

# Coevolution of the language faculty and language(s) with decorrelated encodings

Ted Briscoe

Computer Laboratory  
University of Cambridge  
ejb@cl.cam.ac.uk

<http://www.cl.cam.ac.uk/users/ejb>

## 1 Introduction

Renewed interest in studying human language from an evolutionary perspective stems in large part from a growing understanding that not only the language faculty but also languages themselves can be profitably modelled as evolving systems. Deacon (1997) offers persuasive arguments for, and an extended introduction to, this approach. Working within this framework, a number of researchers have proposed that the language faculty evolved via genetic assimilation (sometimes referred to as the Baldwin Effect) in response to the emergence of (proto)languages – see Briscoe (2003) for a detailed review of these proposals and counter arguments.

Yamauchi (2000, 2001) argues that accounts of the fit between the language faculty and languages which invoke genetic assimilation are suspect because they assume correlation between the genetic encoding of the language faculty and the phenotypic encoding of nativised linguistic constraints. He describes a simulation in which he shows that progressively decorrelating the encodings slows, and finally prevents, genetic assimilation of linguistic information. In this paper, I argue that the decorrelation argument does not undermine the account of the evolution of the language faculty via genetic assimilation, nor the extended coevolutionary account in which the evolving language faculty in turn exerts linguistic selection pressure on languages.

Briscoe (1999, 2000a, 2002) describes a correlated simulation model in which coevolution of the language faculty and language(s) reliably occurs, shortening the learning period and exerting linguistic selection pressure on language change. Here I report the results of new experiments using the extant simulation in which the genotypic encoding of the language faculty and the phenotypic encoding of the starting point for language learning are progressively decorrelated. The results show that decorrelation generally increases the probability of preemptive rather than assimilative evolution of the language faculty. However, because languages coevolve rapidly, preemptive mutations which spread genetically lead to rapid compensatory linguistic changes, so that the ‘fit’ between language and the language faculty remains close. As it is only the detailed timing of genetic and linguistic changes which can discriminate preemptive from assimilative genetic

change, it is likely that these two scenarios will never be discriminated empirically. However, high degrees of decorrelation lead to overall higher failure rates in language acquisition, and eventually either to breakdown in communication or to highly restricted linguistic systems which are largely genetically-encoded. In reality, language acquisition and communication are very robust aspects of human behaviour and languages appear, if anything, to accrue complexity rather than lose expressiveness (e.g. McWhorter, 2001), making the predicted evolutionary dynamic implausible.

Section 2 describes genetic assimilation and its putative role in the evolution of the language faculty, arguing for a coevolutionary model in which languages themselves both influence and are influenced by the evolution of the language faculty. Section 3 describes my correlated simulation model and section 4 summarises results demonstrating genetic assimilation of grammatical information and analyses the critical assumptions behind these results. Section 5 evaluates extant work on decorrelation, describes modifications to my model to allow for progressive decorrelation, and details how this affects the original results. Section 6 argues that consideration of decorrelation strengthens the case for a coevolutionary account of the emergence and evolution of the language faculty based primarily on genetic assimilation.

## 2 From Genetic Assimilation to Coevolution

Genetic assimilation is a neo-Darwinian mechanism supporting apparent ‘inheritance of acquired characteristics’ (e.g. Waddington, 1942, 1975). The fundamental insights are that: 1) plasticity in the relationship between phenotype and genotype is under genetic control, 2) novel environments create selection pressures which favour organisms with the plasticity to allow within-lifetime (so-called epigenetic) developmental adaptations to the new environment, 3) natural selection will function to ‘canalise’ these developmental adaptations by favouring genotypic variants in which the relevant trait develops reliably on the basis of minimal environmental stimulus, providing that the environment, and consequent selection pressure, remains constant over enough generations. For example, humans are unique amongst mammals in their ability to digest milk in adulthood. Durham (1991) argues that the development of animal husbandry created an environment in which this ability conferred fitness, as milk now became a particularly reliable source of nutrition. This created selection pressure for individuals with a genetic make-up for increased ability to digest milk later in life. The apparent ‘feedback’ from environment to genotype is nothing more (nor less) than natural selection for a hitherto neutral variant within the human genotype which became advantageous as a consequence of cultural innovation.

Waddington (1975:305f) suggests that genetic assimilation provided a possible mechanism for the evolution of a language faculty:

‘If there were selection for the ability to use language, then there would be selection for the capacity to acquire the use of language, in an interaction with a language-using environment; and the result of selection for epigenetic responses can be, as we have seen, a gradual accumulation of so many genes with effects tending in this direction that the character gradually becomes genetically assimilated.’

In other words, the ability to learn is a genetic endowment with slight variation between individuals, and individuals with hitherto neutral variants allowing more rapid and/or reliable language acquisition would be selected for in the novel language-using environment.

Briscoe (1999, 2000a) speculates that an initial language acquisition procedure emerged via recruitment (exaptation) of preexisting (preadapted) general-purpose learning mechanisms to a specifically-linguistic cognitive representation capable of expressing mappings from decomposable meaning representations to realisable, essentially linearised, encodings of such representations (see also Bickerton, 1998; Worden, 1998). The selective pressure favouring such a development, and its subsequent maintenance and refinement, is only possible if some protolanguage(s) had already emerged within a hominid population, supporting successful communication and capable of cultural transmission, that is, learnable without a language-specific faculty (e.g. Deacon, 1997; Kirby and Hurford, 1997).

Protolanguage(s) may have been initially similar to those advocated by Wray (2000) in which complete, propositional messages are conveyed by undecomposable signals. However, to create selection pressure for the emergence of grammar, and thus for a faculty incorporating language-specific grammatical bias, protolanguage(s) must have evolved at some point into decomposable utterances, broadly of the kind envisaged by Bickerton (1998). Several models of the emergence of syntax have been developed (e.g. Kirby, 2001, 2002; Nowak *et al* 2000). At the point when the environment contains language(s) with minimal syntax, genetic assimilation of grammatical information becomes adaptive, under the assumption that mastery of language confers a fitness advantage on its users, since genetic assimilation will make grammatical acquisition more rapid and reliable.

