CO-EVOLUTION OF LANGUAGE-SIZE AND THE CRITICAL PERIOD[1]

James R Hurford and Simon Kirby,
Department of Linguistics, University of Edinburgh
jim@ling.ed.ac.uk , simon@ling.ed.ac.uk

to appear in David Birdsong (ed.)
New Perspectives on the Critical Period Hypothesis and Second Language Acquisition
Lawrence Erlbaum.

1 Introduction: gene-language co-evolution.

Species evolve, very slowly, through selection of genes which give rise to phenotypes well adapted[2] to their environments. The cultures, including the languages, of human communities evolve, much faster, maintaining at least a minimum level of adaptedness to the external, non-cultural environment. In the phylogenetic evolution of species, the transmission of information across generations is via copying of molecules, and innovation is by mutation and sexual recombination. In cultural evolution, the transmission of information across generations is by learning, and innovation is by sporadic invention or borrowing from other cultures. This much is the foundational bedrock of evolutionary theory.

But things get more complicated; there can be gene-culture co-evolution[3]. Prior to the rise of culture, the physical environment is the only force shaping biological evolution from outside the organism, and cultures themselves are clearly constrained by the evolved biological characteristics of their members. But cultures **become** part of the external environment, and influence the course of biological evolution. For example, altruistic cultures with developed medical knowledge reduce the cost to the individual of carrying genes disposing to certain pathologies (such as diabetes); and such genes become more widespread in the populations maintaining such cultures. Assortative mating can affect biological evolution, and particular cultures may influence the factors which are sorted for in mating. (For a careful discussion of the effects of cultural evolution on natural selection, see Cavalli-Sforza and Bodmer, 1971:774-804).

This paper examines mechanisms involved in the co-evolution of a biological trait, the

[1]This work was supported by two fellowships at the Collegium Budapest Institute for Advanced Study, and by Economic and Social Research Council research grant R000326551. We thank our colleagues at the Collegium Budapest, Eörs Szathmary and Axel Kowald for their helpful input.

[2]Not every property of an organism is adaptive, of course; spandrels do exist.

[3]Though not uncontroversial, the idea of gene-culture co-evolution has been developed in a variety of models, including Lumsden and Wilson (1981), Boyd and Richerson (1985); Dawkins and and Krebs (1984) propose a coevolutionary mechanism at the root of the evolution of signalling systems, and Deacon (1992) discusses human brain-language coevolution in detail.

critical period for language acquisition, and a property of human cultures, the **size** of their languages. A gene/culture interaction will be shown that can be described as a kind of symbiosis, but perhaps more aptly as an 'arms race'. In this introduction, we will sketch the basic mechanics of the interaction in very broad terms; the rest of the paper will explain and justify the details. The implications of our model for second language acquisition are given toward the end of the paper.

Put simply, the speed at which an individual can learn the language of the community, plus a critical period in which it can be learnt (both biologically given), together determine the maximum size of the language the individual can command as an adult. As this is true for all individuals, a limit on the size of the language as it exists in the community, and the typical age-span in which it can be learnt, are determined by these biological factors. With no biological mutation and no cultural innovation (e.g. invention or borrowing of novel expressions or even of new constructions), the interaction of biology and culture remains static. But mutation and cultural innovation can give the interaction an interesting dynamic.

Assume a biologically uniform population using a language of a fixed size. Say also that, in this hypothetical situation, the biology and the language are 'in harmony', in the sense that all individuals learn at a speed which enables them to learn the community language just by the time their critical period elapses. Taking a simple view of 'size' (see discussion below), the language fits neatly into the time biologically allotted for its acquisition. In this situation, there is no possibility of any lasting cultural innovation, as nobody would have any spare time within their critical period to acquire anything in addition to the existing language.

Now a biological mutant arises, who can acquire language faster, thus arriving at mastery of the community language some time before puberty. If there is an innovation now (perhaps by the mutant itself), there is at least one individual who can acquire it. If the mutant's relevant faster-language-acquisition genes spread through the community, more people will be able to acquire innovations, and the size of the community's language can expand. But the (now faster) innate learning speed and the critical period still put a limit on the possible size of the community's language.

In what follows, computational simulations of these biocultural mechanisms will be described in detail. The key propositions to be established (and which are therefore **not** built in as assumptions) are:

1. There is an evolutionary mechanism locating the age of the end of the critical period at around puberty.

2. The size of the language of a community adjusts itself to coincide with the maximum that can be acquired within the critical period, given the speed at which children can acquire it.

It goes without arguing here that there **is** a critical period for language acquisition, and that it coincides roughly with puberty. Although the critical period for language acquisition in humans varies across individuals, and the mean age of steepest decline in language acquisition capacity is no doubt not exactly at the mean age of puberty (see Long, 1990), we nevertheless believe that the approximate correlation is close enough to warrant exploring possible explanatory mechanisms. This is analogous to noting that the typical cycle of menstruation coincides strikingly with phases of the moon, although in individual women the menstrual cycle may vary and the average period is no doubt not exactly one lunar month

(Knight, 1991). In such cases, it it worth seeing whether a proposed explanatory mechanism can withstand criticism.

We will give some discussion later of what might count as the "size" of a language.

2 Previous work: conclusions and unresolved issues.

2.1 Hurford (1991)

In a previous paper (Hurford, 1991), a mechanism was shown whereby the critical period evolves to fit in the period of life before puberty. Assuming that possession of language confers fitness, it is evolutionarily advantageous to acquire one's whole language before the onset of one's reproductive years. In simulations described in that work, however, the maximum amount of language that could be learnt was a postulated value, externally imposed, and not subject to change.

A space of possible genomes was defined, providing for a range of 'language acquisition profiles', including many deliberately quite implausible ones. A given language acquisition profile would specify how much of the 'whole' language (whose size was fixed) the organism could acquire at each stage of its life. Thus, in principle, an organism could be born with an innate disposition to arrange its lifehistory in such a way that language-acquisition took place near the end of life. Of course, the space of possibilities also allowed for genomes specifying a concentration of language acquisition capacity near the beginning of life.

A simulated population was set up, endowed with random innate language acquisition profiles, in a simulated environment where the language to be acquired was initially set at zero. Modelling the creative acquisition of language in the absence of (good) exemplars, individuals were able to acquire at least some language, even in an environment with the initial zero level of language, but they could never acquire more language at a given stage in life than was permitted by their innate language acquisition profiles. In this way, through a run of the simulation, the language size could grow, but an artificial ceiling of 10 (notional units) was set; no organism could acquire more than 10 'units' of language.

Selective breeding was organized, in such a way that possession of more language conferred reproductive advantage. The populations always evolved to contain only individuals whose language acquisition capacity was concentrated in the period before puberty. In retrospect, this makes obvious sense, as it pays to have all of any reproductively advantageous trait ready for use on time for the period in life when reproduction is possible (i.e. post-puberty). The absence of any language acquisition capacity in post-puberty lifestages in the evolved populations was explained, not as an adaptation, but as due to evolutionary mutation pressure. While there is selective pressure to maintain language acquisition capacity in early life, there is no such pressure in later life, because it can be assumed that language has already been acquired by then[4].

