
Biological Adaptations for Functional Features of Language in the Face of Cultural Evolution

MORTEN H. CHRISTIANSEN,^{1,2} FLORENCIA REALI,³ AND NICK CHATER^{4*}

Abstract Although there may be no true language universals, it is nonetheless possible to discern several family resemblance patterns across the languages of the world. Recent work on the cultural evolution of language indicates the source of these patterns is unlikely to be an innate universal grammar evolved through biological adaptations for arbitrary linguistic features. Instead, it has been suggested that the patterns of resemblance emerge because language has been shaped by the brain, with individual languages representing different but partially overlapping solutions to the same set of nonlinguistic constraints. Here, we use computational simulations to investigate whether biological adaptation for functional features of language, deriving from cognitive and communicative constraints, may nonetheless be possible alongside rapid cultural evolution. Specifically, we focus on the Baldwin effect as an evolutionary mechanism by which previously learned linguistic features might become innate through natural selection across many generations of language users. The results indicate that cultural evolution of language does not necessarily prevent functional features of language from becoming genetically fixed, thus potentially providing a particularly informative source of constraints on cross-linguistic resemblance patterns.

Although the world's languages differ considerably from one another, there are statistical "family resemblance" patterns in how languages are structured and used. Despite potential similarities between any given pair of languages, there is no single set of features common to all languages (for a review, see Evans and Levinson 2009). For example, even when two languages have similar sets of word classes, these languages will nonetheless differ substantially in their morphosyntactic properties, semantic properties, or both (Haspelmath 2007). Determining whether such patterns are the product of biological or cultural evolution is the focus of much debate in language evolution research.

¹Department of Psychology, Uris Hall, Cornell University, Ithaca, NY 14853.

²Santa Fe Institute, 1399 Hyde Park Rd., Santa Fe, NM 87501.

³Institute of Cognitive and Brain Sciences, Tolman Hall, University of California, Berkeley, CA 94720.

⁴Behavioural Science Group, Warwick Business School, University of Warwick, Coventry, CV4 7AL, U.K.

Correspondence to: Nick Chater, Behavioural Science Group, Warwick Business School, University of Warwick, Coventry, CV4 7AL, United Kingdom. E-mail: nick.chater@warwick.ac.uk.

Human Biology, April 2011, v. 83, no. 2, pp. 247–259.

Copyright © 2011 Wayne State University Press, Detroit, Michigan 48201-1309

KEY WORDS: LANGUAGE EVOLUTION, BALDWIN EFFECT, LANGUAGE UNIVERSALS, UNIVERSAL GRAMMAR, COEVOLUTION.

One view sees these cross-linguistic patterns as deriving from an innate Universal Grammar (UG; Chomsky 1965), comprising several domain-specific linguistic properties, including case marking; agreement; and conformity to highly abstract syntactic constraints, such as X-bar theory (Pinker and Jackendoff 2005). Crucially, these features of language are assumed to be arbitrary in that they defy functional explanation; that is, they cannot be explained in terms of constraints on learning, processing, memory, or communication. It has even been suggested that the arbitrary features of UG may actually hinder, rather than facilitate, communication (Chomsky 2005). This abstract body of linguistic constraints is proposed, by some theorists, to have evolved gradually through biological adaptations for increasingly complex grammars (e.g., Briscoe 2003; Jackendoff 2002; Nowak et al. 2002; Pinker 2003; Pinker and Bloom 1990). An alternative view seeks to explain cross-linguistic family resemblance patterns in terms of functional features of language, emerging due to general communicative and cognitive factors that are not language-specific (e.g., Bybee 1998). These features are seen as by-products of linguistic adaptation, in which language itself has been adapted through cultural transmission across many generations of language learners (e.g., Christiansen and Chater 2008; Deacon 1997; Kirby and Hurford 2002; Tomasello 2003).

The Baldwin effect (1896) is the primary evolutionary mechanism by which the arbitrary features of UG are envisioned to have been genetically fixed in the human population (e.g., Pinker and Bloom 1990). Although a Darwinian mechanism, the Baldwin effect resembles Lamarckian inheritance of acquired characteristics, in that traits that are learned or developed over the life span of an individual become gradually encoded in the genome over many generations (see Weber and Depew 2003). That is, if a trait increases fitness, then individuals that, due to random genetic variation, require less exposure to the environment to develop that trait will have a selective advantage. Over generations, the amount of environmental exposure needed to develop this trait will decrease, as individuals evolve increasingly better initial conditions for its rapid development. Eventually, no environmental exposure may be needed; the trait has become genetically encoded. A frequently cited example of the Baldwin effect (e.g., Briscoe 2003) is the ability to develop hard skin on certain areas of the body with relatively little environmental exposure. Over time, natural selection would have favored individuals that could develop hard skin more rapidly (e.g., because it aids in mobility or prevents infection) until it became fixed in the genome, requiring little environmental stimulation to develop. Similarly, it has been suggested that arbitrary linguistic features, which would originally have had to be learned as cultural conventions, gradually became genetically fixed in UG, via the Baldwin effect (Pinker and Bloom 1990).