Given that genetic assimilation only makes sense in a scenario in which evolving (proto)languages create selection pressure, Waddington’s notion of genetic assimilation should be embedded in the more general one of coevolution (e.g. Kauffman, 1993:242f). Waddington, himself, (1975:307) notes that if there is an adaptive advantage to shortening the acquisition period, then we might expect genetic assimilation to continue to the point where no learning would be needed, because a fully-specified grammar had been encoded. In this case acquisition would be instantaneous and fitness would be maximised in a language-using population. However, given a coevolutionary scenario, in which languages themselves are complex adaptive systems (e.g. Kirby, 1998; Briscoe, 2000b), a plausible explanation for continuing grammatical diversity is that social factors favouring innovation

and diversity create conflicting linguistic selection pressures (e.g. Nettle, 1999). Genetic transmission, and thus assimilation, are much slower than cultural transmission. Therefore, continued plasticity in grammatical acquisition is probable, because assimilation will not be able to ‘keep up with’ all grammatical change. Furthermore, too much genetic assimilation, or canalisation to use Waddington’s term, will reduce individuals’ fitness, if linguistic change subsequently makes it hard or impossible for them to acquire an innovative grammatical (sub)system.

### 3 The Coevolutionary Simulation Model

The model is a stochastic computational simulation consisting of an evolving population of language agents (LAgts). LAgts are endowed with the ability to acquire a grammar by learning. However, the starting point for learning, and thus LAgts’ consequent success, is determined to an extent by an inherited genotype. Furthermore, the fitness of a LAg (that is, the likelihood with which LAgts will produce offspring) is determined by their communicative success. Offspring inherit starting points for learning (genotypes) which are based on those of their parents. Inheritance of *starting* points for learning prevents any form of actual (Lamarckian) inheritance of acquired characteristics, but allows for genetic assimilation, in principle. Inheritance either takes the form of crossover of the genotypes of the parents, resulting in a shared, mixed inheritance from each parent, and overall loss of variation in genotypes over generations, and/or random mutation of the inherited genotype, introducing new variation.

#### 3.1 Language Agents

A language agent (LAg) is a model of a language learner and user consisting of a 1) learning procedure,  $LP$ , which takes a definition of a universal grammar,  $UG$ , and a surface-form:logical-form ( $LF$ ) pair or ‘trigger’,  $t$  and returns a specific grammar,  $g$ ; 2) a parser,  $P$  which takes a grammar and a trigger,  $t$ , and returns a logical form,  $LF$ , for  $t$ , if  $t$  is parsable with  $g$ , and otherwise reports failure; and 3) a generator,  $G$ , which given a grammar,  $g$ , and a randomly selected  $LF$  produces a trigger compatible with this  $LF$ .

I have developed several accounts of  $LP$  based on a theory of  $UG$  utilising a generalised categorial grammar and an associated parsing algorithm  $P$  (Briscoe, 2000a). In what follows, I assume the Bayesian account of parametric learning developed in Briscoe (1999, 2002) with minor modifications. Grammatical acquisition consists of incrementally adopting the most probable grammar defined by  $UG$  compatible with the  $n$ th trigger in the sequence seen so far:

$$g = \operatorname{argmax}_{g \in UG} p(g) p(t_n | g)$$

Briscoe (1999, 2002) shows how this formula can be derived from Bayes' theorem and how prior probability distributions can be placed on  $g \in UG$  in terms of the number and type of parameters required to define  $g$ , broadly favouring smaller grammars and greater linguistic regularity. The probability of  $t$  given  $g$  is defined in terms of the posterior probabilities of the grammatical categories required to parse  $t$  and recover the correct  $LF$ . These posterior probabilities are updated according to Bayes' theorem after each new trigger is parsed and  $LP$  has searched a local space or neighbourhood around  $g$ , defined parametrically, to find a parse for  $t$ , if necessary. Although the evidence in favour of a given parameter setting is thus statistical, all parameters have finite, discrete values. So at any point, LAgts use a single grammar, though learning LAgts incrementally reassess the evidence in favour of all possible parameter settings.

In the experiments reported below,  $LP$  does not vary – however, the starting point for learning and the hypothesis space are varied. This starting point is 20 binary-valued individual p-settings, representing principles or parameters, which define possible grammars and the exact prior probability distribution over them.  $P\text{-setting}(UG)$  encodes both prior and posterior probabilities for p-setting values during LAgts learning and thus defines both the starting point for learning (initial p-settings) and which grammar if any a LAgts has currently internalised. Each individual p-setting is represented by a fraction:  $\frac{1}{2}$  maps to an unset parameter with no prior bias on its value;  $\frac{1}{5}$  and  $\frac{4}{5}$  represent default parameters with a prior bias in favour of one or other specific setting. However, this bias is low enough that consistent evidence for the alternative setting during learning will allow  $LP$  to move the posterior probability of this parameter through the  $\frac{1}{2}$  (unset) point to take on its other setting. Principles, which have been nativised, have prior probabilities sufficiently close to 1 or 0, typically  $\frac{1}{50}$  or  $\frac{49}{50}$ , that  $LP$  will not see enough evidence during learning to alter their (absolute) settings.

How a p-setting is initialised for specific LAgts determines their exact learning bias and hypothesis space. The 'weakest' language faculty variant is one in which all p-settings are unset parameters, so there is no prior bias or constraint in favour of any specific grammar. If all p-settings are principles, either a single grammar is already internalised or no grammar is learnable (as some 'off' settings preclude any form of sentence decomposition or are mutually incompatible). Mutation and one-point crossover operators are defined over p-settings and designed not to bias evolution towards adoption of any one of the three types of p-setting. However, if default settings or principles evolve this clearly constitutes genetic assimilation of grammatical information because it creates either learning biases or constraints in favour of subclasses of grammars with specific grammatical properties. This is additional to a general and domain-independent bias in favour of small grammars, and thus linguistic generalisation and regularity, which is a consequence of the Bayesian formulation of  $LP$  (see Briscoe, 1999, 2002).

The space of possible grammars in  $UG$  is defined in terms of canonical constituent order, possible non-canonical ordering, and categorial complexity. The

account of *UG* and associated p-settings is based on the typological literature on attested variation (e.g. Croft, 1990) and treats most variation as, in principle, independent (as each individual default or unset p-setting can be (re)set independently during learning). *UG* defines 70 full languages and a further 200 subset languages of these full languages, generated by subset grammars which have some parameters unset or off so that some triggers from the corresponding full language cannot be generated or parsed. Further details of these grammars and language fragments are given in Briscoe (2000a). Default or absolute p-settings, therefore create clear and concrete forms of specifically grammatical learning bias or constraint in favour of specific constituent orders, and so forth. Simulation of *LP* on samples of the full languages confirms that there are many prior distributions which allow successful acquisition given sufficient triggers, where success is defined as convergence to the target grammar  $g^t \in UG$  with high probability  $P - \varepsilon$ , given an arbitrary sequence of  $n$  triggers drawn randomly from a fair sample for the target language (see Briscoe, 2002 for more details of the learning framework).