But there was a biassing factor in the arithmetic of the notional numbers used to define language size and the language acquisition profiles. The maximum size of a possible language (the ceiling value mentioned above) was set within the limits of what it was possible to acquire within the lifestages before puberty. In principle a language could be so big that it takes a whole lifetime to learn it, and this possibility was excluded in the simulations of Hurford

[4]As Morten Christiansen (1994:147) has crisply noted, "Hurford (1991) also finds the critical period to be a spandrel."

(1991). If the maximum notional language size had been permitted to vary above the limit of what could be acquired before puberty, then it is still likely that a critical period would have emerged, but a much longer one, culminating, for example, say, around middle age.

Couched in the terminology of innate language acquisition profiles, the study in Hurford (1991) might have been identified with a particularly nativist view of language acquisition. In fact, however, nothing was argued in that study about domain specificity; the same evolutionary mechanism could equally apply to the acquisition of any advantageous skill. And that study also said nothing about any maturational factors that might also be involved in language acquisition. This a matter taken up by Elman's article.

2.2 Elman (1993)

Elman's approach is not evolutionary, but ontogenetic. He shows, with elegant experiments with the training schedules of neural nets, that an organism whose syntax acquisition resources 'start small' can successfully acquire a language with the human-like characteristics of nested long distance dependencies. A network without the maturational 'starting small' strategy could not be trained to acquire such a language. The 'starting small' strategy involves concentrating at first only on very short stretches of input, and gradually expanding the window of attention to longer and longer stretches. The network learns basic facts about the input language, such as Noun/Verb classification, before it even sees (and risks being confused by) evidence for more complex aspects of linguistic organization, such as long distance dependencies.

The key to successful learning, in Elman's study, is the maturational schedule whereby the window of attention (which he calls 'working memory') gradually expands. The implication is that this is what happens in children, and that the schedule of expansion is nicely timed to allow just the right amount of time at each stage for the acquisition of enough language to provide a firm foundation for the next stage in learning. Elman tried various schedules and found one that worked, given the pre-defined learning task. Interestingly, it needed a longer period at the first 'narrow window' stage than at later stages. Elman's explanation for the existence of a critical period for language acquisition relies on the maturational schedule being built in to development. If, as in the tragic case of Genie, a child is given no language input in early life, the maturational schedule expanding the size of the relevant attention span carries on regardless. A person who is only exposed to language abnormally late will not have the advantage of 'starting small', and will not be able to learn the language.

It might be thought that the two explanations of the critical period just presented (Hurford, 1991 and Elman, 1993) were incompatible rivals, but they are not. Hurford gives an evolutionary mechanism by which (with certain assumptions about language size) the critical period ends near a specific lifehistory event (puberty), but mentions no psycholinguistic mechanisms relating to the learning process or to the structure of the language being learnt. Elman's study, on the other hand, makes specific and interesting proposals about the relation of the learning process to linguistic structure, but does not deal at all with the lifehistory timing of the maturational schedule that he invokes, and proposes no evolutionary mechanism. The two accounts are complementary. It is suggested in Hurford (1998) that the two accounts could in fact be combined; evolutionary processes giving rise to the kind of maturational schedule that Elman describes could be sought, and one could attempt to show how such a schedule becomes calibrated to relate to crucial lifehistory events, such as puberty. This is what the next study does.

2.3 Kirby and Hurford (1997)

This study assumes the relationship between incremental learning of language and the steady expansion of a resource, such as Elman's 'working memory'. That it, is is assumed that for each stage of language learning to be successful, the previous stage must have been successfully completed, and the resource drawn on by the learning mechanism, the 'working memory', must be expanded one notch to the right 'size' for the next stage to begin. The study examines the evolutionary processes by which this expansion of resource could be programmed into development. It uses computer simulations of evolution.

The simulations defined a population with lifehistory characteristics of birth, puberty and death; reproduction was only possible after puberty, and the probability of thus passing on ones genes was related to the amount of language the prospective parent had acquired. In the simulations, evolution was given two different ways of expressing the expansion of working memory: as a function of chronological age, or as a function of exposure to input (or as some mixture of these two). The structure of the simulated genome allowed in principle for the construction of various phenotypes. In one possible phenotype, the maturation of language-relevant working memory was tied to chronological age; in another possible phenotype, the expansion of working memory was triggered by language input; and various mixed possibilities were also allowed. One such (hypothetical) mixed strategy, for example, would be an individual whose working memory expansion was triggered by chronological age early in life, but by exposure to language later in life.

What emerged from these simulations was a mixed genome, in which expansion of working memory is triggered by exposure to language early in life, and by chronological age later in life. In other words, with such a genome, some slight delay in exposure to language would not be disastrous; the expansion of working memory could wait around for a while, during childhood, and it would be possible for a child starting to learn language a little late to catch up. But too great a delay would be disastrous, as, sooner or later, the age-determined expansion of working memory would kick in and take working memory to a size where the possibility of 'starting small' had disappeared. The evolved genomes in this study all encoded a switch from input-sensitive expansion of the resource to age-related expansion. In some conditions, this switch even happened to coincide with puberty.

But the switch from input-sensitivity to age-related growth did not coincide with puberty in all the simulations. Instead, it coincided with a time in lifehistory which was a construct of various parameters chosen for the running of the simulations. In all these simulations, as with the earlier study (Hurford 1991), an arbitrarily chosen 'size' for language was fixed. Also fixed was a measure of 'input quality', simulating a degree of unreliability in the environment. This variable could be set, for example, at 50%, expressing the fact that a learner only had a 50% chance of actually getting any language input at a particular lifestage. The actual number of lifestages taken to acquire the full language was a function of language size and input quality. Naturally, individuals exposed to more language, or exposed to the same amount of language less reliably, learnt it more slowly. For example, to learn a language of notional size 10 at 50% input quality would take 20 lifestages; to learn a language of notional size 9 at input quality 75% would take 12 (= 9/0.75) lifestages. Systematic variation of these parameters, independent of the age set for puberty, showed that the critical period emerged at the age of typical acquisition of full language, and not necessarily always at puberty.

A clear graphic way of depicting a critical period effect is with a graph of final attainment against length of deprivation. Figure 1 is one such graph.

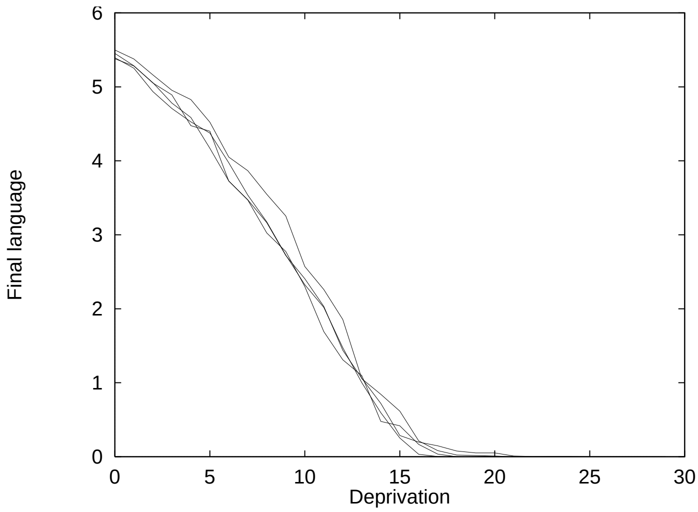

Figure 1: How language competence declines with age of initial exposure to a language.