In previous work (Chater et al. 2009), we have shown that there are strong restrictions on the conditions under which the Baldwin effect can operate and genetically embed arbitrary linguistic constraints. Specifically, the Baldwin effect only emerges when language provides a stable target for natural selection,

presupposing that there is no linguistic change. However, given that there is nothing to stabilize the arbitrary features, they are likely to drift over time. Indeed, historical linguistics shows that linguistic conventions, like other cultural products, change much faster than biological evolution (Gray and Atkinson 2003). Thus, whereas linguistic changes occur over centuries or millennia, biological evolution operates over hundreds of thousands of years. Chater et al. (2009) also show that allowing genes to control language learnability and hence to influence the direction of linguistic change in a process similar to Baldwinian niche construction (e.g., Odling-Smee et al. 2003), does not lead to the operation of the Baldwin effect. Indeed, the degree of genetic pressure needed to overcome the effect of linguistic change and drive the genes to assimilate to the language was so high that it would effectively predetermine the nature of language, such that the population would converge to the language even without selection, as a function of genetic drift.

Together with related computational work (e.g., Kirby and Hurford 1997; Munroe and Cangelosi 2002; Reali and Christiansen 2009; Yamauchi 2001), these results cast doubt on the possibility of biological adaptations for arbitrary properties of language. It remains to be explained what gives rise to the patterns of family resemblance observable across the languages of the world. Recent modeling has begun to provide plausible explanations for some of these patterns in terms of linguistic adaptations to preexisting learning biases. For example, Reali and Christiansen (2009) showed how sequential learning biases, amplified by cultural transmission across generations of neural network learners, can result in the emergence of consistent word order (for a similar, rule-based model, see Kirby 1999). In a more general analysis, Kirby et al. (2007) have shown that the language of a cross-generational population of Bayesian learners, sampling from a prior distribution of grammars, will converge to match the priors encoded by the learners. Other modeling work has demonstrated how interactions between language users within a population also can result in linguistic adaptation. Along these lines, it has been demonstrated that words and perceptually grounded categories may coevolve with one another (Steels and Belpaeme 2005), leading to the emergence of a hierarchical system in which linguistic categories can overcome individual differences in perceptual discrimination (Puglisi et al. 2008).

Most of the previous modeling work that has highlighted the potential contribution of cultural evolution to explanations of linguistic structure has either focused on vertical or horizontal transmission of language (Jäger et al. 2009). Thus, whereas some models have focused on linguistic adaptation to the limitations arising from vertical transmission of language *across* generations of learners (e.g., Chater et al. 2009; Kirby et al. 2007), others have centered on the emergence of structure from horizontal transmission of information *within* a generation of learners (e.g., Puglisi et al. 2008; Steels and Belpaeme 2005). Often coinciding with the vertical and horizontal emphasis is the issue of whether there is cross-generational learning of a linguistic system between

single “teachers” and learners versus communicative interaction between multiple agents. Crucially, both vertical and horizontal models have suggested that the role of biological adaptation for language may be minimal, in the light of cultural evolution (for discussion, see Jäger et al. 2009). Here, we investigate the possibility of biological adaptations for *functional* features of language in simulations involving both horizontal and vertical transmission of linguistic information in a population of agents with communicative and cognitive biases. Thus, we seek to model functional aspects of language—potentially including vocabulary size, emphasis on local linguistic processes, layered digital codes of phonological and syntactic structure, and compositional semantics—that may be stable aspects of the linguistic environment, precisely because of their functional role in subserving effective communication. The results indicate that such functional features of language can become genetically fixed in the population when they promote better communicative abilities.

Adaptation for Functional Features of Language

Our simulations seek to explore whether the Baldwin effect might occur for functional aspects of language, that is, aspects of language that facilitate communication. Thus, these simulations go beyond much previous work investigating the potential role of the Baldwin effect in language evolution (e.g., Briscoe 2003; Chater et al. 2009; Hinton and Nowlan 1987; Kirby and Hurford 1997; Munroe and Cangelosi 2002; Reali and Christiansen 2009) in that communication exchanges between agents are key to both learning and reproductive fitness. In addition, the simulations incorporate (1) socio-pragmatic constraints requiring some amount of parity between agents for learning and communication to take place, (2) cognitive constraints in the form of a comprehension/production asymmetry that allows comprehension to precede production to some degree, and (3) a communicative bias only to learn from more competent communicators. Importantly, the simulations have no distinction between communication and learning insofar as every communicative event also provides a learning opportunity (in contrast to much previous work, e.g., Chater et al. 2009; Kirby and Hurford 1997).