In addition, each LAgt has an age, between 1 and 10, and a fitness, between 0 and 1. LAgts can learn until they exceed age 4 and pairs of LAgts can linguistically interact (INT i.e. parse or generate) with whatever grammar they have internalised between 1 and 10. The simplest version of fitness measures LAgts' communicative success as a ratio of successful to all interactions, but other factors can be included in the fitness function, such as the degree of expressiveness of the grammar acquired. A successful interaction (SUCC-INT) occurs when the trigger generated by one LAgt can be parsed by the other LAgt to yield the same *LF*. This does not necessarily require that the LAgts share identical grammars. In summary, a language agent has the following components:

LAgt:	
$LP(UG, t)$	$= g$
$P(g, t)$	$= LF$
$G(g, LF)$	$= t$
Age :	[1 – 10]
Fitness :	[0 – 1]

## 3.2 Populations and Speech Communities

The operations which define the simulation model for an evolving population of linguistically interacting LAgts, such as INT and SUCC-INT above, and the other (capitalised) ones described below are summarised at the end of this section.

A population (POP) is a changing set of LAgts. Time steps of the simulation consist of interaction cycles (INT-CYCLE) during which each LAgt participates in a prespecified number of interactions. On average each LAgt generates (parses)

for half of these interactions. LAGts from the entire population interact randomly without bias. After each cycle of interactions, the age of each LAGt is incremented and those over age 10 are removed, the fitness of each LAGt over that interaction cycle is computed, and LAGts aged 4 or over who have greater than mean fitness reproduce (REPRO) a new LAGt by single-point crossover of their *P-setting* with another such LAGt with whom they have successfully interacted at least once. The resulting p-setting may also optionally undergo random unbiased single-point mutation creating new p-setting values at specific loci. The number of new LAGts per timestep is capped to prevent the proportion of learning LAGts exceeding one-third of the overall population. Capping is implemented by random selection from the pool of offspring created from the fitter, interacting LAGts. Alternatively, in order to simulate the situation in which there is no natural selection for LAGts based on communicative success, a prespecified number of new LAGts can be created by unbiased random selection of parent LAGts.

The mean number of interactions (INT) per interaction cycle is set (typically, at 30 INT per LAGt per INT-CYCLE) so that, despite stochastic sampling variation, accurate grammatical acquisition by all learning LAGts is highly probable from many of the possible initialisations of *UG*, including the ‘weakest’ language faculty in which all p-settings are unset parameters.

If a simulation run is initialised with no mutation and a mixed age population of adult LAGts sharing the same initial p-setting and same internalised full grammar, then grammatical acquisition by subsequent generations of learning LAGts will be >99% accurate, and communicative success (i.e. the proportion of successful interactions) will average 98%, the 2% accounting for learners who have temporarily internalised a subset grammar. Under these conditions, the population constitutes a stable homogeneous speech community, in which no significant grammatical variation is present and no grammatical change takes place.

If grammatical variation is introduced into such a speech community, then linguistic drift, analogous to genetic drift, means that the population will reconverge on one variant within around  $2N$  time steps (where  $N$  is population size) due to sampling effects on LAGts’ input during learning (Briscoe, 2000b). Grammatical variation can be introduced by initialising the simulation with LAGts who have internalised different grammars or by periodic migrations (MIGRATE) of groups of such adult LAGts. Such migrations (crudely) model contact between speech communities. However, the dynamic of the simulation is always to recreate a single homogeneous speech community with a high overall communicative success because all LAGts in the current population interact with each other with equal probability, regardless of the grammar they have internalised, their provenance, or their age.

Linguistic selection, as opposed to drift, occurs whenever any factor, other than the proportion of LAGts who have internalised a grammatical variant, plays a role in this variant’s ability to be passed on to successive generations of learning LAGts. Such factors might be the relative parsability of variants and their

consequent learnability, the probability with which they are generated, the degree to which any learning bias or constraint militates for or against them, their expressiveness, social prestige, and so forth. In simple cases of linguistic selection (Briscoe, 2000a), the population typically converges on the more adaptive variant within  $N$  timesteps (where  $N$  is population size). In this simulation model, once linguistic variation is present there is a tendency for populations to converge on subset grammars and the languages associated with them. These grammars require fewer parameters to be set and thus can be learnt faster. However, if all LAgts utilise the same subset language, then communicative success will remain high. This tendency can be countered by introducing a further factor into LAgT fitness which adds an extra cost for utilising a subset grammar each time a LAgT generates a sentence, as in the fitness function defined below. This creates selection for grammars able to economically express the widest range of  $LF$ s (see Briscoe, 2000a for more detailed discussion).

Linguistic selection can occur without natural selection for, or any genetic evolution of, LAgts so long as their initial *P-setting* contains principles or default parameters and thus creates learning constraint or bias. If genetic mutation is enabled and reproduction is random, then simulation runs inevitably end with populations losing the ability to communicate because accumulated genetic drift in p-settings eventually prevents learning LAgts acquiring any compatible grammars. However, if LAgT reproduction is fitness-guided and genetic evolution is possible via crossover with or without mutation, then there is 1) modest natural selection for p-settings which shorten the learning process and increase fitness at age 4, since LAgts who have yet to acquire a full grammar are unlikely to have higher than average fitness and thus reproduce at the end of this interaction cycle, and 2) more severe natural selection for p-settings which allow reliable, accurate grammatical acquisition by the end of the learning period at age 5, since non-communicators will not reproduce at all and subset language speakers are likely to have below average fitness for the remainder of their lives.