This figure shows the results of four different runs of the simulations. In these runs, the language size was set at 7 units, and the input quality was set at 0.5 (50%). With this language size and this input quality, the expectation is that with normal exposure one will acquire one's whole language by the end of the 14th lifestage. The four curves give the predicted results of simulated Genie-experiments on individuals with the genome that evolved in the main (non-deprived-environment) simulations. They show that with deprivation of language input up to the end of the 14th lifestage, the expected attainment is close to zero. What is important to note for our purposes here is that these four runs were conducted with puberty set at very different lifestages, 5, 10, 15 and 20; Figure 1 shows no effect of puberty on the expected end of the critical period.

In K&H's study, the lifestage at which normal language acquisition is complete is a construct derivative of language size; in the simulations, language size is simply given, and the critical period evolves to coincide with this construct. This is an example of the Baldwin Effect (Baldwin, 1896; Hinton and Nowlan, 1987), by which the presence of learning can in fact guide phylogenetic evolution. In an environment stable over millennia, some constant aspect of the learning process becomes innate. In this case, the constant aspect of the learning process which gets biologized is its timing in relation to lifehistory.

3 How is the size of a language fixed?

What comes out of these previous studies is that the evolutionary emergence of a critical period can be modelled, but so far all models suffer from two related defects, namely:

1. No strongly convincing mechanism is shown linking the critical period to puberty without building in a fixed ('God-given', as it were) quantity for language size. Hurford (1991) does show that, given a certain assumption about language size, the critical period evolves to fit within puberty. But on the other hand, Kirby & Hurford (1997) show

that if this fixed quantity is varied, then the critical period can be systematically 'made' to evolve at lifestages **other than** puberty.

2. No suggestions are made at all regarding how the size of a language can become fixed, or evolve.

Languages exist in communities. The language acquired by children of one generation is the language that they, in their turn, transmit to the children of the next generation. Transmission is not always perfect, of course, as languages evolve over the course of history. The approach we take here is that the **size** of a language is also an aspect that is held in the community, and subject to the slight changes with each generation that accumulate to constitute the history of the language. As with other features of a language, its size has to pass through the filter of language-acquirers each generation. If (hypothetically) a language were too big for the innate L.A.D. to cope with, its size would not be preserved into the next generation, just as a (hypothetical) language containing an impossibly complex construction would not be faithfully transmitted to the next generation. Indeed, the size of a language can be thought of as the total of all its complexities.

Linguists usually believe that no language is, overall, any more complex than any other. Admittedly, no metric exists to test this common assumption, but it is one that we will accept here. But while languages may not vary in their **overall** complexity (read "size"), it is well known that languages vary in the complexity of their subsystems. Some languages, for example, have no numeral system at all, but may have complex kinship naming systems; a language may have a very complex case system, but a relatively simple aspect system; or a rich lexical tone system with simple phonotactics; or complex constraints on word order but no case system; and so on. We claim that a language in which all the most complex subsystems were put together would be unviable, for a combination of psychological and social reasons. Imagine a language with: rich Arabic template-type derivational morphology; a set of noun suffixes as detailed as those of Finnish; vowel harmony as in Turkish; a lexical tone system as rich as that of Cantonese; consonantal phonotactics as permissive as Russian; as many vowel distinctions as British English; the click phoneme inventory of a Khoisan language, such as Nama; pharyngeals, uvulars, velars, palatals, palato-alveolars, alveolo-palatals, alveolars, dentals, labiodentals, and bilabials; three degrees of voicing; a combined aspect/mood system as complex as that of Navaho; both head-marking and dependent-marking; a rich system of deixis as in Inuit; singular, dual, paucal and plural number; as many noun classes as a Bantu language; four varieties of past tense; definite and indefinite conjugations as in Hungarian; mixed ergativity; switch reference; .... That would be an impossibly large language; the reasons for its impossibility would be a combination of the social and the psychological. It would put a heavy burden on the acquirer, and the communicative load of much of its complexity would be relatively low, as messages would be massively redundant.

The sense in which, we claim, the size of a language is "held in the community" does not rely on an ontological category such as (Durkheimian) social fact. The size of a language can be taken to be (depending on ones purpose) the average, or the maximum, of the language competences stored in the heads of individuals in the community and expressed by them in behaviour. In a similar way, one could talk about the average loudness, or the average pitch, or the average speed in syllables per second, of the voices of individuals in a crowd. The set of language competences acquired by the individuals in one generation determines their linguistic output, which is the basis of the language acquisition by the next generation.

In the next section, we will describe the implementation of a model in which language size is not imposed by the programmer, but is determined by the (simulated) biological factors involved in language acquisition and the (simulated) social/historical factors involved in the constant cycle of acquisition and re-transmission of language to the next generation. In this model, language size is constantly adjusted during the history of a community. Given a particular innate speed of language acquisition, and a socially presented language of a particular size, there is a typical age for complete language acquisition. Previous results, summarized above (Kirby & Hurford, 1997), show that the biologically-given critical period for language acquisition tracks the age of complete language-acquisition. The outcome is that the end of the critical period approximates to the age of puberty.

4 New simulations

The explanatory scheme for this work is given in Figure 2.

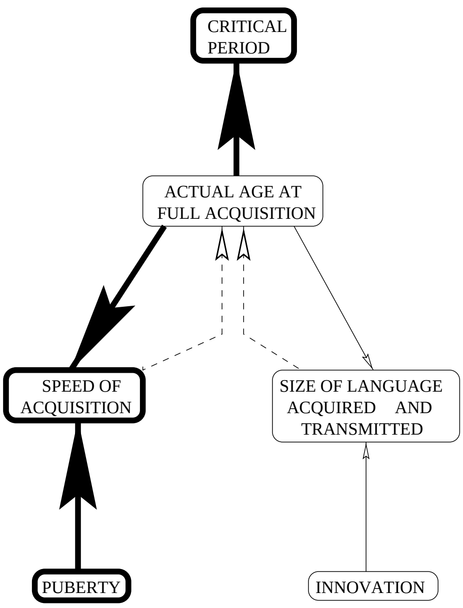

Figure 2: Co-evolution mechanism for speed of acquisition and language size

In this diagram[5]:

- the heavily outlined boxes represent genetically encoded properties of human language acquirers;

[5]One should not get too carried away by "box and arrow" diagrams; they are expository schemes intended to clarify the issues involved. There are alternative ways of schematizing any domain with boxes and arrows. We hope our way illuminates the problem area.

- the lightly outlined boxes represent social processes or constructs, which occur or exist in the social communities of language users;
- the heavy arrows represent phylogenetic evolutionary processes, determined by natural selection;
- the light arrows represent social processes, operating within the social communities of language users;
- the dashed arrows represent a logical definitional relationship.

Two of the boxes in Figure 2 have no arrows entering them. These two, puberty and innovation, are taken as given in this study, and are not explained. But a few words about each are in order.