Methods

Setup of Simulations. The simulations aim to model potential gene-language coevolution in a small band of hominids, focusing on the emergence of potential adaptations for better communicative abilities due to functional pressures. An individual’s language is defined in terms of a set of n communicative features, $F_1 \dots F_n$, taking the values $+_F$ or $-_F$. The corresponding genes, $G_1 \dots G_n$, each have three alleles, $+_G$, $-_G$, and $?_G$, with the

first two alleles encoding biases toward $+_F$ or $-_F$ and the final allele being “neutral.”⁵ The number of genes and communicative features, n , was set to 20. The strength of the genetic bias of $+_G$ and $-_G$ for sampling the corresponding $+_F$ and $-_F$ was 95% (i.e., for $+_G$ there is a 95% chance of getting an $+_F$). For the neutral allele, $?_G$, there is a 50% chance of getting either a $+_F$ or $-_F$. The simulations start with x neutral alleles, where x in the current simulations is 0%, 50%, or 100%; the remaining biased alleles are set randomly to $+_G$ or $-_G$. Group size was set to 100, which is within the hypothesized range of hominid band sizes (Dunbar and Shultz 2007). For each generation, pairs of agents were randomly picked for 500 interactions. The 50 agents with the highest communication scores were selected for reproduction through sexual recombination of the genomes of two randomly chosen “parents.” The alleles at each “child” locus were randomly sampled from the same locus on either of the two parent genomes. The child genome was subject to random mutation at the rate of m , which in the current study varied between 1%, 0.5%, and 0.1%. The effect of mutation was to reassign the value $+_G$, $-_G$, or $?_G$ with equal probability.

Communicative Interaction. Communication is only possible between agents who have a majority of the same kinds on communicative features (either $+_F$ or $-_F$). Only the number of $+_F$ and $-_F$ features matters for the purpose of communicative success, and not their order with respect to one another.⁶ Thus, communication requires some degree of sign parity between the interacting individuals; they must be able to understand each other, at least in part. That is, an agent, a_1 , whose language [$+_F$, $+_F$, $+_F$, $+_F$, $-_F$] has four $+_F$ ’s would be able to communicate with an agent, a_2 , with a [$+_F$, $+_F$, $+_F$, $+_F$, $+_F$] language containing five $+_F$ ’s but not with agent a_3 that has a [$+_F$, $-_F$, $-_F$, $-_F$, $+_F$] language with only two $+_F$ ’s. Agents benefit mutually from successful communication in proportion to the overlap in their features. The successful *bidirectional* interaction between a_1 and a_2 would result in an increase in both agents’ communication scores by 9 (the combined number of $+_F$ ’s in their two languages). Given that fitness is depending on an agent’s communication score, there is a selective pressure toward having genes encoding the language features of the same kind, the sign of which will depend on what is most prevalent in the population. This can be seen as a general functional pressure toward more expressivity in the language.

⁵Because Chater et al. (2009) had previously shown that the Baldwin effect does not work for arbitrary linguistic features, we adopted their simplified one-to-one mapping between genes and language to demonstrate that functional pressures nonetheless *can* lead to genetic assimilation under similar simulation conditions. Note that these genetic features signify strong biases, rather than rigid binary constraints as in, for example, Kirby and Hurford (1997).

⁶This pressure toward feature sign parity loosens the one-to-one mapping between genes and language by analogy with protein-protein networks. Thus, although a gene may be selected for its associated protein, selection can also work at the level of protein networks (e.g., Kim et al. 2007).

Production/Comprehension Asymmetry. In language acquisition, there is a general tendency for comprehension to precede production. As noted by Clark (1993: 246), “Logically, comprehension must precede production. How else can speakers know which words to use to convey a particular meaning?” (for a recent discussion, see Hendriks and Koster 2010). The simulations therefore incorporate a production/comprehension asymmetry that provides a cognitively motivated way for agents to learn from more language proficient individuals: even though a_1 can only “produce” four $+_F$ ’s, this agent can “comprehend” a_2 ’s five $+_F$ ’s. However, if the difference between the productive abilities of two agents is more than one unit, then the less competent “speaker” will not be able to understand its more proficient communication partner, resulting in a *unidirectional* interaction. In this case, the proficient speaker receives the combined communication score (as before because this speaker can understand both parts of the interaction), whereas the less competent agent only receives its own contribution to that score. Hence, if a_2 interacted with a_4 , whose language is $[+_F, -_F, +_F, -_F, +_F]$, a_2 would increase its communication score by 8, whereas a_4 ’s score would only increase by 3.