Figure 3.2 illustrates how the model incorporates both biological evolution of LAgT p-settings and linguistic evolution of languages, represented as sets of triggers which form the data for successive generations of learners. The model's principal components, operations and parameters are defined more succinctly below:

LAgT:  $\langle P\text{-setting}(UG), \text{Parser}, \text{Generator}, \text{Age}, \text{Fitness} \rangle$

AGE: 1-4 LAgT = Learning lLAgT / 4-10 = adult reproducing aLAgT

POP <sub>$n$</sub> : {LAgT<sub>1</sub>, LAgT<sub>2</sub>, ... LAgT <sub>$n$</sub> }



INT: (LAg<sub>t<sub>i</sub></sub>, LAg<sub>t<sub>j</sub></sub>),  $i \neq j$ , Gen(LAg<sub>t<sub>i</sub></sub>, t<sub>k</sub>), Parse(LAg<sub>t<sub>j</sub></sub>, t<sub>k</sub>)

SUCC-INT: Gen(LAg<sub>t<sub>i</sub></sub>, t<sub>k</sub>)  $\mapsto$  LF<sub>k</sub>  $\wedge$  Parse(LAg<sub>t<sub>j</sub></sub>, t<sub>k</sub>)  $\mapsto$  LF<sub>k</sub>

INT-CYCLE:  $\approx$  30 ints. / LAg<sub>t</sub>

REPRO: (aLAg<sub>t<sub>i</sub></sub>, aLAg<sub>t<sub>j</sub></sub>),  $i \neq j$ ,  
Create-lLAg<sub>t</sub>(Mutate(Crossover(P-setting(aLAg<sub>t<sub>i</sub></sub>, P-setting(aLAg<sub>t<sub>j</sub></sub>))))))

MIGRATE: (POP<sub>n</sub>), For i=0 to n/3  
Create-aLAg<sub>t</sub>(Flip(P-setting(Dominant-Lg-LAg<sub>t</sub>(POP<sub>n</sub>))))

**LAg<sub>t</sub> Fitness** (Costs/Benefits per INT):

1. Generate cost: 1 (GC)
2. Generate subset cost: 1 (GSC)
3. Parse cost: 1 (PC)
4. Success benefit: 1 (SI)
5. Fitness function:  $\frac{SI}{GSC+GC+PC}$

## 4 Results and Critical Assumptions

Previous experiments (Briscoe, 1999, 2000a, 2002) a variety of different initial *P-setting* configurations and several variants of *LP* have demonstrated that genetic assimilation occurs with natural selection for communicative success and that populations continue to utilise full grammars and associated languages if there is also natural selection for expressiveness. Inducing rapid linguistic change through repeated migrations does not prevent genetic assimilation, though it does cause it to asymptote rather than reach a point where the entire population fixates on a *P-setting* defining a single nativised grammar. Rapid linguistic change also creates a preference for the genetic assimilation of default parameters over principles, since the latter are potentially more damaging when subsequent linguistic change renders a principle maladaptive for learners. If the population were exposed to the entire space of grammatical variation within the time taken for a variant *p-setting* to go to fixation, then assimilation would not occur. However, for this to happen, the rate of linguistic change would be so great that communication

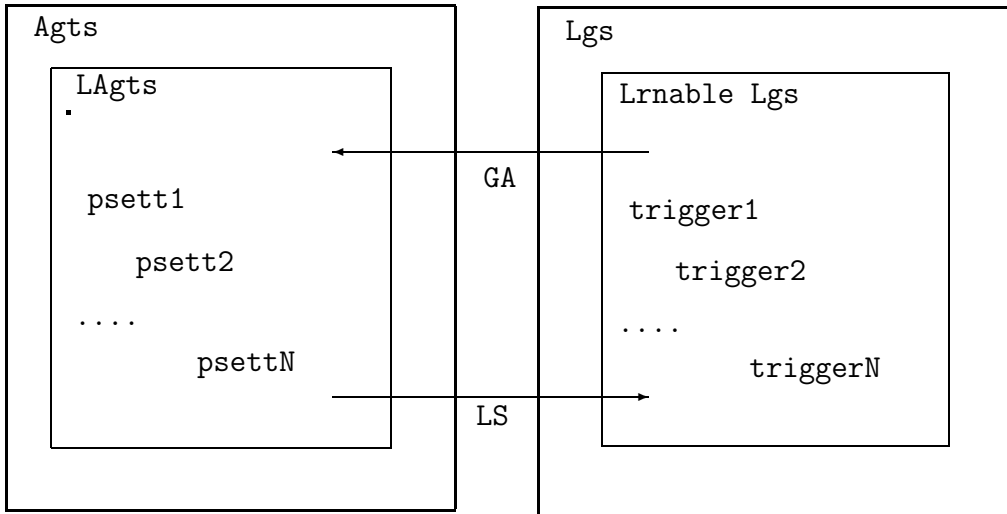


Figure 1: Simulation Model: Linguistic Selection (LS), Genetic Assimilation (GA), Initial P-settings (psettN), Trigger sentence types (triggerN)

would breakdown and the population would not constitute a speech community in which the majority of interactions are successful. Below I describe one such experiment using the *LP* and simulation model outlined above, first reported in Briscoe (1999).

Populations of adult LAgts were created with initial p-settings consisting of 3 principles and 17 unset parameters. In each simulation run, the first generation of LAgts all utilised one of seven typologically-attested full grammars (see Briscoe, 2000a for the linguistic details). 70 runs were performed – 10 under each condition. Simulation runs lasted for 2000 interaction cycles (about 500 generations of LAgts). Constant linguistic heterogeneity was ensured by migrations of adult LAgts speaking a distinct full language at any point where the dominant (full) language utilised by the population accounted for over 90% of interactions in the preceding interaction cycle. Migrating adults accounted for approximately one-third of the adult population and were set up to have an initial *P-setting* consistent with the dominant settings already extant in the population; that is, migrations were designed to increase linguistic, and decrease genetic, variation.

Over all these runs, the mean increase in the proportion of default parameters was 46.7%. The mean increase in principles was 3.8%. Together these accounted for an overall decrease of 50.6% in the proportion of unset parameters in the initial p-settings of LAgts. Qualitative behaviour in all runs showed increases in default parameters and either maintenance or increase in principles. Figure 2 shows the relative proportions of default parameters, unset parameters and principles in the overall population and also mean fitness for one such run. Overall fitness increases as the learning period is truncated, though there are fluctuations caused by migrations and/or by higher proportions of learners.

