Modelling socio-economic status effects on language development

Date submitted: 11 April 2011

Abstract

Socioeconomic status (SES) is an important environmental predictor of language and cognitive development but the causal pathways by which it operates are unclear. We used a computational model of development to explore the adequacy of manipulations of *environmental information* to simulate SES effects in past-tense acquisition, in a data set provided by Bishop (2005). To our knowledge, this is the first application of computational models of development to SES. The simulations addressed three new challenges: (1) to combine models of development and individual differences in a single framework; (2) to expand modelling to the population level; and (3) to implement both environmental and genetic/intrinsic sources of individual differences. The model succeeded in capturing qualitative patterns of regularity effects in both population performance and the predictive power of SES that were observed in the empirical data. The model suggested that the empirical data are best captured by relatively wider variation in learning abilities and relatively narrow variation in (and good quality of) environmental information. There were shortcomings in its quantitative fit, which are discussed. The model made several novel predictions. One of these was that SES should reliably predict gifted performance in children but not delayed performance. This novel prediction was supported by the Bishop (2005) data set. Finally, the model demonstrated limits on the inferences that can be drawn about developmental mechanisms on the basis of data from individual differences.

Socio-economic status (SES) is a well-known environmental measure that predicts significant individual differences in cognitive and language development (e.g., Hart, Petrill, Deckard & Thompson, 2007; Petrill, Pike, Price & Plomin, 2004), and even some measures of brain function, such as hemispheric specialisation (Raizada, Richards, Meltzoff & Kuhl, 2008). SES effects on neural processes (e.g., as measured by event-related potentials) have been observed even when there are no overt differences in behaviour (e.g., D'Angiulli et al., 2008). The exact definition of SES is controversial, typically involving factors such as parental education, income, and occupation (Hackman & Farah, 2009). SES indexes a range of confounded factors, among them differences in prenatal care, nutrition, exposure to environmental toxins (such as lead), chronic stress, neglect, depression, wealth, transient or chronic social hierarchy effects, cognitive stimulation (including quality of schooling), parenting styles, parental education, social support, and availability of books in the home (Hackman & Farah, 2009; Raizada & Kishiyama, 2010; Stevens, Lauinger & Neville, 2009). It is not clear by which causal route SES effects operate on developmental mechanisms, or the relative weighting across the routes if (as is likely) several routes contribute.

One way to evaluate causal theories is via computational modelling. Implemented models of the developmental process can demonstrate the sufficiency of causal accounts to explain observed data. In this paper, we present a set of simulations investigating how SES influences the course of language development. We focus on one particular aspect, that of past-tense acquisition. To our knowledge, this represents the first attempt to simulate SES effects on language development.

SES and language development

In a recent review, Hackman and Farah (2009) noted how the effect of SES is uneven across cognitive domains, affecting language development and executive functioning more than memory, spatial cognition, and visual cognition. When Farah et al. (2006) created a language composite score of performance on vocabulary and phonological processing tasks in first graders (5-6 year olds), they were able to predict 32% of the behavioural variance using SES. Other reported effect sizes are more modest. For example, Petrill et al. (2004) tested expressive vocabulary and grammatical complexity via parental report in a sample of 6000-8000 twins. SES predicted 3.2% of the variance at age 3 and 3.6% at age 4.

Ginsborg (2006) listed the candidate factors by which SES effects might operate on language development, including parental education, home environment, the relationship between principal caregiver and child, the language environment experienced by the child and the nature of the interaction between mother and child, including the quantity of the speech addressed to the child and the nature of child-directed speech. Although few in number, longitudinal observation studies of the development of spoken language in young children have addressed the role of SES. In a landmark study, Hart and Risley (1995) followed children in 13 professional/managerial families, 23 working class families, and 6 families living on benefits, between the ages of 8 months and 3 years, recording all language produced by the child or available around the child for 1 hour per month. The most salient difference at different SES levels was not the nature of the language spoken to the child but the quantity. Professional families addressed 2100 words to their child in the average hour compared to 600 in the welfare families. In addition, higher SES was associated with a greater incidence of affirmative feedback and lower incidence of prohibitions. The importance of language input is further implicated by the finding of Huttenlocher, Vasilyeva, Cymerman and Levine (2002) that SES did not predict the frequency of complex speech produced by 4- to 5-year-old children once parental input and teacher input were controlled for (see Ginsborg, 2006).