Differences in Linguistic Ability. Learning is mediated by communicative interactions, in which less competent agents are able to learn from more competent agents (with stronger bias toward $+_F$ ’s or $-_F$ ’s); this is meant to reflect the tendency for children to learn much of their language from others with greater language skills than themselves (e.g., adults or older children). Learning can only happen when bidirectional communication is possible (as described above), allowing the less competent agent, based on the biases in its genome, to resample the first bit in its language that differs from the more competent agent’s language. For example, in a communicative interaction between a_1 and a_4 , the latter would resample its second language bit. If a_4 ’s genome encoded an innate bias ($+_G$ or $-_G$), then there would be a 95% chance of getting this bit expressed; but if the genome encoded the neutral allele, $?_G$, the chance of either value would be 50%. Thus, genes provide potentially strong biases on learning, determining how easy it will be to change the communicative feature for a given locus.

Language Change. Our previous work has shown how genetic assimilation of linguistic features can be strongly affected by language change (Chater et al. 2009). To mirror the effect of language change on learning, we introduced noise into the learning process at a rate 10 times higher than the mutation rate. During 10% of the learning opportunities, a random feature in the learner’s language is chosen for potential reassignment (given the learner’s genetic bias for that feature) instead of the first bit that deviated from the competent speaker’s language. This parallels the random change in the target language in our previous work, while also incorporating Baldwinian niche construction (Odling-Smee et al. 2003), whereby the genetic biases of the learners affect the direction of change (although in this case on an individual basis rather

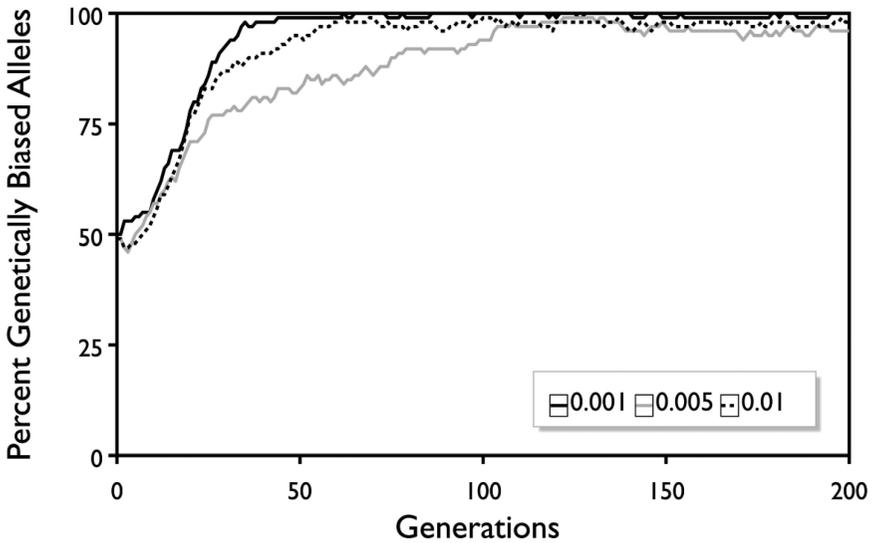


Figure 1. Percentage of genetically biased alleles across generations for different mutation rates, averaged across agent in the population, starting with 50% neutral alleles.

than at the group level as in Chater et al. 2009). That is, the randomly chosen language feature, F_i , will be reassigned based on the learner's genetic bias, G_i , ($+_G$, $-_G$, or $?_G$) independently of the value that this particular feature has in the competent speaker's language. For the different rates of biological mutation, m , the relative speed of linguistic change was always $10m$ (to model the effect of language change from Chater et al. 2009, as closely as possible).

Results

In our simulations, we investigated whether a combination of vertical and horizontal transmission of language, together with functional pressures toward increased communicative ability would lead to biological adaptations for language.⁷ Figure 1 shows, across generations, the percentage of the genome, at the population level, that has a genetic bias (toward either $+_G$ or $-_G$) when the agents start out with 50% neutral alleles ($?_G$). Using 95% as a criterion for successful genetic assimilation, the Baldwin effect emerges robustly across different rates of mutation after 33–144 generations. To determine how this changing genetic bias may affect the overall language used by the population of agents, we computed the mean sign across the 20 language features. The result of this analysis is presented in Figure 2 in terms of the percentage of $+_F$ features across generations, showing that the average phenotype for a population—what might be construed as a common expressivity bias—quickly settles on either

⁷The figures show data from single simulation runs but the results have been replicated across multiple different runs.