Puberty is the biological lifestage after which an organism is capable of reproduction. We assume that the age of puberty evolved long ago in our prehominid ancestors, well before language entered the picture. A given age of puberty is the biological foundation on which our explanations rest.

Innovation is a social force (operating within biological constraints, of course) by which individuals may sporadically introduce new social constructs into their communities. In the linguistic domain, a tiny example of innovation would be the coining of a new word; a more substantial innovation would be the "invention" of a new grammatical construction. We assume that innovation is possible, but rare. Without innovation by individuals there could be no mechanism by which the social constructs we know as languages could have arisen. An innovation, though originating with an individual, needs to be adopted by the community as a whole. It is not necessary for our argument that innovations adopted by the community be functional, but it seems natural to assume that they are. We also assume that the effect of an innovation is to increase the size of a language, that is, in effect, to place a greater burden on subsequent acquirers of the community's historically evolving language.

The remaining four boxes in Figure 2 all have arrows entering them. The arrows entering a box all, in some sense, explain its contents. We will deal with each explanandum box, or factor, in turn, in separate subsections. The first two factors discussed, age at acquisition and the critical period, can be dealt with quite briefly. Our main focus will be on the mechanisms explaining the other two factors, speed of acquisition and size of language.

4.1 Actual age at acquisition

By definition, the variables **size of language acquired** and **speed of acquisition** give the **actual age at full acquisition**, at any stage in history. This is not an empirical claim, but merely a tautology; just as in a physical situation $Time = \frac{Distance}{Speed}$, so here $AcquisitionAge = \frac{Size}{Speed}$. This is the most elementary part of our story, with which there can be no argument, with the one obvious reservation that if $\frac{Size}{Speed}$ is greater than $AgeAtDeath$, then $AcquisitionAge = AgeAtDeath$.

Our model will show the co-evolution of the two variables, size of language and speed of acquisition. The factor mediating their co-evolution is their joint outcome, age at acquisition. They co-evolve in such a way as to keep age at acquisition close to puberty, as will be shown. The dotted arrows in Fig. 2 show this relationship, whereby speed of acquisition and size of language simply **define** the age by which language is acquired.

4.2 Critical period

The connection between age at acquisition and the critical period is a result argued for in detail in Kirby and Hurford (1997). We have summarized the work leading to this result in a previous section and we simply assume it here. The heavy arrow leading into the Critical Period box in Figure 2 represents this correlation between age at acquisition and the critical period. To accept this correlation is not to build in a solution to our problem. On the contrary, it is precisely this correlation that seems to **be** our problem, as K&H98 show that the critical period can be 'made' to vary independently of puberty, given various settings of language size and the average rate at which it is acquired.

4.3 Speed of acquisition

The heavy arrows into the Speed of Acquisition box in Figure 2 show an evolutionary effect of puberty and actual age of language acquisition. There is phylogenetic evolutionary pressure to adapt speed of acquisition so that actual completion of full acquisition, for a language of a given size, occurs by puberty. This gives an evolved lower bound on speed of acquisition. We now describe simulations modelling this effect.

For a simulated population, facts including the following were defined:

Size of population	50 individuals
Age of death	40 lifestages
Age of puberty	variously 6.5, 12.5 and 18.5 lifestages
Initial language maximum size	various, from 10 to 1000 'units'
Initial speed of acquisition	various, from 0 to 5 units per lifestage

50 is a large enough population size for this kind of simulation; larger numbers would not change the outcome. The scale of lifestages by which death and puberty are defined was chosen to be reminiscent of years. Puberty, an immutable genetic property of individuals, was set at various ages, so that an effect of the age of acquisition relating to puberty could be demonstrated. The puberty numbers are the halfway points between the last lifestage at which an individual could not reproduce and the first lifestage at which s/he could. The initial language maximum size was simply an arbitrary number used to get the simulations started. The maximum language size changed constantly during the course of the simulations, determined by the greatest size of language acquired by any adult in the population, at any given cycle in the simulation. The speed of acquisition was a genetically encoded property of each individual, inherited from a parent, and subject to sporadic mutation in newborns after initialization of the simulation; this property was represented as a single number.

The simulations went through cycles, outlined briefly as follows (some details will be elaborated later):

1. Individuals at lifestage 40 die.

2. A subset (usually 25%) of the adult population (i.e. those past puberty) are selected as prospective parents for the next generation. This selection is made on the basis of how much language the individuals have acquired at the time.

3. Parents are selected at random from this elite subset and breed sexually, producing enough new individuals to keep the population constant. At this point, there is a small

chance of a random mutation, so that speed of acquisition in the genes of a newborn may differ from that of the parents.

4. Language acquisition: Any individual whose language is not yet at the maximum established by the community acquires as much language as his genetically given speed of acquisition will allow in one lifestage, up to the limit of the community maximum.

5. All individuals advance one lifestage.

To show how language size is affected during the simulation, we will give an example with some specific numbers. Take a case where the initial (arbitrary) language size is set at 1000 'units', and the individuals are genetically endowed with an acquisition speed of 5 units per lifestage. An individual newborn at the beginning of a run has 40 lifestages ahead, during each of which just 5 units of language can be acquired. This individual will live for 40 lifestages, acquiring language throughout life (in this artificial initial situation). At the end of this individual's life, she will have acquired a language of size 200 (5×40), and 200 now becomes the maximum language size in the community. No subsequent individual can acquire any more than 200 units of language.

Alternatively, take the case where the initial (arbitrary) language maximum size is set relatively low, say at 10 units, with the same initial innate speed of acquisition in all individuals, 5 units per lifestage. In this case, all individuals will acquire the full language by the second lifestage. In this version of the simulation, the maximum language size will remain fixed at 10 for the rest of the run.

This fixing of the maximum community language size happens rapidly, of course, always within the first 40 cycles of a run; it is still something of an artefact, being either the initial arbitrary language size or the product of the initial genetically given speed of acquisition and age at death, whichever is the lesser. In the simulations described in this subsection, language size gets rapidly fixed in this manner in the very first cycles of a run and does not change further during a long run. In the next subsection, we will show how the introduction of an 'innovation' factor allows language size to change constantly during a simulation. But now we focus on how the speed of acquisition may change to affect age at acquisition.

As the simulation progresses, mutants arise whose innate speed of acquisition is either faster or slower than that of the rest of the population. Assuming the community maximum language size is still such that it takes nonmutant individuals a whole lifetime to acquire it, a mutant slower acquirer will get to the end of life without acquiring the whole language. At any stage in life such a mutant slow acquirer will possess less language than his lifestage cohorts, and will be relatively disadvantaged in the competition for selection for parenthood. (If, most improbably, such a slow-acquirer mutant gene were to wholly invade the population, then the community maximum language size would, of course, drop to the size acquirable within a lifetime by such a population of slow learners.) By contrast, a mutant fast acquirer will acquire the whole community language before the last lifestage, and throughout life be at a relative advantage in the competition to be a parent. Such faster mutant genes are likely to spread through the population. But, as there is (so far) no mechanism for increasing the size of the community's language, the mean age at which the whole language is acquired will be reduced.