The study of children raised by their biological parents permits the possibility that correlations between parental and child language are mediated genetically (that is, SES is actually a marker of genetic transmission). Environmental variation could therefore play a weaker role than the statistical correlations imply. Research has explored variation in the school environment, specifically correlating gains in language development with the quality of teacher language over the school year. Since the teacher is genetically unrelated to the children, correlations are more likely to represent genuine environmental effects. Data have tended to support the view the environmental variation in language indeed plays a causal role in development (Huttenlocher et al., 2002; Klibanoff et al., 2006; Vasilyeva et al., 2006). For example, the proportion of complex sentences produced by the teacher predicted 18% of the variance in the improvement in performance on a syntax comprehension task over the school year (Huttenlocher et al., 2002).

However, SES effects may be uneven across different parts of language: when Vasilyeva, Waterfall and Huttenlocher (2008) investigated the effects of SES on the emergence of productive syntax between 22 and 42 months of age, they found SES effects in the mastery of complex sentences but not simple sentences. Moreover, the causal pathways of SES may alter across development: when Aikens and Barbarin (2008) examined SES effects on early reading development between 5 and 8 years (i.e., kindergarten to 3rd grade), family characteristics predicted more of the SES-linked variability in initial reading ability, but school and neighbourhood conditions explained more of the SES-linked variability in subsequent improvements in reading.

Despite such complexities, data from the predictive power of SES has been used to draw inferences about developmental mechanisms within language acquisition. For example, Rice, Wexler, and Hershberger (1998) followed the acquisition of the English past tense longitudinally in 21 children with Specific Language Impairment and 20 typically developing children over a 3-year period. Rice and colleagues used growth curves to model individual developmental trajectories for this aspect of morphosyntax. Additionally, they assessed whether SES (as measured by maternal education) accounted for any of the individual variation between the children's growth curves. They found that SES was a non-significant predictor (less than 1% of the variance), with linear and non-linear growth functions explaining the majority of the variance across age. Based on the failure of an environmental variable to explain differences between individuals, Rice and colleagues inferred that the growth (development) was explained by *maturational mechanisms*, where changes in behaviour over time are due to the aging process rather than experience-dependent learning; and that the difference between typically developing children and those with developmental language impairment lies in genetic differences in the specification of the timing of linguistic properties.

In sum, SES has reliable but uneven effects on language development, and one candidate pathway for this influence is the input to the child. The most salient difference is the quantity of input, with higher SES parents talking more with their children. Some researchers have used evidence of a *lack* of SES effects on language development to infer that maturational mechanisms are responsible for the growth in language skills.

Challenges in constructing a computational model of SES effects on development

Computational models of development involve the interaction of a learning system with a training environment (Mareschal & Thomas, 2007). Usually, models are applied to capturing the developmental profile of the average child, and less frequently to individual differences. Some consideration has been given to altering the computational properties of the learning system to explain individual differences in intelligence (e.g., Garlick, 2002; Richardson et al., 2006a, 2006b) or the deficits observed in developmental disorders (e.g., Thomas & Karmiloff-Smith, 2003a, 2003b). To our knowledge, no models have sought to simulate individual differences that may stem from variations in the socio-economic status of the families in which children are raised.

The construction of a computational model of SES effects necessitates committing to a particular implementation of environmental variation, and then evaluating its adequacy in capturing observed behavioural data. Indeed, the virtue of modelling is the theoretical clarity required by such commitments. We initially considered two possible avenues of implementation. First, the environment may influence the computational properties of the learning system, perhaps via SES-related influences such as nutrition and stress. Second, SES may operate by influencing the nature of the information available in the environment. This might occur in three ways. It might alter the *nature* of the information available in the environment. It might alter the *quantity* of the information in the environment. Or it might influence the motivation of the child to engage with the information available, perhaps through differences in reward and punishment; in effect, this would modulate the subjective information that the child actually exploits from that objectively available in the environment. Based on Hart and Risley's (1995) findings of dramatic differences in the linguistic environment between families with different SES levels, we began by investigating the adequacy of manipulations in the information content of the environment for simulating SES effects on language development.