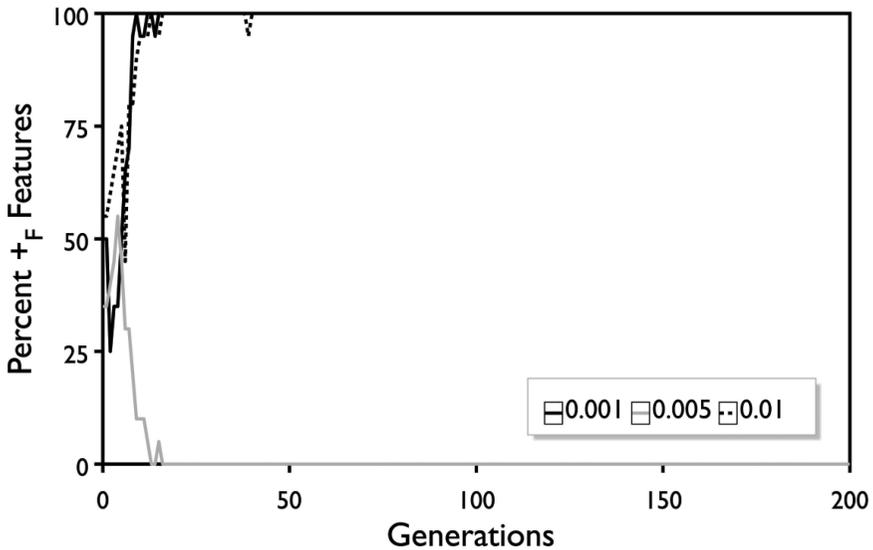


Figure 2. Percentage of $+_F$ features across generations for different mutation rates, averaged across individual agent languages in the population for agents starting with 50% neutral alleles.

having primarily all $+_F$'s or all $-_F$'s. This change in the language occurs within only 16 generations of learners and thus sets the stage for the direction of the genetic bias that follows at a slower pace. However, because the functional pressures pushed the language toward one of the two fitness peaks (all $+_F$'s or all $-_F$'s), the Baldwin effect is able to catch up even in the face of the considerable random linguistic change at the individual agent level.

Figure 3 illustrates that the Baldwin effect also occurs even if the learners start out with all neutral alleles, although the genetic assimilations take considerable longer (note the increase in the number of generations on the ordinate axis). Across different mutation rates, the Baldwin effect emerges between 252 and 585 generations. The protracted period of genetic assimilation is paralleled by a comparable slowing of the linguistic change toward either $+_F$'s or $-_F$'s. Thus, as shown in Figure 4, the language does not settle completely into one of the fitness maxima until after 70–339 generations, although still well ahead of the genetic change.

Our previous simulations involving arbitrary features, which by definition do not have functional pressures to keep them stable and thus become “moving targets” for the Baldwin effect, showed that even when the genome starts out with all genetically biased alleles, neutral alleles will quickly invade the population as they provide for faster learning of a rapid-changing language (Chater et al. 2009). In the current simulations with a starting state involving only genetically biased alleles, we found that neutral alleles never got a foothold in the genome because the genetic biases were functionally adaptive, providing a bias toward better communicative ability. Figure 5 shows that the average language

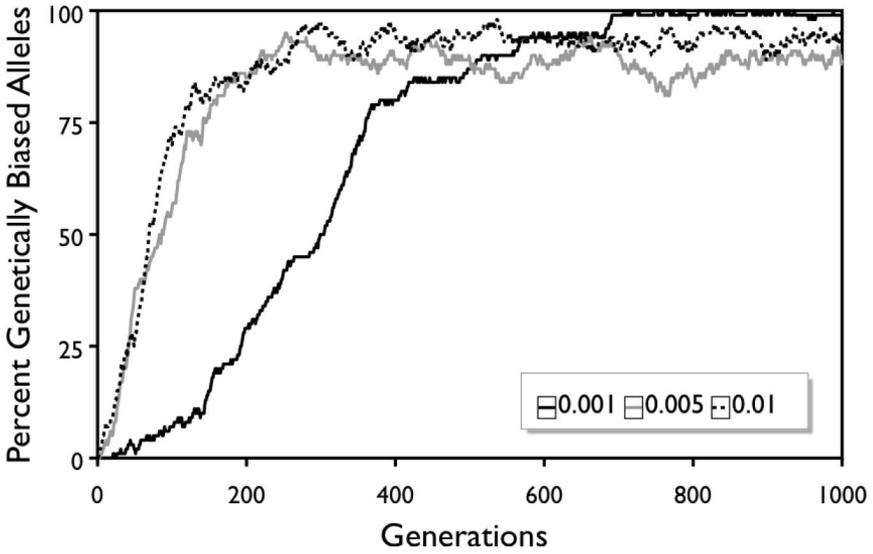


Figure 3. Percentage of genetically biased alleles across generations for different mutation rates, averaged across agent in the population, starting with 100% neutral alleles.

across the population rapidly converges on either all $+_F$'s or all $-_F$'s after only 6 to 10 generations. Thus, the Baldwin effect emerges robustly across variations in both mutation rate and the initial number of neutral alleles.