The competition to be a parent takes place among those who have passed puberty. In that adult population, individuals who have relatively more language will be at an advantage. Thus there is selection pressure in favour of genes which will speed up language acquisition to acquire the whole language before puberty. Speed of acquisition adapts to be at least fast enough to get the job done before puberty. There is nothing to prevent speed of acquisition from becoming (by mutation) even faster, so that language acquisition could be complete well before puberty, but there is no pressure on pre-puberty individuals to learn any faster than they 'need' in order to be finished in time for their entry into the mating competition.

These effects, which occurred in our simulations, are shown in Figures 3 - 5 .

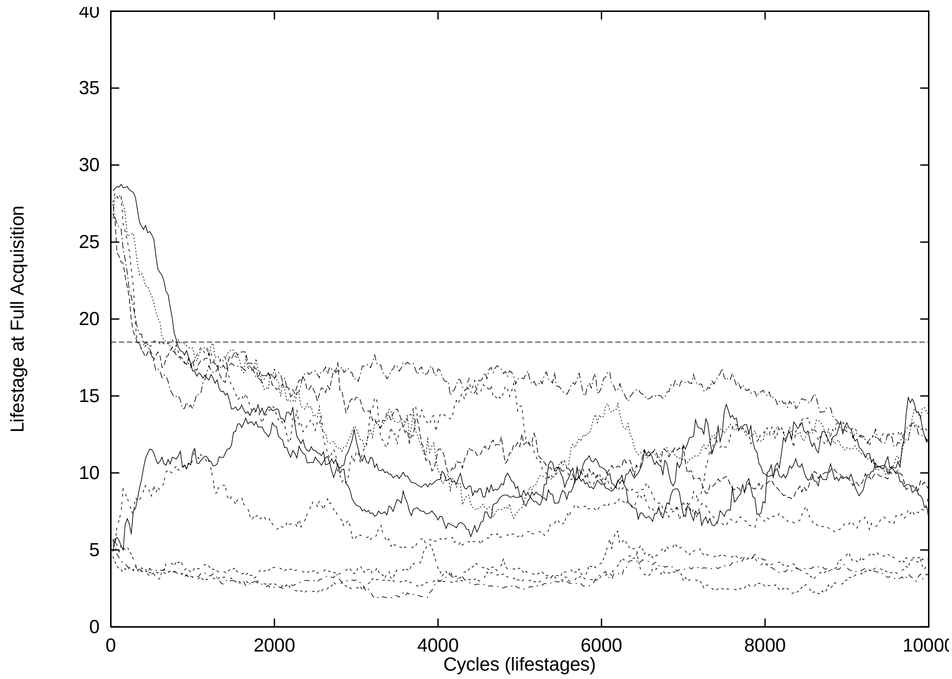

Figure 3: Age at acquisition settles and walks randomly below puberty (18.5).

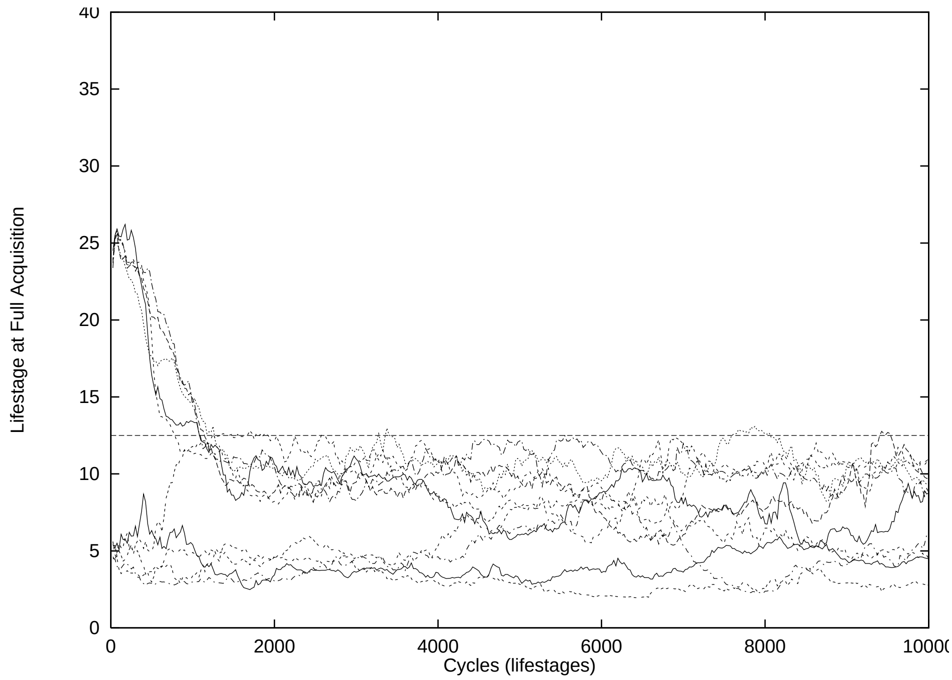

Figure 4: Age at acquisition settles and walks randomly below puberty (12.5).

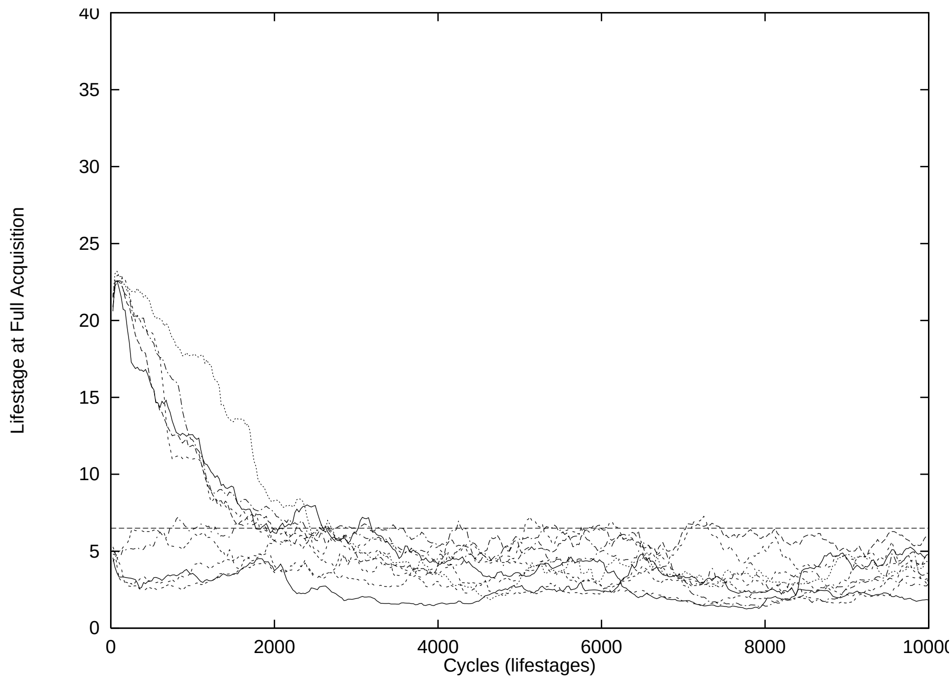

Figure 5: Age at acquisition settles and walks randomly below puberty (6.5).