We employed the domain of past-tense acquisition for two reasons: first, there is a long history of computational modelling of this aspect of morphosyntax, and there is therefore some consensus on what the 'normal' learning system and training environment should look like. Second, we had available to us a data set of SES effects on the past-tense formation of 270 6-year-old children, as a target for our computational simulations. These data were originally published as part of Bishop (2005), although the SES effects were not reported in that work.

The simulation of SES effects on language development provides a novel challenge because SES effects are not a property of an individual but of a population. Moreover, it is widely held that individual differences of a genetic origin are responsible for a significant proportion of individual differences in behaviour, including language development (Plomin et al., 2008; Smith, 2007). Therefore we must simulate a large population of individuals, and implement both environmental and genetic contributions to individual differences. We expand on these challenges in the modelling section, where we also establish our criteria for success or failure of the simulations. First, we turn to the target data set.

Empirical data: English Past-Tense Formation

Children's acquisition of tense formation, along with other aspects of inflectional morphology, has been the focus of a great deal of empirical research. In part, this is due to the quasi-regular nature of the domain. Past tense comprises a regular rule (add –ed to form the past tense; e.g., *talk-talked*), which is readily extended to novel forms (*wug-wugged*), and also a set of irregular past tenses that are exceptions to the rule (*go-went*, *sing-sang*, *hit-hit*). The investigation of the processing structures necessary to acquire a quasi-regular domain have led to an extended debate (Pinker, 1999; Rumelhart & McClelland, 1986; see Thomas & McClelland, 2008, for a review). Some of the key data involve children's greater ease in acquiring the past tense of regular verbs compared to irregular verbs, and the presence of 'over-generalisation' errors, where children mistakenly apply the regular rule to the exception forms (*thinked*).

Tense acquisition has been considered more widely within the theoretical framework of the Optional Infinitive stage (Wexler, 1994, 1996). Young children pass through a phase where they sometimes omit grammatical morphemes, such as those marking tense, in contexts where the morphemes are obligatory for grammatical correctness. Where finite inflected forms are expected, children sometimes produce infinitival forms (in English, unmarked verb stems; e.g. *yesterday I talk to my friend*). Wexler suggested that in this phase of acquisition, children regard the infinitive as an optional form of the verb. Notably, in children with SLI, such infinitival forms are observed at ages where typically developing children have ceased to use them. This has led to the proposal that in SLI, there is an extended optional infinitive stage, so that problems in tense marking might be diagnostic of the disorder (Rice, Wexler, & Cleave, 1995).

In 2001, Rice and Wexler published a diagnostic test for early grammatical impairment in which tense marking was one of the key aspects of the assessment. The past-tense elicitation subtest assessed accuracy levels in producing regular and irregular past tenses, as well as the level of over-generalisation errors, where the regular rule is mistakenly applied to irregular verbs. Since such over-generalisation errors still represent finite forms (just the wrong one), Rice and Wexler used the three scores to compute a fourth, the proportion of regular and irregular verbs that were produced in the finite form. The finiteness measure was equal to the sum of correct regular, correct irregular, and incorrect over-regularised irregular verbs, divided by the total number of regular and irregular verbs attempted. Performance of a group of children on the Rice-Wexler past-tense task form the target empirical data for the computational simulations.

Bishop (2005) gave the Rice-Wexler past-tense sub-test to a population of 442 6-year-old children, for whom SES information was additionally available (Bishop, 2005; Petrill et al., 2004). These data were originally collected as part of a twin study and published in composite form (Bishop, 2005). The author kindly made the raw data available to us, including accuracy levels on regular and irregular verbs, over-generalisation error rates, and the computed finiteness measure, along with the SES measure. The Bishop (2005) data allowed us to assess the predictive power of SES on two different verb types and one error type at one particular age. To our knowledge, they represent the largest data set on SES effects on English past tense acquisition.