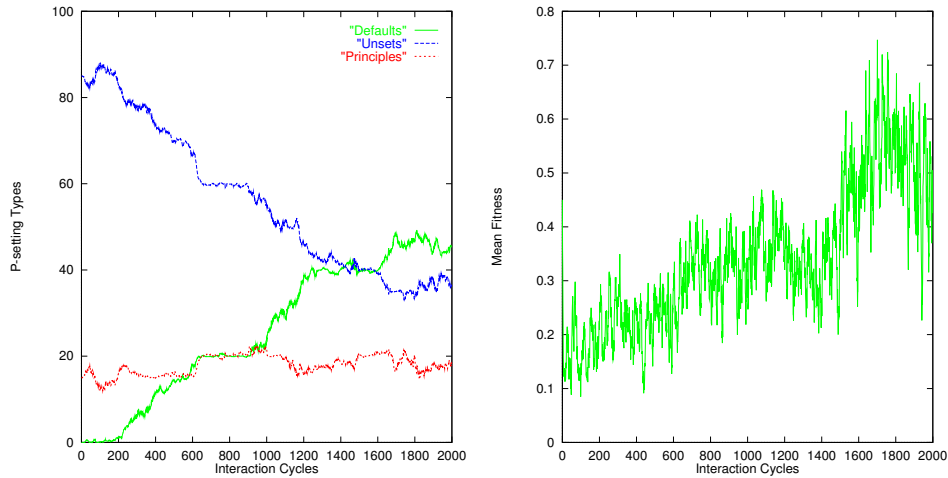


Figure 2: Proportions of P-setting Types and Mean fitness

In these experiments, linguistic change (defined as the number of interaction cycles taken for a new parameter setting to be acquired by all adult LAGts in the population) is about an order of magnitude faster than the speed with which a genetic change (i.e. a new initial p-setting) can go to fixation. Typically, 2-3 grammatical changes occur during the time taken for a principle or default parameter to go to fixation. Genetic assimilation remains likely, however, because the space of grammatical variation (even in this simulation) is great enough that typically the population is only sampling about 5% of possible variations in the time taken for a single p-setting variant to go to fixation (or in other words, 95% of the environmental pressure is constant during this period).

Many contingent details of the simulation model are arbitrary and unverifiable, such as the size of the evolving population, the number of learnable grammars, and relative speed at which LAGts and languages can change. These have been varied as far as possible to explore whether they affect the results. Importantly, it seems likely that the simulation model massively *underestimates* the true space of grammatical possibilities. Thus, there would very probably have been more opportunity to restrict the hypothesis space by genetic assimilation than is predicted by the simulation model as more of this space would have gone unsampled during the period of adaptation for the language faculty. Nevertheless, there is a limit to genetic assimilation in the face of ongoing linguistic change: in simulation runs with LAGts initialised with all default parameters, populations evolve away from this ‘fully-assimilated’ *P-setting* when linguistic variation is maintained. Briscoe (2000a) reports variant experiments and discusses these issues in more detail.

## 5 Decorrelation

The relationship between genes and traits is rarely one-to-one, so epistasis (or ‘linkage’ of several genes to a single trait) and pleiotropy (or ‘linkage’ of a single gene to several traits) are the norm. In general, one effect of epistasis and pleiotropy will be to make the pathways more indirect from selection pressure acting on phenotypic traits to genetic modifications increasing the adaptiveness of those traits. Therefore, in general terms, we would expect a more indirect and less correlated genetic encoding of a trait to impede, slow or perhaps even prevent genetic assimilation. Mayley (1996) explores the effects of manipulating the correlation between genotype (operations) and phenotype (operations) on genetic assimilation. In his model, individuals are able to acquire better phenotypes through ‘learning’ (or another form of within-lifetime plasticity), thus increasing their fitness. However, Mayley demonstrates that in his model the degree to which the acquired phenotype can be assimilated into the genotype of future generations, thus increasing overall fitness, depends critically on the degree of correlation.

Yamauchi (2000, 2001) replicates Turkel’s (2002) simulation demonstrating genetic assimilation of grammatical principles. However, he then manipulates the degree of correlation in the encoding of genotype and phenotype. He represents grammar space as a sequence of  $N$  principles or parameters but determines the initial setting at each locus of the phenotype from a look-up table which uses  $K$  0/1s (where  $K$  can range from 1 to  $N - 1$  to encode each setting (one of two binary values or unset, i.e. a parameter), and to ensure that all possible phenotypes can be encoded. A genotype is represented as a sequence of  $N$  0/1s. A translator reads the first  $K$  genes from the genotype and uses the look-up table to compute the setting of the first locus of the phenotype. To compute, the setting of the second locus of the phenotype, the  $K$  genes starting at the second locus of the genotype are read and looked up in the table, and so on. The translator ‘wraps around’ the genotype and continues from the first gene locus when  $K$  exceeds the remaining bits of the genotype sequence.

Yamauchi claims, following Kauffman (1993), that increases in  $K$  model increases in pleiotropy and epistasis. Increased  $K$  means that a change to one locus in the genotype will have potentially more widespread and less predictable effects on the resulting phenotype (as the translator will ‘wrap around’ more frequently). It also means that there is less correspondence between a learning operation, altering the value of single phenotypic locus, and a genetic operation. The latter may potentially alter many phenotypic loci in differing ways, or perhaps alter none, depending on the look-up table. For low values of  $K$ , genetic assimilation occurs, as in Turkel’s and my model, for values of  $K$  around  $N/2$  genetic assimilation is considerably slowed, and for very high values ( $K = N - 1$ ) it is stopped.

Yamauchi does not consider how the progressive decorrelation of phenotype

from genotype affects the language acquisition process, the degree of communicative success achieved, or how linguistic systems might be affected. In part, the problem here is that the abstract nature of Turkel’s simulation model does not support any inference from configurations of the phenotype to concrete linguistic systems. Yamauchi, however, simply does not report whether increasing decorrelation, that is higher values of  $K$ , affects the ability of the evolving population to match phenotypes via learning. The implication, though, is that, for high values of  $K$ , the population cannot evolve to a state able to match phenotypes and thus to support reliable communication.

Kauffman’s original work with the  $NK$  model was undertaken to find optimal values of  $K$  for given  $N$  to quantify the degree of epistasis and pleiotropy likely to be found in systems able to evolve most effectively. Both theoretical analysis and experiments which allow  $K$  itself to evolve suggest intermediate values of  $K$  are optimal (where the exact value can depend on  $N$  and other experimental factors). To stop assimilation, Yamauchi decorrelates his model to an extent to which Kaufmann’s (1993) results suggest will yield a dynamically chaotic and evolutionarily highly suboptimal system. But despite these caveats, Yamauchi’s simulation suggests that (lack of) correlation of genotype and phenotype with respect to the language faculty is as important an issue for accounts of genetic assimilation of grammatical information as it is for accounts of genetic assimilation generally.

My modified model supporting decorrelation does not distinguish genotype and phenotype, instead it utilises a single  $P$ -*setting* which encodes both the initial and subsequent states of the learning process. Initial p-settings (i.e. those encoded by the genotype) are defined by a sequence of fractions which define the prior probability of each of three possible p-setting types: unset parameter, default parameter and absolute principle. Arbitrary manipulation of the denominators and numerators of these fractions is very likely to result in values outside the range 0-1. For example, a  $NK$ -like scheme based on a binary sequential ‘genetic’ encoding of these fractions with single-point mutation by bit flipping will nearly always produce new absolute principles (under the fairly natural assumption that values outside the range 0-1 are interpreted this way).