Each of those figures shows the population's average lifestage at complete language acquisition, as it evolved through 10,000 cycles of a simulation. Each figure shows 10 separate runs, 5 runs with an initial high language size and 5 runs with an initial low language size. Each figure shows runs with puberty set at a different level (18.5, 12.5, 6.5). In all cases, it can be seen that where language size started high, the average lifestage at full language acquisition was well above puberty, but quickly dropped, by selection pressure on speed of acquisition, to below puberty. Once the average lifestage of full language acquisition was below puberty (or if it had started below puberty) it simply wandered randomly (due to the effects of random mutations), there being no pressure on pre-puberty individuals to acquire language any faster.

So far, then, we have half of the story. We have shown (a model of) speed of language acquisition evolving in such a way that the average age of acquisition is below puberty. And we have shown how language size can be stored in the community, acquired and re-transmitted by successive generations. So far the only demonstrated effect on language size is the relatively trivial constraint that it cannot be greater than any individual can learn within a lifetime. But we have shown how this upper bound on language size plays a role in the co-evolutionary process whereby language size and speed of learning are together adjusted so that language is learnt by puberty. We have not yet shown what prevents the average age of language acquisition from wandering well below the age of puberty, as it does in many of the runs shown in Figures 3 - 5. We will fill these gaps in the next subsection.

4.4 Size of language acquired and transmitted

We postulate that innovative potential, fuelled by social and communicative considerations, provides pressure to increase the size of language acquired. In a second version of our simulations, we added this factor of innovation. All other conditions remained as before, but now an individual who had acquired the full language of the community was permitted, at very rare intervals, to add to the language by a small increment. In this way the language of the community actually expanded.

We now have two factors at work: a biological factor of mutation affecting speed of acquisition; and a social factor of innovation periodically increasing the size of the language to be acquired. Fast learner mutants are advantaged, up until the point where language is acquired by puberty, as we have seen. Consider a population with a uniform evolved speed of acquisition such that the community language is acquired just by puberty. Now, if the language size increases, the age of language acquisition, for all individuals, will increase to an age above puberty. A new mutant, with an even faster speed of language acquisition, will now have an advantage, and this mutation will tend to spread through the population.

In our simulations, we made innovation a constant force, exerting constant upward pressure on language size. Thus, for a population with an evolved speed of acquisition such that all individuals acquire the language well before puberty, this upward pressure on language size will tend to make the age at acquisition rise (because there is more to acquire). So long as the age at acquisition stays below puberty, biological selection pressure will not react to this, by speeding up acquisition, as only adults are involved in the selection process. Biological selection pressure to speed up language acquisition only bites where the language size has increased to such a size that it cannot be completely acquired by puberty.

The results of these simulations are shown in Figures 6 - 8, again with puberty set at various ages – 18.5, 12.5 and 6.5.

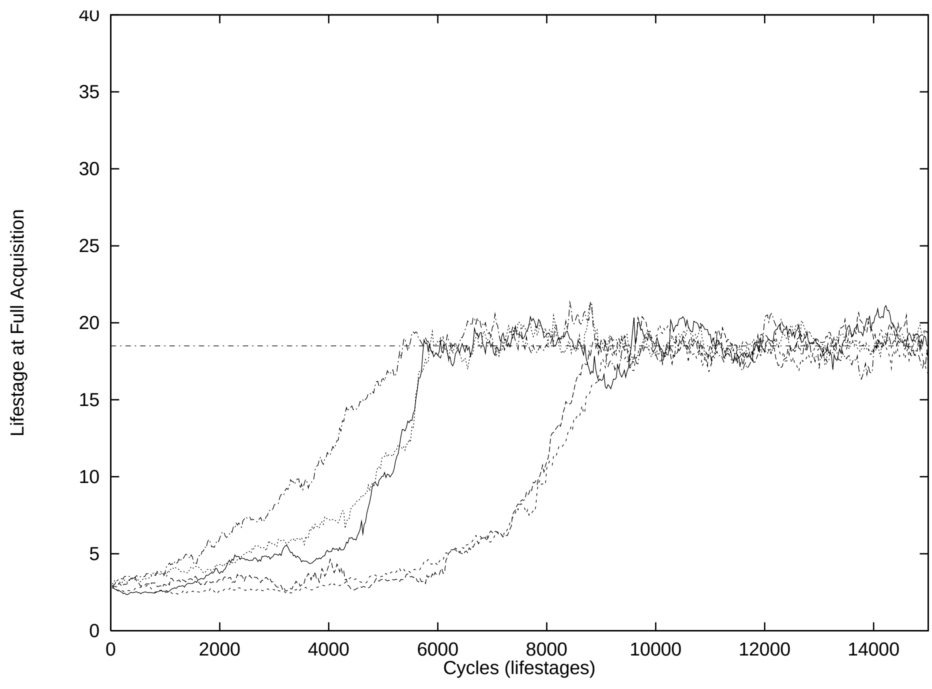

Figure 6: Language size expands and age at acquisition settles at puberty (18.5).

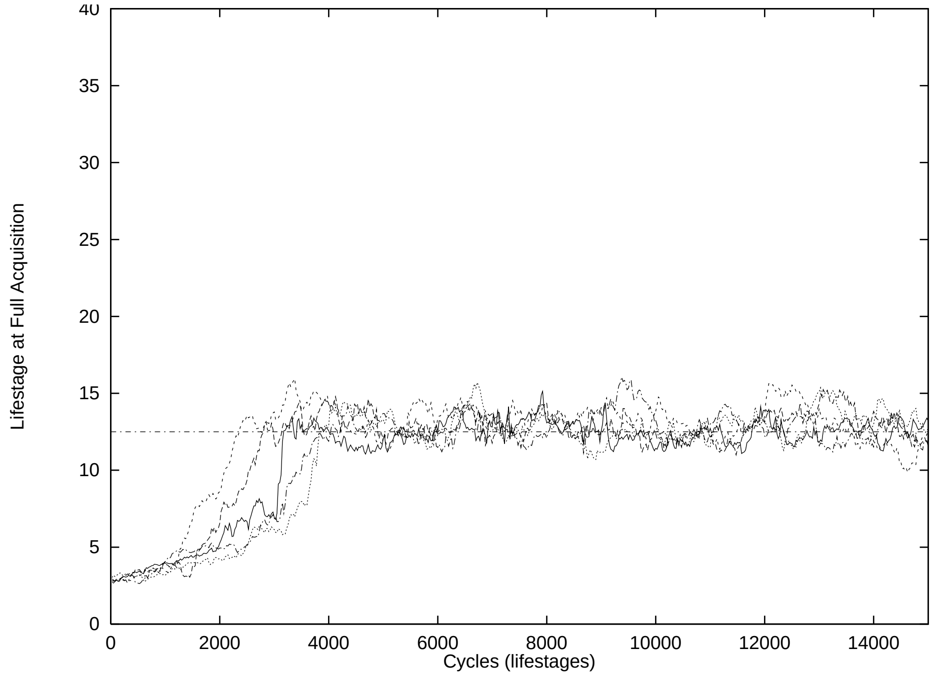

Figure 7: Language size expands and age at acquisition settles at puberty (12.5).

