of farming into southern Africa. Rexová et al. (2006) perform maximum parsimony and Bayesian analyses of the same lexical data in combination with phonological and grammatical characters. They also find a single clade for the east, southeast, and southwest groups and

conclude that there was a single migration from the equatorial rainforests to the areas south and east. Holden & Gray (2006) use network models to ascertain why the Bantu group is not treelike in certain respects. They conclude that West Bantu languages radiated rapidly but without much contact, whereas East Bantu and east Central Bantu languages appear to have had much contact leading to borrowing. Holden & Mace (2003, 2005) use the maximum parsimony tree to argue that matrilineal descent was lost as cattle were adopted in prehistoric southern Africa.

Gray & Jordan (2000) use a parsimony analysis on 77 Austronesian languages with 5185 lexical items from Blust's unpublished *Austronesian Comparative Dictionary* to argue that the structure of the language family is strongly treelike, with greatest diversity in Taiwan, and that this analysis supports the express train hypothesis of relatively rapid colonization of Oceania by the ancestral Austronesians. However, Greenhill & Gray (2005) show that applying bootstrapping to the original study demonstrates it is not that robust. They use Bayesian methods to construct a tree that fits more closely with the traditional historical linguistic analysis, which still places the origin in Taiwan but does not necessarily entail a rapid expansion across Oceania. A network analysis indicates that the major groups in Austronesian are treelike but that the deeper branchings are less treelike because of the lack of signal in the data set.

Dunn et al. (2005) use a maximum parsimony analysis of typological traits, rather than lexical traits, to construct phylogenies of Oceanic and Papuan languages of Island Melanesia. They compare their analysis to the phylogeny of the Oceanic languages in that area and find a high degree of corroboration. The Papuan languages in the area do not display much lexical resemblance, but Dunn et al. argue that the use of typological traits reveals a phylogeny that is apparently lost owing to lexical replacement. Typological traits are avoided in comparative linguistics because they have few possible values (and are thus highly prone to chance resemblance), they diffuse through contact relatively easily, their values are often not independent (e.g., they are linked by implicational universals), and their values are often externally (functionally) motivated (another source of convergence). Some of these issues are raised in a critique by Donohue & Wichmann (2007); see also the response by Dunn et al. (2007). Although the result from Dunn et al. (2005) is surprising to a historical linguist, it may be that a cluster of typological traits will provide more precision in classification than will individual traits; also some typological traits are quite stable and therefore may be useful indicators of phylogeny. However, Gray (2005) notes some weaknesses in the analysis from Dunn et al., and the use of typological traits in phylogeny reconstruction remains to be investigated further.

As with the employment of qualitative concepts from evolutionary biology in theories of language change, the application of quantitative methods from evolutionary biology to phylogeny reconstruction in comparative linguistics is in its infancy. The application to accepted language families using established cognates provides a new perspective on outstanding problems in those families and allows for a link to human prehistory. However, further progress in adapting methods to linguistic phenomena is required before they can be used confidently to investigate controversial or as yet undiscovered language families.

DISCLOSURE STATEMENT

The author is not aware of any biases that might be perceived as affecting the objectivity of this review.

LITERATURE CITED

Atkinson QD, Gray RD. 2005. Curious parallels and curious connection: phylogenetic thinking in biology and historical linguistics. *Syst. Biol.* 54:513–26

Atkinson QD, Gray RD. 2006. How old is the Indo-European family? Illumination or more moths to the flame? See Forster & Renfrew 2006, pp. 91–109

Atkinson QD, Nicholls GK, Welch D, Gray RD. 2005. From words to dates: water into wine, mathemagic or phylogenetic inference? *Trans. Philol. Soc.* 103:193–219

Bastin Y, Coupez A, Mann M. 1999. Continuity and divergence in the Bantu languages: perspectives from a lexicostatistic study. *Ann. Sci. Hum.* 162:315–17

Baxter GJ, Blythe RA, Croft W, McKane AJ. 2006. Utterance selection model of linguistic change. *Phys. Rev. E* 73.046118