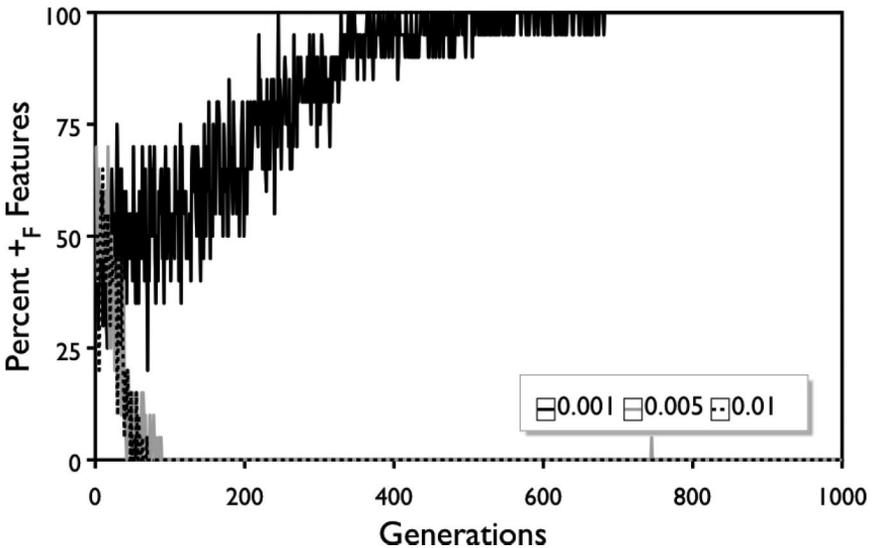


Figure 4. Percentage of $+_F$ features across generations for different mutation rates, averaged across individual agent languages in the population for agents starting with 100% neutral alleles.

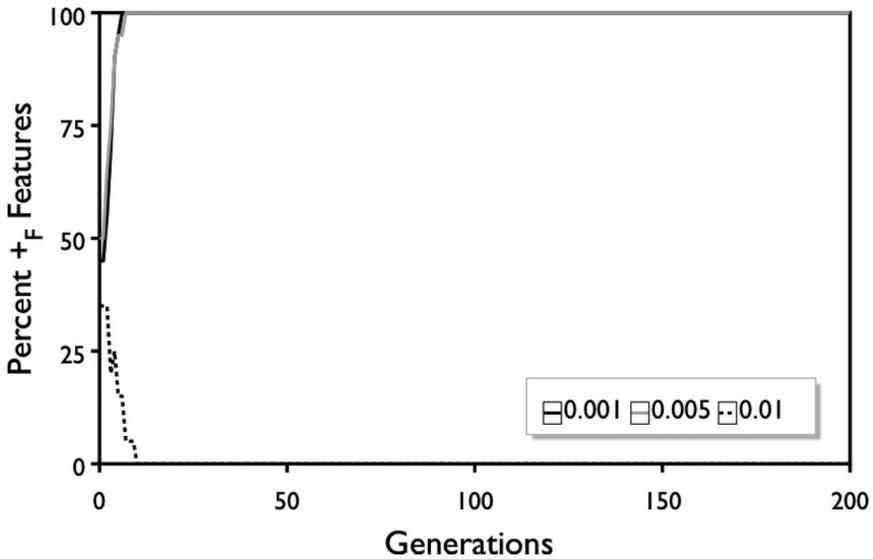


Figure 5. Percent of $+_F$ features across generations for different mutation rates, averaged across individual agent languages in the population for agents starting with 0% neutral alleles.

Discussion

Past modeling work has shown that both vertical (e.g., Kirby et al. 2007; Reali and Christiansen 2009) and horizontal (e.g., Puglisi et al. 2008; Steels and Belpaeme 2005) transmission of information between language users can lead to the emergence of linguistic structure (for a review, see Jäger et al. 2009). In some cases, there are even demonstrations that such cultural evolution may prevent biological adaptations for arbitrary features of language (e.g., Chater et al. 2009; Smith and Kirby 2008). The current simulations extend this prior work by combining both vertical and horizontal transmission of linguistic information and incorporating cognitive and communicative constraints on linguistic interactions. Our results indicate that selective pressures toward increased communicative abilities can lead to biological adaptations for functional features of language even in the context of cultural evolution of language (toward primarily having either $+_F$'s or $-_F$'s). That is, functional features that improve communication can become genetically fixed in the population, despite the possibility of rapid linguistic change driven by cultural transmission.

This work can be viewed as building on prior work relating to the interaction of cultural and biological factors in the evolution of language. In particular, the innovative work of Kirby and Hurford (1997) first considered the potential importance of functional pressures (in their model, parsability) as acting to stabilize the language through cultural evolution. Such a stable language can thus provide a target against which biological evolution can operate to drive the relevant aspects of linguistic structure into the genes through the Baldwin effect.