Method

The English past-tense data were derived from the Past Tense probe subtest of the Rice-Wexler Test of Early Grammatical Impairment (2001). In the Past Tense probe subtest, children are shown a picture of an action being carried out, while the experimenter describes the picture using a sentence containing the target verb in the present tense. The child is then shown a second picture with the action completed. The child is encouraged to describe the second picture. The task is designed to elicit the past tense in a full sentence with an overt subject. The test comprises 2 practice trials with regular verbs (*rake*, *skate*), and then 18 test trials including 10 regular verbs (*paint*, *brush*, *clean*, *kick*, *climb*, *jump*, *pick*, *plant*, *tie*, *lift*) and 8 irregular verbs (*catch*, *make*, *write*, *ride*, *dig*, *eat*, *blow*, *give*). Rice and Wexler selected the verbs on two criteria: to be familiar to young children, as indexed by appearance in Hall, Nagy and Linn's (1984) compilation of spoken words by children ages 4.05-5.00; and so that they could be clearly depicted by line drawings or pictures (Rice & Wexler, 2001, p.52). Comparison of regular and irregular verbs using the MRC Psycholinguistic Database (Coltheart, 1981) found that the verb types did not reliably differ on ratings of familiarity, imageability, Kucera-Francis (1967) word frequency, and length in phonemes. Bishop (2005) used a pre-publication version of the test, which employed 11 regular verbs (*brush*, *clean*, *climb*, *colour*, *jump*, *kick*, *paint*, *pick*, *plant*, *play*, and *wash*) and 8 irregular verbs (*catch*, *dig*, *fall*, *make*, *ride*, *swim*, *throw*, and *write*). Children's responses were coded according to whether they were correct, incorrect, over-regularised in the case of irregular verbs, or not attempted. The Past Tense Probe score is derived by dividing the total of verbs either correct or over-regularised divided by the total number of past-tense forms attempted. The Probe assesses the proportion of responses in finite form, and henceforth we refer to it as the *Finiteness* measure. The test/retest reliability for the Past Tense probe subtest is r = .82 (Rice & Wexler, 2001).

Participant details: The original Bishop (2005) sample contained 442 children with mean age 6 years and 6 months (range: 6 years and 0 months to 7 years and 1 month), comprising 250 boys and 192 girls. The children were originally recruited as twin pairs and in this sample, 224 were MZ twin pairs and 218 were DZ twin pairs. This group was over-sampled for risk of language disorder. Based on a parental questionnaire carried out when the children were aged four, 215 of the children had been flagged as at risk for language impairment. The questionnaire combined responses indicating whether the children were talking in full sentences, whether they had low vocabulary, and whether the parents were worried that the child's language was developing slowly (see Bishop, 2005, for details). At six years of age, 304 of the 442 children were viewed as normal (i.e., had not been diagnosed with any disorder, language or otherwise), implying that markers of risk had disappeared in about a third of the original at-risk children. For these children, SES data were available for a subset of 270 children. The final sample of children without a disorder and for whom SES data were available comprised 141 boys and 129 girls, with a mean age of 6 years 6 months (range 6 years 0 months to 7 years 0 months).

Socio-economic status: Demographic information was obtained via questionnaire from the first contact with the family at age 4. Five pieces of information were collected: the father's highest educational level and occupational status, the mother's highest educational level and occupational status, and the age of mother at birth of the eldest child. From these data, an index of SES was created based on a factor analysis. Using principal components analysis, a single-factor solution yielded an eigenvalue of 2.51, accounting for 50% of the variance; based on these results, a single composite was created by standardising the five variables and summing them using unit weights (Petrill et al., 2004, p.448). This method yielded a scale ranging from -1.57 (low SES) to +1.54 (high SES), with a mean of -0.16 and a standard deviation of 0.72.