Instead the mutation operator was modified to create unbiased movement of parameters between default and unset settings at *multiple* random points in a  $P$ -*setting*. The maximum number of p-settings that could be modified in a single mutational event is the *decorrelation factor* varied in these simulation runs. The fractional values defining prior probabilities remained prespecified, as defined in section 3.1 above, but the exact number modified, the points in the p-setting modified and the resultant settings, are all independent stochastic variables of each such event. Just as increasing  $K$  in Yamauchi’s model increases the maximum number of possible changes to the phenotype given a genetic mutation, so does increasing the decorrelation factor. Similarly, the possible *values* of phenotypic loci are not biased or otherwise altered by decorrelation in either scheme.

The actual number of loci altered by any given mutation depends on the look up table in stochastic interaction with the mutated gene in Yamauchi's simulation and on stochastic variables in mine. However, in both cases this number cannot exceed the absolute value of  $K$  or of the decorrelation factor.

New simulation runs were performed, identical to those reported in section 4 except that the degree of decorrelation between mutation of p-settings and parameter (re)setting was varied. Half the runs did not include migrations, as their general effect is to add greater linguistic variation and therefore to increase the potential for linguistic selection. Linguistic selection will also occur in a population in which the language faculty is evolving because different p-setting variants can force even a homogeneous speech community to shift to a new language. For example, if a new LAgT inherits a mutated p-setting which alters a default parameter setting, that learner may acquire a variant grammar compatible with the new default setting if the input sample does not exemplify the non-default setting reliably enough to override it. If that LAgT and some of its descendents achieve better than mean fitness, because the new setting is only relevant for a small subset of possible triggers, or because these LAgTs reset it successfully during learning, then the default initial setting may spread through the population. The likelihood of such LAgTs achieving better than mean fitness is lower in an environment where the remaining population are learning accurately and efficiently, but is increased in one in which other LAgTs are also inheriting mutated p-settings some of which disadvantage them more seriously.

The main effect of progressively decorrelating the mutation operator is to increase the rate of linguistic selection and, despite natural selection on the basis of expressiveness, to cause populations to converge on successively less expressive subset languages. Often, linguistic change is coextensive with a few LAgTs appearing who fail to learn any language. However, swift shifts to other (often less expressive) languages mean that other genetically similar LAgTs do acquire a language. Thus, although decorrelation modestly increases the number of subset learning, mislearning and nonlearning LAgTs, this, in turn, creates linguistic selection for other more learnable languages. When the decorrelation rate is very high, potentially affecting all of the p-settings during one mutational event, then the number of non-learners appears to go through a phase transition increasing about a thousand-fold over the previous increment. The mean percentage of mislearning LAgTs who do not acquire a full grammar or any grammar manifest in the environment is under 1% for low rates of decorrelation, rising to 4.5% for intermediate rates and to 24% for the highest rate. In 100% of the high decorrelation runs, populations converge on a minimal subset language, which is acquired by setting three p-settings, and in most cases, the population has fixated on correct default parameters for several of these settings. For intermediate rates of decorrelation less than 5% of runs end with the majority of the population acquiring a subset language, and for low rates none do.

Tracking the rate at which default parameters replace unset ones over these

runs reveals that this rate *increases* by about 5% over runs with correlated encodings, as measured by the number of default parameters in the population at the end of each run. This increase is broadly constant across all the runs regardless of the level of decorrelation and the presence or absence of migrations. However, as the decorrelation rate increases, the standard deviation of the mean also increases, reflecting the randomness of the potential changes induced by the increasingly dramatic mutational events. That is, for higher rates of decorrelation, distinct runs diverge more as the stochastic factors in the mutational operator affect the exact behaviour of individual runs to a greater extent. Similarly, even without migration and starting from a homogeneous linguistic environment, the mean number of distinct languages acquired by LAGts across a run increases with decorrelation from a mean of 20 for lower rates to 40 for intermediate rates, to 63 for the highest rate.

An increase in the number of default parameters in the language faculty only counts as genetic assimilation if the mutated defaults are compatible with the language(s) in the environment. Examining the timing of changes in the initial p-setting with linguistic changes reveals that decorrelated mutation is often the *cause* of a linguistic change, rather than assimilatory. These preemptive, non-assimilative mutations which spread and become adaptive are ones which drive rapid linguistic change, so that they rapidly become indistinguishable from assimilative ones. A default setting which is correct, and thus assimilative, in the current linguistic environment reduces the number of parameter settings required to learn the language, shortening the acquisition period and making it more robust against sampling variation in learner input.

If a correct default setting emerges via mutation, then it is likely to spread through the population, creating added linguistic selection pressure for subsequent linguistic change to remain compatible with the default setting. If a mutated default is incompatible with the current linguistic environment but manages to spread to other LAGts, either because grammatical acquisition is generally less accurate or because sampling variation allows enough learners to override the default without significant fitness cost, then it will exert increasing linguistic selection pressure, both because more learners will have the default setting and because less LAGts will generate the counterexamples that would cause the default setting to be overridden. For low rates of decorrelation a mean 20% of mutations going to fixation are preemptive, so 80% remain assimilatory. The mean percentage of preemptive mutations going to fixation rises to 45% for intermediate rates and to 99% for the highest rate of decorrelation.

The lefthand plot in Figure 3 shows the rate of increase of default parameter settings within the population for a low and high degree of decorrelation in two runs with no migrations and otherwise identical initialisations. The righthand plot shows the corresponding decrease in the number of parameters which are (re)set by learners in the same two runs. Although the overall increase in defaults is consistently higher, and the number of (re)sets is mostly correspondingly lower