Baxter GJ, Blythe RA, Croft W, McKane AJ. 2008. Modeling language change: an evaluation of Trudgill's theory of the emergence of New Zealand English. *Lang. Var. Change.* In press

Baxter WH. 1995. "A stronger affinity…than could have been produced by accident": a probabilistic comparison of Old Chinese and Tibeto-Burman. In *The Ancestry of the Chinese Language*, J. Chin. Ling. Monogr. Ser. 8, ed. WSY Wang, pp. 1–39. Berkeley: Univ. Calif. Press

Baxter WH, Manaster Ramer A. 1996. Review of Donald A. Ringe, Jr., *On Calculating the Factor of Chance in Language Comparison*. *Diachronica* 13:371–84

Baxter WH, Manaster Ramer A. 2000. Beyond lumping and splitting: probabilistic issues in historical linguistics. See Renfrew et al. 2000, pp. 167–88

Ben Hamed M. 2005. Neighbour-nets portray the Chinese dialect continuum and the linguistic legacy of China's demic history. *Proc. R. Soc. London B* 272:1015–22

Ben Hamed M, Wang F. 2006. Stuck in the forest: trees, networks and Chinese dialects. *Diachronica* 23:29–60

Bergsland K, Vogt H. 1962. On the validity of glottochronology. *Curr. Anthropol.* 3:115–53

Blackmore S. 1999. *The Meme Machine*. Oxford, UK: Oxford Univ. Press

Blevins J. 2004. *Evolutionary Phonology: The Emergence of Sound Patterns*. Cambridge, MA: Cambridge Univ. Press

Blust R. 2000. Why lexicostatistics doesn't work: the "universal constant" hypothesis and the Austronesian languages. See Renfrew et al. 2000, pp. 311–31

Blythe RA, McKane AJ. 2007. Stochastic models of evolution in genetics, ecology and linguistics. *J. Stat. Mech.* 07-P07018

Boyd R, Richerson PJ. 1985. *Culture and the Evolutionary Process*. Chicago, IL: Univ. Chicago Press

Bryant D. 2006. Radiation and network breaking in Polynesian linguistics. See Forster & Renfrew 2006, pp. 111–18

Bryant D, Filimon F, Gray RD. 2005. Untangling our past: languages, trees, splits and networks. See Mace et al. 2005, pp. 67–83

Bybee JL. 2001. *Phonology and Language Use*. Cambridge, UK: Cambridge Univ. Press

Cann RL, Stoneking M, Wilson AC. 1987. Mitochondrial DNA and human evolution. *Nature* 329:111–12

Cavalli-Sforza LL, Feldman MW. 1981. *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton, NJ: Princeton Univ. Press

Cavalli-Sforza LL, Piazza A, Menozzi P, Mountain J. 1988. Reconstruction of human evolution: bringing together genetic, archaeological, and linguistic data. *Proc. Natl. Acad. Sci. USA* 85:6002–6

Covington MA. 1996. An algorithm to align words for historical comparison. *Comp. Ling.* 22:481–96

Covington MA. 1998. Alignment of multiple languages for historical comparison. *Proc. COLING-ACL'98: Annu. Meet Assoc. Comput. Ling., 36th, Montreal, Canada, and Int. Conf. Comput. Ling., 17th, Montreal, Canada*, pp. 275–80. Morristown: Assoc. Comput. Ling.

Croft W. 2000. *Explaining Language Change: An Evolutionary Approach*. Harlow, Essex, UK: Longman

Croft W. 2002. The Darwinization of linguistics. *Selection* 3:75–91

Croft W. 2003. Mixed languages and acts of identity: an evolutionary approach. In *The Mixed Language Debate*, ed. Y Matras, P Bakker, pp. 41–72. Berlin: Mouton de Gruyter

Croft W. 2006. The relevance of an evolutionary model to historical linguistics. In *Competing Models of Linguistic Change: Evolution and Beyond*, ed. O Nedergård Thomsen, pp. 91–132. Amsterdam: John Benjamins

Croft W. 2008. The origins of grammaticalization in the verbalization of experience. *Linguistics*. In press

Dawkins R. 1982. *The Extended Phenotype*. Oxford, UK: Freeman

Dawkins R. 1989. *The Selfish Gene*. New York: Oxford Univ. Press. 2nd ed.