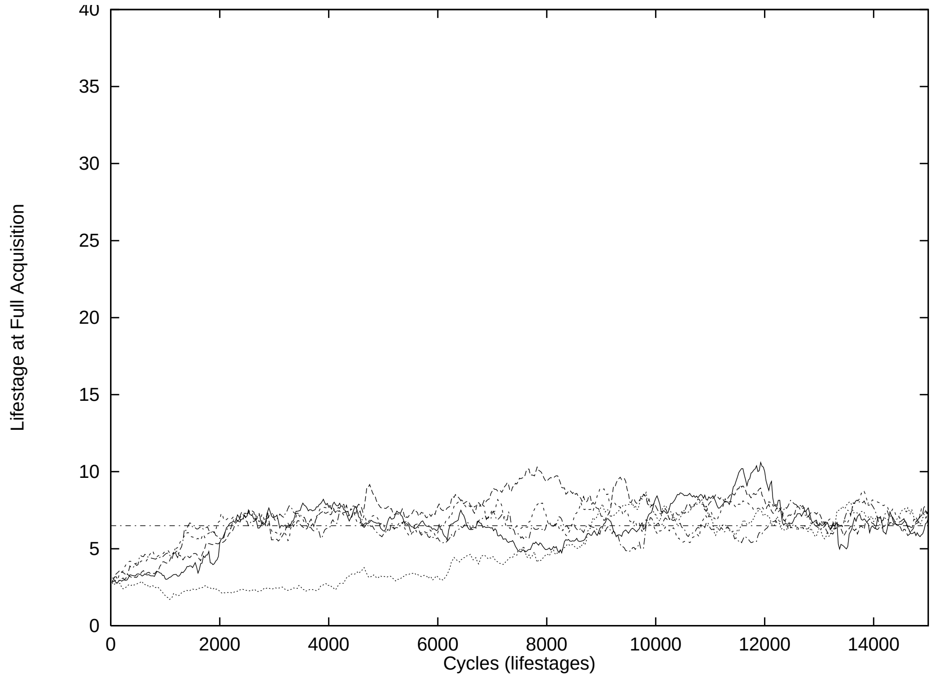

Figure 8: Language size expands and age at acquisition settles at puberty (6.5).

As before, these figures plot average age at acquisition against 'time' (in cycles of the simulation). These graphs each show the results of five separate runs. In these runs, the initial language size was set low, so that, initially, age at full acquisition was well below puberty. There is a clear tendency for age of acquisition initially to rise, due to the slow but constant force of innovation, which enlarges the target language of acquirers. During this initial phase, the genetically given speed of acquisition is under no selection pressure. When the language size reaches the point where $\frac{Size}{Speed} = Puberty$, biological selection begins to bite, and an 'arms race' between language size and speed of acquisition begins, keeping age of acquisition near the age of puberty.

Figures 6 - 8 should be compared with the previous figures, 3 - 5. The random walk below puberty in the earlier figures is replaced in the later figures by a balancing act in which counterposed forces (biological speed of acquisition and social innovation) keep age of full language acquisition near puberty.

A final figure, Figure 9, plots speed, language size and acquisition age, on a log scale, taking results from a single run.

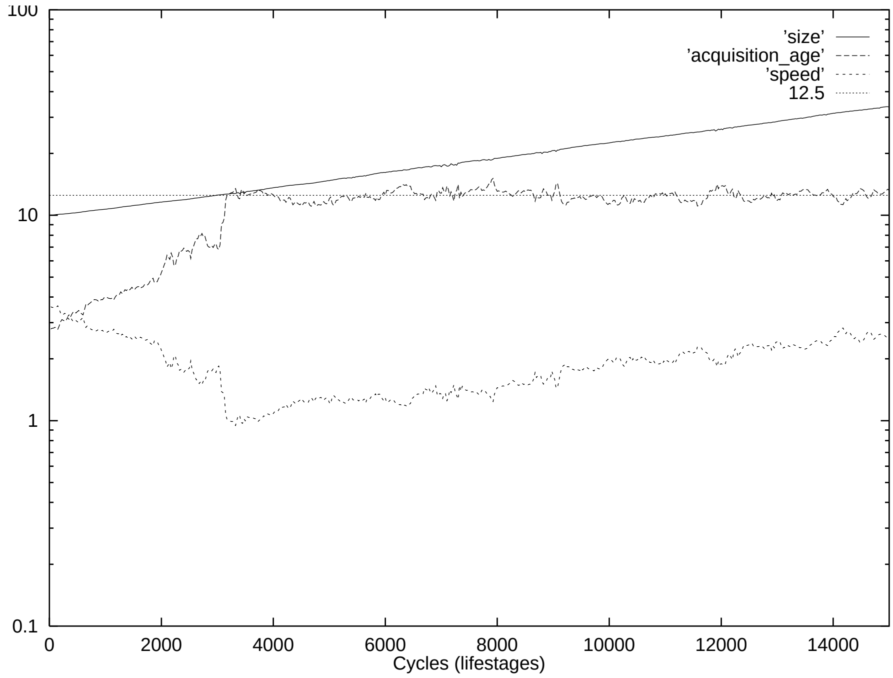

Figure 9: Size increases: when age at acquisition hits puberty, speed also increases (log scale).

The x axis of this graph gives 'time' (in cycles of the simulation). The y axis actually represents several different kinds of quantities, measured in different units. The numbers on the y axis (0.1, 1, 10, 100) express age, in lifestages; the straight line at 12.5 expresses the age of puberty on this scale; the curve for acquisition age, which rises to about 12.5 and then flattens out, is also drawn to this numerical scale. Innovation increases language size exponentially, as shown by the straight diagonal line sloping upwards from 10; this line expresses the size of the language in notional 'units' (not lifestages). The lower curve plots speed of acquisition in terms of units per lifestage.

The significant point to note in Figure 9 is the relationship between the curves for speed and age at acquisition. The curve for speed starts in this case by wandering downward; simultaneously, the curve for age at acquisition moves upward, due both to the (intentionally) slowly increasing language size and the (so far, randomly) decreasing speed of acquisition. At exactly the point where the curve for age at acquisition moves above the puberty line at 12.5 lifestages (at around 3,200 cycles), the curve for speed of acquisition responds by turning upwards, and from then on it follows an upward course essentially parallel to the curve of language size. Of these two parallel curves, size and speed, one is straight (on this log scale) and the other is wiggly. This is because, in this simulation, innovation proceeded with a very regular beat, and always in the same upward direction; the biological mutations for speed were also introduced at completely regular intervals, but the direction of mutation (acceleration or deceleration) was chosen randomly, and the progress of any particular mutant gene through the population was affected by the random processes involved in the simulated sexual reproduction. Somewhat different treatments could have been adopted, but there is no reason to suppose that the outcome would have been significantly different.

There is a coincidence in Figure 9 that is **not** significant. The fact that the acquisition-age and speed curves change direction at more or less exactly the time when the size curve crosses the puberty line is not at all significant. Size and puberty are 'measured' differently: size in 'units', puberty in lifestages. This is an accident of the scaling of the size numbers. In similar diagrams from other runs, this coincidence does not occur.