Although there is considerable overlap in the conclusions drawn from their simulations and ours—despite having very different formal characteristics (e.g., learning operating by a triggering mechanism)—there are nonetheless also important differences in fundamental assumptions. Thus, whereas Kirby and Hurford's model has a strict separation of learning and communication, our simulations assume that learning occurs only in the context of communication. This key difference probably plays a role in the different conclusions drawn from the respective simulations. Kirby and Hurford suggest that complex functional features of language (in their simulations, involving alignment between features) cannot be genetically assimilated. By contrast, our simulations show that a complex functional feature (sameness of feature values across parameters) can be assimilated by the Baldwin effect.

Methodologically, we see our simulations as a first step toward a more realistic paradigm for simulating the evolution of language. Similar to other current computational modeling of language evolution, our simulations nonetheless incorporate several simplifying assumptions. For example, we used a sudden generational turnover every 500 interactions, rather than a gradual replacement of individuals. Neither did we incorporate any population structure into the simulations nor potential mating strategies (e.g., assortative mating; Thiessen and Gregg 1980). Future work is needed to determine the extent to which these simplifications impose limitations on the generality of our results, as well as to explore other outstanding questions, such as the potential interaction between arbitrary and functional features in language evolution.

On a theoretical level, the picture that emerges from the current and previous modeling of the cultural evolution of language is consistent with the account of language evolution put forward in Christiansen and Chater (2008) (for a brief summary, see Christiansen et al. 2009). On this view, an innate universal grammar, consisting of abstract features of language that started out as arbitrary linguistic conventions, can be ruled out on evolutionary grounds. Arbitrary linguistic features are free to change in whatever direction conceivable because there is nothing to keep them stable over time. By definition, each arbitrary convention is as good as the next, and none is better than any other. Thus, arbitrary features are characterized by a flat fitness landscape. In contrast, only a few functional features provide better communicative abilities, whereas most do not and some might even prevent communication altogether. The fitness landscape for functional features is therefore full of valleys and peaks, and the evolutionary hill-climbing process will drive the system toward the peaks (although not necessarily the highest, i.e., the globally optimal, peak). In the simulations presented here, there were two functional fitness peaks that allow for the best communicative ability characterized by genomes with either all $+_G$'s or all $-_G$'s. Depending on the specific genetic make-up of the successful communicators in the initial populations, one of these peaks quickly emerges as the stable functional target for selection by way of their expression in the

language phenotype (Figures 2, 4, and 5). Thus, we can construe the functional pressure as being toward increased expressivity in the communicative interactions between agents.

Our simulations suggest that cultural evolution of language is compatible with the notion that there may have been some functional adaptations for language. However, these biological adaptations do not have to be specific to language. For example, vocabulary learning is likely to rely on innate domain-general abilities for establishing reliable mappings between forms and meanings (e.g., Bloom 2002). As such, the ability to acquire a large vocabulary may have become gradually innate by way of the Baldwin effect because it would have increased communicative abilities. From this perspective, then, the patterns of family resemblance among languages, in terms of their structure and use, may be explained by cultural evolution of language constrained by the brain (Christiansen and Chater 2008). These constraints include socio-pragmatic considerations, the nature of our thought processes, perceptuo-motor factors, and cognitive limitations on learning, memory and processing. Most of these constraints are likely to predate the emergence of language but, as suggested by our results, some may have been subject to subsequent biological adaptation due to functional pressures toward better communicative abilities.

Received 1 April 2010; revision accepted for publication 28 September 2010.

Literature Cited

- Baldwin, J. M. 1896. A new factor in evolution. *Am. Nat.* 30:441–451.
- Bloom, P. 2002. *How Children Learn the Meanings of Words*. New York, NY: Oxford University Press.
- Briscoe, E. 2003. Grammatical assimilation. In *Language Evolution*, M. H. Christiansen and S. Kirby, eds. New York, NY: Oxford University Press, 295–316.
- Bybee, J. 1998. A functionalist approach to grammar and its evolution. *Evol. Commun.* 2:249–278.
- Chater, N., F. Reali, and M. H. Christiansen. 2009. Restrictions on biological adaptation in language evolution. *Proc. Natl. Acad. Sci. USA* 106:1015–1020.
- Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Chomsky, N. 2005. Three factors in language design. *Linguist. Inq.* 36:1–22.
- Christiansen, M. H., and N. Chater. 2008. Language as shaped by the brain. *Behav. Brain Sci.* 31:487–558.
- Christiansen, M. H., N. Chater, and F. Reali. 2009. The biological and cultural foundations of language. *Commun. Integr. Biol.* 2:221–222.
- Clark, E. 1993. *The Lexicon in Acquisition*. Cambridge, U.K.: Cambridge University Press.
- Deacon, T. 1997. *The Symbolic Species: the Coevolution of Language and the Brain*. New York, NY: Norton.
- Dunbar, R. I. M., and S. Shultz. 2007. Evolution in the social brain. *Science* 317:1344–1347.
- Evans, N., and S. Levinson. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behav. Brain Sci.* 32:429–492.
- Gray, R. D., and Q. D. Atkinson. 2003. Language-tree divergence times support the Anatolian theory of Indo-European origin. *Nature* 426:435–439.
- Haspelmath, M. 2007. Pre-established categories don't exist: Consequences for language description and typology. *Linguist. Typol.* 11:119–132.