Donohue M, Wichmann S. 2007. Typology and the linguistic macrohistory of Island Melanesia. *Ocean. Linguist.* 46:348–87

Dunn M, Foley RA, Levinson SC, Reesink G, Terrill A. 2007. Statistical reasoning in the evaluation of typological diversity in Island Melanesia. *Ocean. Linguist.* 46:388–403

Dunn M, Terrill A, Reesink G, Foley RA, Levinson SC. 2005. Structural phylogenetics and the reconstruction of ancient language history. *Science* 309:2072–75

Durham WH. 1991. *Coevolution: Genes, Culture and Human Diversity*. Stanford, CA: Stanford Univ. Press

Durie M, Ross M, eds. 1996. *The Comparative Method Reviewed: Regularity and Irregularity in Language Change*. Oxford, UK: Oxford Univ. Press

Dyen I, Kruskal JB, Black P. 1992. An Indoeuropean classification: a lexicostatistical experiment. *Trans. Am. Philos. Soc.* 82:1–132

Forster P, Renfrew C, eds. 2006. *Phylogenetic Methods and the Prehistory of Languages*. Cambridge, UK: McDonald Inst. Archaeol. Res.

Gray RD. 2005. Pushing the time barrier in the quest for language roots. *Science* 309:2007–8

Gray RD, Atkinson QD. 2003. Language-tree divergence times support the Anatolian theory of Indo-European origin. *Nature* 426:435–39

Gray RD, Jordan FM. 2000. Language trees support the express-train sequence of Austronesian expansion. *Nature* 405:1052–55

Greenberg JH. 1957. Language, diffusion and migration. In *Essays in Linguistics*, pp. 66–74. Chicago, IL: Univ. Chicago Press

Greenberg JH. 1992. Preliminaries to a systematic comparison between biological and linguistic evolution. In *The Evolution of Human Languages*, ed. JA Hawkins, M Gell-Mann, pp. 139–58. Redwood City, CA: Addison Wesley

Greenberg JH. 1993. Observations concerning Ringe's *Calculating the Factor of Chance in Language Comparison*. *Proc. Am. Philos. Soc.* 137:79–90

Greenberg JH. 2001. The methods and purposes of linguistic genetic classification. *Lang. Linguist.* 2:111–36

Greenberg JH. 2005. *Genetic Linguistics: Theory and Method*, ed. W Croft. Oxford, UK: Oxford Univ. Press

Greenberg JH, Ruhlen M. 1992. Linguistic origins of Native Americans. *Sci. Am.* 267(5):94–99

Greenhill SJ, Gray RD. 2005. Testing population dispersal hypotheses: Pacific settlement, phylogenetic trees and Austronesian languages. See Mace et al. 2005, pp. 31–52

Hinnebusch TJ. 1999. Contact and lexicostatics in comparative Bantu studies. In *Bantu Historical Linguistics: Theoretical and Empirical Perspectives*, ed. J-M Hombert, LM Hyman, pp. 173–205. Stanford, CA: Cent. Study Lang. Inf.

Holden CJ. 2002. Bantu language trees reflect the spread of farming across sub-Saharan Africa: a maximum-parsimony analysis. *Proc. R. Soc. London Ser. B* 269:793–99

Holden CJ, Gray RD. 2006. Rapid radiation, borrowing and dialect continua in the Bantu languages. See Forster & Renfrew 2006, pp. 19–31

Holden CJ, Mace R. 2003. Spread of cattle led to the loss of matrilineal descent in Africa: a coevolutionary analysis. *Proc. R. Soc. London Ser. B* 270:2425–33

Holden CJ, Mace R. 2005. "The cow is the enemy of matriliny": using phylogenetic methods to investigate cultural evolution in Africa. See Mace et al. 2005, pp. 217–34