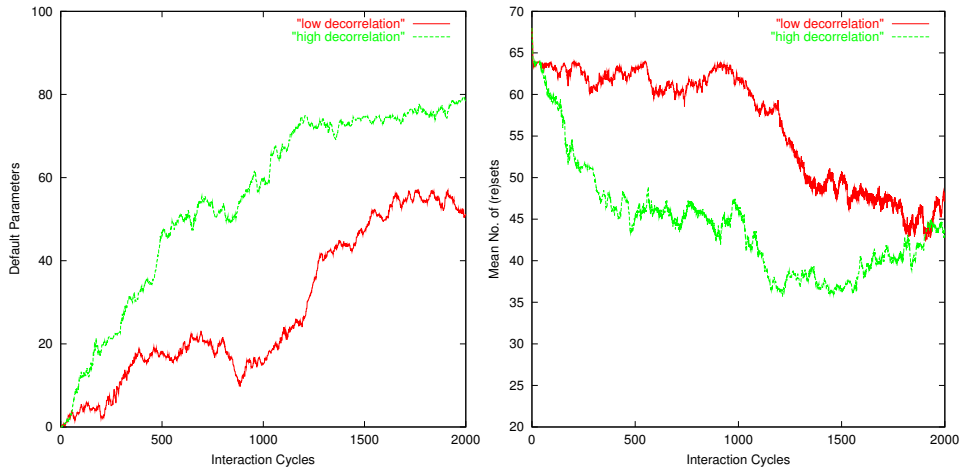


Figure 3: Change in Defaults and Resets with Low/High Decorrelation

with more decorrelation, in these runs the number of resets converge towards the end, because the ‘fit’ between the languages of the speech community and the language faculty tends to become less close with higher degrees of decorrelation. This is also a tendency in other runs with no migrations. However, the effect is removed and to some extent even reversed in runs with migrations, presumably because migrations provide linguistic variation supporting more rapid linguistic selection of variants more compatible with the continuously mutating language faculty.

One conclusion that can be drawn from these experiments is that if non-assimilatory random mutations were a factor in the evolution of the language faculty, then these mutations would rapidly mesh with linguistic systems, because of the greater speed and responsiveness of linguistic selection. Subsequently such mutations will appear to be cases of genetic assimilation of grammatical information, unless one has access to the precise nature and timing of the mutational event, linguistic environment and any subsequent linguistic change – something one cannot realistically hope to have access to outside the simulation ‘laboratory’. Preemptive mutations are quite compatible with the coevolutionary approach to the evolution of the language faculty and languages presented here. However, the experiments also suggest that high degrees of decorrelation (and thus non-assimilatory mutations) are unlikely, in line with Kauffman’s (1993) more general results. The predicted consequence of such changes is that linguistic (pre)history would be punctuated by the periodic emergence of mislearners and nonlearners sometimes coupled with bursts of rapid linguistic change, often in the direction of less expressive languages. This is contrary to what most non-assimilationists have argued (see Briscoe, 2003) and certainly not supported by the available evidence, which suggests that languages, if anything, accrue expressiveness and



hence complexity, and that acquisition remains robust.

A simulation model in which greater expressiveness, or acquisition of innovative grammatical variants, outweighed learnability in LAgt fitness might predict natural selection for a less restrictive language faculty. A model which integrated some of the social pressures maintaining linguistic diversity discussed, for example, by Nettle (1999), might counterbalance the tendency for learnability to outweigh expressiveness. However, given the inalienable relationship between the size of the hypothesis space and the amount of data required to reliably acquire a specific grammar (e.g. Nowak *et al* 2001), a model in which expressiveness regularly overcame learnability would predict that the learning period would increase over time or, if a critical period had been nativised, that the reliability of grammatical acquisition would degrade. The existence of a critical period for grammatical acquisition, the accuracy of grammatical acquisition, and its selectivity in the face of variant input (e.g. Lightfoot, 1999) all suggest that this is an implausible evolutionary dynamic. Nevertheless, integration of a more realistic account of expressiveness into the simulation model would certainly be a worthwhile extension of these experiments.

A further set of similar experiments was undertaken in which the mutational operator was modified so that the fractional values defining initially unset parameters mutated randomly by increasingly large amounts. The increasing bias of this operator is to create more principles as the base of the fractions increases and as they exceed the 0-1 range (see discussion above), so that *LP* becomes unable to move them through the  $\frac{1}{2}$  threshold which alters a setting. Unsurprisingly, in these experiments, there were many more cases of nonlearners, since principles rather than just default settings were acquired. This mutation operator is exceedingly unlikely to create new unset parameters, and increasingly likely to only create principles with greater degrees of decorrelation. Overall the rate of increase in defaults and principles was slightly higher in these experiments. However, just as in the previous ones, many of the mutational events are preemptive rather than assimilative, and where the preemption results in principles incompatible with the linguistic environment, a learning LAgt has less chance of reproducing, unless the overall accuracy of grammatical acquisition in the population has degraded significantly. Thus, as in the previous experiments, the trend in linguistic change is towards successive reconvergence on subset languages until the population is speaking a minimal subset language compatible with the principles or any remaining default parameter settings on which the population has converged.

In general, the greater the degree of decorrelated mutational events involving preemptive non-assimilatory changes, the more the simulation model predicts that the coevolutionary dynamic would bias the hypothesis space until only one grammatical system remained. If the mutation operator is prevented from creating principles or increasingly stronger defaults, as in the first series of experiments, then there is a limit to this effect, but removing this, as in the second series of experiments, simply strengthens this tendency.

## 6 Conclusions

In summary, extant models predict that genetic assimilation of grammatical information would have occurred given three crucial but plausible assumptions. Firstly, communicative success via expressive languages with compositional syntax conferred a fitness benefit on their users. Secondly, the linguistic environment for adaptation of the language faculty consistently provided grammatical generalisations to be genetically assimilated. Thirdly, at least some of these generalisations were neurally and genetically encodable with sufficient correlation to enable genetic assimilation. The simulation experiments I describe show that these assumptions can be incorporated in a precise model, that reliably shows genetic assimilation under conditions which support ongoing successful communication with an expressive language.

The simulation model predicts that if language confers no fitness benefit but the acquisition procedure is under genetic control, genetic drift would result in evolution of nonlearners and thus non-linguistic populations. It also predicts that, for there to be no grammatical generalisations capable of being genetically assimilated, language change would need to be impossibly rapid, to the point where speech communities would breakdown. Finally, it predicts that, if genotype and phenotype are decorrelated to the extent that assimilation is blocked, the resultant coevolutionary dynamic would lead inexorably towards simpler but less expressive subset languages. Thus, the case for genetic assimilation as the primary mechanism of the evolution of the language faculty remains, in my opinion, strong.

Nevertheless, the coevolutionary perspective on genetic assimilation of grammatical information raises two important caveats. Firstly, as languages themselves are adapted to be learnable (as well as parsable and expressive) and as languages change on a historical timescale, some of the grammatical properties of human languages were probably shaped by the process of cultural transmission of (proto)language via more general-purpose learning (e.g. Kirby, 1998, Brighton, 2002) prior to the evolution of the fully-formed language faculty. Secondly, whether the subsequent evolution of this faculty was assimilative, encoding generalisations manifest in the linguistic environment, or preemptive, with mutations creating side-effects causing linguistic selection for new variants, the fit between the learning bias of the language faculty and extant languages is predicted to be very close.