To summarize this section, and express our central point:

- Language size evolves socially, by innovation.
- Speed of acquisition evolves biologically by natural selection.
- The two evolutionary processes coordinate in such a way that

$$\frac{Size}{Speed} \rightsquigarrow Puberty.$$

The implications for SLA of the models developed in Kirby and Hurford (1997) and in the present paper can be summarized as follows. The effect of applying the language-acquisition resources to a first language, early in life, is to build up a store of knowledge about the language acquired. The knowledge of a first language might now be considered an additional resource, which can be deployed in the acquisition of a second language. It is, however, theoretically clarifying to distinguish between two different kinds of "resource", that is, between the facility for processing data and turning it into (more) knowledge, on the one hand, and actual knowledge of language, on the other. The model of K&H97 showed the likely evolution of a certain kind of genetic control over the processing facility subserving language acquisition. This form of control dictates that exposure to language early in life alters the parameters of this facility, making it progressively more adapted to acquisition of the higher reaches of linguistic knowledge, but probably, following Elman's and Newport's ideas on 'starting small' and 'less is more', less well adapted to the demands of beginning a new language. It is an open, and empirical, question to what extent the acquired knowledge of a first language can substitute for the lost processing facility. To the extent that a second language is similar to the first language acquired, the amount of knowledge to be gained by the L2 acquirer is diminished. The adult second language acquirer has less to learn, but has impoverished processing resources with which to learn.

The main consequence of the model developed in the present paper for SLA, as for first language acquisition, concerns the life-history timing of the switch away from the early facility with language input. It shows how a co-evolutionary process can be expected to define the age at which one becomes an "adult language learner" to coincide closely with the age at which one becomes reproductively adult. Onset of sexual maturity is not logically connected with language acquisition capacity; this paper has shown the possibility of an evolutionary connection.

5 In conclusion

It might be thought that the issues discussed here are something of a specialism. The notions of language size and speed of acquisition are barely discussed in such general terms as here. But, although we have been utterly unspecific about the details of exactly what structures children acquire, and at what stages in development, our model deals with the fundamental dimensions of human language. The introductory literature on both generative grammar and language acquisition always stresses the impressive richness (complexity, size) of the competence acquired, and the impressive speed of acquisition. Naturally, speed and size are never quantified – just asserted to be impressive, as indeed they are. This raw speed and size are what language acquisition is fundamentally all about; it is what makes the subject so important. Detailed studies of language acquisition never engage in discussion of raw speed and size, just as practising terrestrial mapmakers don't preface their maps with discussion of the fundamental dimensions North-South and East-West. The geographic dimensions are given by the nature of the earth's rotation controlled by a balance of physical forces. We have proposed that the linguistic dimensions of size and speed evolved to maintain a balance between natural selection for greater advantage by breeding age and social construction of systems with greater representational power.

The model we have presented is clearly idealized in a number of ways, as are all such models. We aimed to capture the mechanism central to language evolution. This model differs from a common early type of model of language evolution (e.g. Hurford (1989, 1991)). Early models treated the evolution of the language capacity as a biological phenomenon essentially unaffected by any changes in the cultural environment created by the community of language users. Clearly, such interactions of the biological and social have to be treated with caution. For one thing, it is essential to try to ascertain whether the rates at which the biological and social processes happen can be coordinated in the ways suggested. At present, our notion of language size is so simple, and our knowledge of the genetic contribution to speed of learning so incomplete, that this study remains highly speculative.

Venturing to speculate even further, our kind of model of gene-language coevolution might conceivably hold some potential for explaining the spectacularly fast increase in brain size between the australopithecines and ourselves. The essential idea would be that by some potential for culturo-linguistic innovation, a new, non-physical type of environment (rudimentary languages) arose, susceptible to change at a much faster rate than the physical environment, which had previously been paramount in steering the course of evolution. No correlation between brain size and language size is likely to be straightforward, but it would be surprising if there were no correlation at all.

Culture in the modern era is evolving at a much faster rate than at any previous stage. The Twentieth Century, in particular, has seen dramatic changes in the human (largely human-made) environment. Even if the rates of biological and social evolution were ever coordinated in the way our model proposes, it is quite possible that they have become uncoordinated in the modern era. If we include acquisition of the conventions of written language as part of language acquisition, modern language acquisition takes longer than the time to puberty (see Miller and Weinert (1998)). This may be an instance of cultural evolution racing along so fast that it is now impossible for biological evolution to adapt.

Our model depicts a self-feeding spiral of language size responding to increases in speed of acquisition, and speed of acquisition in turn responding to increased language size. Where will it end? It can't go on for ever. At some stage, considerations external to the closed system of our model will exert an influence. Language acquisition cannot be speeded up indefinitely; there must be some cost. And languages will not expand indefinitely in size; there must be some principle of diminishing returns for increased size. It is possible that the evolution of human language in the modern age has reached the point where such external factors of cost and benefit, which we have not modelled, have come into play.

6 References

Baldwin , J M 1896 "A New Factor in Evolution", *American Naturalist*, 30:441-451.

Boyd , Robert, and Peter J Richerson, 1985 *Culture and the evolutionary process*, University of Chicago Press, Chicago.

Cavalli-Sforza , L L, and W F Bodmer, 1971 *The Genetics of Human Populations*, W H Freeman and Company, San Francisco.

Christiansen , Morten Hyllekvist, 1994 *Infinite Languages, Finite Minds: Connectionism, Learning and Linguistic Structure*, PhD thesis, University of Edinburgh.

Dawkins , Richard and John R Krebs, 1984 "Animal signals: Mind-reading and manipulation", in *Behavioral Ecology: An Evolutionary Approach* (second edition) edited by John R Krebs and Nicholas B Davies, Blackwell Scientific Publications, Oxford. pp.380-402.

Deacon , Terrence W, 1992 "Brain-Language Coevolution", in *The Evolution of Human Languages* edited by John A Hawkins and Murray Gell-Mann, Proceedings Volume XI, Santa Fe Institute Studies in the Sciences of Complexity, Addison-Wesley Publishing Company, Redwood City California. pp.49-83.

Elman , Jeffrey L 1993 "Learning and Development in Neural Networks: the importance of starting small", *Cognition* 48:71-99.

Hinton , Geoffrey, and S Nowlan, 1987 "How Learning can Guide Evolution" *Complex Systems*, 1:495-502.

Hurford , James R, 1991 "The evolution of the critical period for language acquisition", *Cognition* 40:159-201.

Hurford , James R 1998 "Functional Innateness: Explaining the Critical Period for Language Acquisition", *Proceedings of the UWM Linguistics Symposium on Functionalism and Formalism*, Benjamins, Amsterdam.

Kirby , Simon, and James R Hurford 1997 “The Evolution of Incremental Learning: Language, Development and Critical Periods” submitted for publication; available as University of Edinburgh Occasional Paper in Linguistics EOPL-97-2, at website http://www.ling.ed.ac.uk/˜eopl .

Knight , Chris, 1991 *Blood Relations: Menstruation and the Origins of Culture*, Yale University Press, New Haven.

Lumsden , Charles J, and Edward O Wilson, 1981 *Genes, mind, and culture : the coevolutionary process*, Harvard University Press, Cambridge, Mass.

Miller , Jim and Weinert, Regina, 1998 *Spontaneous Spoken Language*, Oxford University Press, Oxford.