Holden CJ, Meade A, Pagel M. 2005. Comparison of maximum parsimony and Bayesian Bantu language trees. See Mace et al. 2005, pp. 53–65

Hull DL. 1988. *Science as a Process: An Evolutionary Account of the Social and Conceptual Development of Science*. Chicago, IL: Univ. Chicago Press

Hull DL. 2001. *Science and Selection: Essays on Biological Evolution and the Philosophy of Science*. Cambridge, UK: Cambridge Univ. Press

Joos M. 1964. Glottochronology with retention rate homogeneity. *Proc. Int. Congr. Linguist.*, *9th*, *Cambridge, Mass.*, ed. HG Lunt, p. 237. The Hague: Mouton

Justeson JS, Stephens LD. 1980. Change cognation: a probabilistic model and decision procedure for historical inference. *Pap. Int. Conf. Hist. Linguist.*, *4th*, *Stanford, Calif.*, ed. EC Traugott, R LaBrum, S Shepherd, pp. 37–46. Amsterdam: John Benjamins

Kessler B. 1995. Computational dialectology in Irish Gaelic. *Proc. EACL-95: Conf. Eur. Chapter Assoc. Comput. Linguist.*, *6th*, *Dublin, Ireland*, pp. 60–67. San Francisco: Morgan Kaufman

Kessler B. 2001. *The Significance of Word Lists*. Stanford, CA: Cent. Study Lang. Inf.

Kondrak G. 2002. *Algorithms for language reconstruction*. PhD thesis, Univ. Tor. **http://www.cs.ualberta. ca/∼kondrak**

Kondrak G. 2003. Phonetic alignment and similarity. *Comput. Hum.* 37:273–91

Kruskal JB, Dyen I, Black P. 1973. Some results from the vocabulary method of reconstructing language trees. In *Lexicostatistics in Genetic Linguistics*, ed. I Dyen, pp. 30–55. The Hague: Mouton

Lass R. 1990. How to do things with junk: exaptation in language change. *J. Linguist.* 26:79–102

Lass R. 1997. *Historical Linguistics and Language Change*. Cambridge, UK: Cambridge Univ. Press

Lightfoot DW. 2006. *How Languages Emerge*. Cambridge, UK: Cambridge Univ. Press

Mace R, Holden CJ, Shennan S, eds. 2005. *The Evolution of Cultural Diversity: A Phylogenetic Approach*. London: UCL Press

Manaster Ramer A, Hitchcock C. 1996. Glass houses: Greenberg, Ringe, and the mathematics of comparative linguistics. *Anthropol. Linguist.* 38:601–19

McMahon AMS, Heggarty P, McMahon R, Slaska N. 2005. Swadesh sublists and the benefits of borrowing: an Andean case study. *Trans. Philol. Soc.* 103:147–70

McMahon AMS, McMahon R. 2003. Finding families: quantitative methods in language classification. *Trans. Philol. Soc.* 101:7–55

McMahon AMS, McMahon R. 2006. *Language Classification by Numbers*. Oxford, UK: Oxford Univ. Press

Milroy L. 1987. *Language and Social Networks*. Oxford: Basil Blackwell. 2nd ed.

Minett JW, Wang WS-Y. 2003. On detecting borrowing: distance-based and character-based approaches. *Diachronica* 20:289–330

Mufwene S. 2001. *The Ecology of Language Evolution*. Cambridge, UK: Cambridge Univ. Press

Mufwene S. 2005. *Créole, Écologie Sociale, Évolution Linguistique*. Paris: L'Harmattan

Nakhleh L, Ringe DJ Jr, Warnow T. 2005. Perfect phylogenetic networks: a new methodology for reconstructing the evolutionary history of natural languages. *Language* 81:382–420

Nerbonne J, Heeringa W. 1997. Measuring dialect distance phonetically. *Proc. SIGPHON-97: Meet. ACL Spec. Interest Group Comput. Phonol.*, *3rd*, *Madrid, Spain*. **http://www.cogsci.ed.ac.uk/sigphon**