It is important to emphasise that modelling and simulation, however careful and sophisticated, are not enough to establish the truth of what remains a partly speculative inference about prehistoric events. The value of the simulations, and related mathematical modelling and analysis, lies in uncovering the precise set of assumptions required to predict that genetic assimilation of grammatical information will or will not occur. Some of these assumptions relate to cognitive abilities or biases which remain in place today – these are testable. Others,

such as the relative weight of factors relating to learnability and expressiveness in the LAgT fitness function in my simulation model remain largely speculative, though not, in principle, untestable, since they should, for example, affect attested grammatical changes, including those under intensive study right now (e.g. Kegl *et al* 1999). Other assumptions, such as the degree of correlation between genetic and neural encoding are theoretically plausible but empirically untestable using extant techniques.

## References

- Bickerton, D. (1998) ‘Catastrophic evolution: the case for a single step from protolanguage to full human language’ in J. Hurford, M. Studdert-Kennedy and C. Knight (ed.), *Approaches to the evolution of language*, Cambridge University Press, Cambridge, pp. 341–358.
- Brighton, H. (2002) ‘Compositional syntax through cultural transmission’, *Artificial Life*, vol.8.1, 25–54.
- Briscoe, E. (1999) ‘The Acquisition of Grammar in an Evolving Population of Language Agents’, *Electronic Trans. of Art. Intelligence (Special Issue: Machine Intelligence, 16)*. (ed) Muggleton, S., vol. Vol 3(B), [www.etaij.org](http://www.etaij.org), 44–77.
- Briscoe, E. (2000a) ‘Grammatical Acquisition: Inductive Bias and Coevolution of Language and the Language Acquisition Device’, *Language*, vol.76.2, 245–96.
- Briscoe, E. (2000b) ‘Evolutionary perspectives on diachronic syntax’ in Susan Pintzuk, George Tsoulas and Anthony Warner (ed.), *Diachronic Syntax: Models and Mechanisms*, Oxford University Press, Oxford, pp. 75–108.
- Briscoe, E. (2002) ‘Grammatical acquisition and linguistic selection’ in E. Briscoe (ed.), *Language acquisition and linguistic evolution: formal and computational approaches*, Cambridge: Cambridge University Press, pp. 255–300.
- Briscoe, E. (2003) ‘Grammatical Assimilation’ in M. Christiansen and S. Kirby (ed.), *Language evolution: the states of the art*, Oxford: Oxford University Press.
- Deacon, T. (1997) *The symbolic species: coevolution of language and brain*, MIT Press, Cambridge MA.
- Durham, W. (1991) *Coevolution, Genes, Culture and Human Diversity*, Stanford University Press, Palo Alto, Ca..
- Kauffman, S. (1993) *The Origins of Order: Self-Organization and Selection in Evolution*, Oxford University Press, New York.
- Kegl, J., A. Senghas, and Coppola, M. (1999) ‘Creation through contact: Sign language emergence and sign language change in Nicaragua’ in M. DeGraff (ed.), *Language Creation and Language Change: Creolization, Diachrony, and Development*, MIT Press, Cambridge MA.

- Kirby, S. (1998) ‘Fitness and the selective adaptation of language’ in J. Hurford, M. Studdert-Kennedy and C. Knight (ed.), *Approaches to the evolution of language*, Cambridge University Press, Cambridge, pp. 359–383.
- Kirby, S. (2001) ‘Spontaneous evolution of linguistic structure: an iterated learning model of the emergence of regularity and irregularity’, *IEEE Transactions on Evolutionary Computation*, vol.5.2, 102–110.
- Kirby, S. (2002) ‘Learning, bottlenecks and the evolution of recursive syntax’ in E. Briscoe (ed.), *Language Acquisition and Linguistic Evolution: Formal and Computational Approaches*, Cambridge University Press, Cambridge, pp. 173–204.
- Kirby, S. and Hurford, J. (1997) ‘Learning, culture and evolution in the origin of linguistic constraints’ in Phil Husbands and Imran Harvey (ed.), *4th European Conference on Artificial Life*, MIT Press, Cambridge, MA., pp. 493–502.
- Lightfoot, D. (1999) *The Development of Language: Acquisition, Change, and Evolution*, Blackwell, Oxford.
- Mayley, G. (1996) ‘Landscapes, learning costs and genetic assimilation’ in Peter Turney, Whitley, D., and Anderson, R. (ed.), *Evolution, learning and instinct: 100 years of the Baldwin effect*, MIT Press, Cambridge MA.
- McWhorter, J. (2001) *The Power of Babel: A Natural History of Language*, William Heinemann, London.
- Nettle, D. (1999) *Linguistic diversity*, Oxford: Oxford University Press.
- Nowak, M., Plotkin, J., and Jansen, V. (2000) ‘The evolution of syntactic communication’, *Nature*, vol.404, 495–498.
- Nowak, M., Komarova, N., and Niyogi, P. (2001) ‘Evolution of universal grammar’, *Science*, vol.291, 114–118.
- Turkel, W. (2002) ‘The learning guided evolution of natural language’ in E. Briscoe (ed.), *Language acquisition and linguistic evolution: formal and computational approaches*, Cambridge University Press, Cambridge, pp. 235–254.
- Waddington, C. (1942) ‘Canalization of development and the inheritance of acquired characters’, *Nature*, vol.150, 563–565.
- Waddington, C. (1975) *The evolution of an evolutionist*, Edinburgh: Edinburgh University Press.
- Worden, R. (1998) ‘The evolution of language from social intelligence’ in J. Hurford, M. Studdert-Kennedy and C. Knight (ed.), *Approaches to the evolution of language*, Cambridge University Press, Cambridge, pp. 148–168.
- Wray, A. (2000) ‘Holistic utterances in protolanguage: the link from primates to humans’ in C. Knight, M. Studdert-Kennedy and J. Hurford (ed.), *The Evolutionary Emergence of Language*, Cambridge University Press, Cambridge, pp. 285–302.
- Yamauchi, H. (2000) *Evolution of the LAD and the Baldwin Effect*, MA Dissertation, University of Edinburgh, Dept. of Linguistics.

Yamauchi, H. (2001) 'The difficulty of the Baldwinian account of linguistic innateness' in J. Keleman and P. Sosik (ed.), *Advances in Artificial Life: Proceedings of the 6th European Conference on Artificial Life*, Springer-Verlag, Heidelberg, pp. 391–400.