- Hendriks, P., and C. Koster. 2010. Production/comprehension asymmetries in language acquisition. *Lingua* 120:1887–1897.
- Hinton, G. E., and S. J. Nowlan. 1987. How learning can guide evolution. *Complex Systems* 1:495–502.
- Jackendoff, R. 2002. *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford, U.K.: Oxford University Press.
- Jäger, H., A. Baronchelli, E. Briscoe et al. 2009. What can mathematical, computational and robotic models tell us about the origins of syntax? In *Biological Foundations and Origin of Syntax, Strüngmann Forum Reports, Vol. 3*, D. Bickerton and E. Szathmáý, eds. Cambridge, MA: MIT Press, 385–410.
- Kim, P. M., J. O. Korbel, and M. B. Gerstein. 2007. Positive selection at the protein network periphery: Evaluation in terms of structural constraints and cellular context. *Proc. Natl. Acad. Sci. USA* 104:20274–20279.
- Kirby, S. 1999. *Function, Selection and Innateness: The Emergence of Language Universals*. Oxford, U.K.: Oxford University Press.
- Kirby, S., M. Dowman, and T. Griffiths. 2007. Innateness and culture in the evolution of language. *Proc. Natl. Acad. Sci. USA* 104:5241–5245.
- Kirby, S., and J. Hurford. 1997. Learning, culture and evolution in the origin of linguistic constraints. In *Fourth European Conference on Artificial Life*, P. Husbands and I. Harvey, eds. Cambridge, MA: MIT Press, 493–502.
- Kirby, S., and J. Hurford. 2002. The emergence of linguistic structure: An overview of the iterated learning model. In *Simulating the Evolution of Language*, A. Cangelosi and D. Parisi, eds. London, U.K.: Springer-Verlag, 121–147.
- Munroe S., and A. Cangelosi. 2002. Learning and the evolution of language: The role of cultural variation and learning cost in the Baldwin effect. *Artif. Life* 8:311–339.
- Nowak, M. A., N. L. Komarova, and P. Nyogi. 2002. Computational and evolutionary aspects of language. *Nature* 417:611–617.
- Odling-Smee, F. J., K. N. Laland, and M. W. Feldman. 2003. *Niche Construction: The Neglected Process in Evolution*. Princeton, NJ: Princeton University Press.
- Pinker, S. 2003. Language as an adaptation to the cognitive niche. In *Language Evolution*, M. H. Christiansen and S. Kirby, eds. New York, NY: Oxford University Press, 16–37.
- Pinker, S., and P. Bloom. 1990. Natural language and natural selection. *Behav. Brain Sci.* 13:707–784.
- Pinker, S., and R. Jackendoff. 2005. The faculty of language: What’s special about it? *Cognition* 95:201–236.
- Puglisi, A., A. Baronchelli, and V. Loreto. 2008. Cultural route to the emergence of linguistic categories. *Proc. Natl. Acad. Sci. USA* 105:7936–7940.
- Real, F., and M. H. Christiansen. 2009. Sequential learning and the interaction between biological and linguistic adaptation in language evolution. *Interact. Stud.* 10:5–30.
- Smith, K., and S. Kirby. 2008. Cultural evolution: Implications for understanding the human language faculty and its evolution. *Philos. Trans. R. Soc. B* 363:3591–3603.
- Steels, L., and T. Belpaeme. 2005. Coordinating perceptually grounded categories through language: A case study for colour. *Behav. Brain Sci.* 28:469–489.
- Thiessen, D., and B. Gregg. 1980. Human assortative mating and genetic equilibrium: An evolutionary perspective. *Ethol. Sociobiol.* 1:111–140.
- Tomasello, M. 2003. On the different origin of symbols and grammar. In *Language Evolution*, M. H. Christiansen and S. Kirby, eds. New York, NY: Oxford University Press, 94–110.
- Weber, B. H., and D. J. Depew. 2003. *Evolution and Learning: The Baldwin Effect Reconsidered*. Cambridge, MA: MIT Press.
- Yamauchi, H. 2001. The difficulty of the Baldwinian account of linguistic innateness. In *ECAL01*, J. Kelemen and P. Sosik, eds. Prague, Czech Republic: Springer, 391–400.

Copyright of Human Biology is the property of Wayne State University Press and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.