Nettle D. 1999. *Linguistic Diversity*. Oxford, UK: Oxford Univ. Press

Nicholls GK, Gray RD. 2006. Quantifying uncertainty in a stochastic model of vocabulary evolution. See Forster & Renfrew 2006, pp. 161–71

Nichols J. 1996. The comparative method as heuristic. See Durie & Ross 1996, pp. 39–89

Oakes MP. 2000. Computer estimation of vocabulary in protolanguage from word lists in four daughter languages. *J. Quant. Linguist.* 7:233–43

Ohala JJ. 2003. Phonetics and historical phonology. In *Handbook of Historical Linguistics*, ed. B Joseph, R Janda, pp. 669–86. Oxford, UK: Blackwell

Oswalt RL. 1970. The detection of remote linguistic relationships. *Comp. Stud. Hum. Verbal Behav.* 3:117–29

Oswalt RL. 1991. A method for assessing distant linguistic relationships. In *Sprung from Some Common Source: Investigations into the Prehistory of Languages*, ed. SM Lamb, ED Mitchell, pp. 389–404. Stanford, CA: Stanford Univ. Press

Pagel M. 2000. Maximum-likelihood models for glottochronology and for reconstructing linguistic phylogenies. See Renfrew et al. 2000, pp. 189–207

Pagel M, Atkinson QD, Meade A. 2007. Frequency of word-use predicts rates of lexical evolution throughout Indo-European history. *Nature* 449:717–20

Pagel M, Meade A. 2005. Bayesian estimation of correlated evolution across cultures: a case study of marriage systems and wealth transfer at marriage. See Mace et al. 2005, pp. 235–56

Pagel M, Meade A. 2006. Estimating rates of lexical replacement on phylogenetic trees of languages. See Forster & Renfrew 2006, pp. 173–82

Pierrehumbert JB. 2001. Exemplar dynamics: word frequency, lenition and contrast. In *Frequency and Emergence in Grammar*, ed. JL Bybee, P Hopper, pp. 137–57. Amsterdam: John Benjamins

Renfrew C, McMahon AMS, Trask RL, eds. 2000. *Time Depth in Historical Linguistics*. Cambridge, UK: McDonald Inst. Archaeol. Res.

Rexová K, Bastin Y, Frynta D. 2006. Cladistic analysis of Bantu languages: a new tree based on combined lexical and grammatical data. *Naturwissenschaften* 93:189–94

Rexová K, Frynta D, Zrzavy J. 2003. Cladistic analysis of languages: Indo-European classification based on lexicostatistical data. *Cladistics* 19:120–27

Richerson PJ, Boyd R. 2005. *Not By Genes Alone: How Culture Transformed Human Evolution*. Chicago, IL: Univ. Chicago Press

Ringe DA Jr, Warnow T, Taylor A. 2002. Indo-European and computational cladistics. *Trans. Philol. Soc.* 100:59–129

Ringe DA Jr. 1992. *On Calculating the Factor of Chance in Language Comparison*. Trans. Am. Philos. Soc. Vol. 82, Part 1. Philadelphia: Am. Philos. Soc.

Ringe DA Jr. 1993. A reply to Professor Greenberg. *Proc. Am. Philos. Soc.* 137:91–109

Ritt N. 2004. *Selfish Sounds and Linguistic Evolution: A Darwinian Approach to Language Change*. Cambridge, UK: Cambridge Univ. Press

Ross M, Durie M. 1996. Introduction. See Durie & Ross 1996, pp. 3–38

Vigilant L, Stoneking M, Harpending H, Hawkes R, Wilson AC. 1991. African populations and the evolution of human mitochondrial DNA. *Science* 253:1503–7

Wang WS-Y, Minett JW. 2005. Vertical and horizontal transmission in language evolution. *Trans. Philol. Soc.* 103:121–46

Warnow T, Evans SN, Ringe DA Jr, Nakhleh L. 2006. A stochastic model of language evolution that incorporates homoplasy and borrowing. See Forster & Renfrew 2006, pp. 75–87

Wedel A. 2006. Exemplar models, evolution and language change. *Linguist. Rev.* 23:247–74

Contents