Results

Table 1 includes the mean and standard deviation for accuracy of production of regular and irregular verbs, the rate of over-generalisation errors, and the composite finiteness measure. The expected advantage for regular verbs was present, while performance on irregular verbs was fairly poor, with high rates of over-generalisation (regulars > irregulars, $t(269)=35.76$, $p<.001$). At 46%, the over-generalisation rates were higher than those found in other studies of past-tense elicitation using different stimulus sets and procedures with 6-year-old children. For example, Thomas et al. (2001) and van der Lely and Ullman (2001) each collected control groups of 6-year-old children in past-tense elicitation studies (N=10 and N=12, respectively). In these studies, over-regularisation rates varied between 24 and 31%. Rates of regular accuracy were also lower for those samples (49%-80%), while irregular accuracy levels were comparable (19-45%). This disparity suggests either some task effect of the Rice-Wexler test that encouraged regular responses, or that the Bishop (2005) population sample differed from the (much smaller) control groups. The children were over-sampled for risk of language impairment and despite not displaying overt language deficits, some subtle residual difficulty might have remained.

===================

Insert Table 1 about here

===================

Figure 1 displays scatter plots linking past-tense performance with the SES measure for the four dependent variables. SES predicted regular verb performance at marginal significance, and irregular verb performance more robustly; it predicted irregular performance significantly more strongly than regular (regular: R^2=.013, F(1,268)=3.56, p=.060; irregular: R^2=.047, F(1,268)=13.13, p<.001; interaction: F(1,536)=5.90, p=.011). One concern with these data is the large number of children at ceiling on regular verbs, which might account for the reduced predictive power of SES. Figure 2 displays the data removing all children for whom regular verbs were at ceiling but irregulars were not (reduced sample N=64; mean performance levels for this sub-sample are shown in Table 1). The relationship between SES and verb performance remained of the same size, although the reduced participant numbers meant that the relationships were no longer statistically significant (regular: R^2=.013, F(1,62)=.82, p=.368; irregular: R^2=.047, F(1,62)=3.08, p=.084; interaction: F(1,124)=.987, p=.323). The fact there was a consistent relationship when ceiling effects were removed implies that the difference in the predictive power of SES between regular and irregular verbs is a real one, and provides more evidence for the differential effect of SES across parts of language (Vasilyeva, Waterfall and Huttenlocher, 2008). The low proportion of variance accounted for by SES in regular past-tense formation is consistent with the findings of Rice, Wexler, and Hershberger (1998), a result which those authors took to support a maturational theory of the acquisition of regular morphosyntax.

In sum, then, a regularity effect is present in both mean performance and the predictive power of SES, with SES picking up between 1-5% of population variance.

======================

Insert Figures 1 and 2 about here

=======================

Computational modelling

The computational modelling of SES effects proceeded as follows. We specified a base or 'normal' model of the acquisition of English past tense. We then designed a manipulation to the training environment, corresponding to the family in which simulated children are raised. We next designed a manipulation to the efficiency of the learning system, corresponding to the genetic contribution to individual differences. Henceforth, we refer to this as *intrinsic* rather than genetic variability, because it refers to the property of the past-tense learning system. The property must be an outcome of a prior developmental process that constructed the learning system, a process that will have both genetic and environmental contributions (whether cognitive or biological). Finally, we generated a large population of simulated individuals, each of whom underwent a developmental process of acquiring the English past tense. Population performance and the predictive power of environmental variations were then assessed.

We encountered two challenges in pursuing this design, one practical and one theoretical. First, population modelling by its nature necessitates the simulation of large numbers of individuals. Practically, this required some simplification to the base model. (Each simplified model took two hours to train and test, and we report on the results of 6000 networks in the following sections.) The simplification led to some shortcomings in the performance of the base model, but given the advanced state of past-tense modelling, these are at least well understood. The simplifications involved the use of an artificial language-like training set, rather than English monosyllabic verbs (following Plunkett & Marchman, 1991, 1993, rather than Joanisse & Seidenberg, 1999); and training solely on the past-tense paradigm for verbs, rather than simulating a system that learns all inflection types across multiple grammatical classes (Karaminis & Thomas, 2010).

The second, theoretical problem is that we do not know *a priori* the relative range of variability of environmental and intrinsic factors in our population of real children. Do actual environments vary just a little bit, with most environments providing decent information for the children, while intrinsic factors vary more widely, from very poor to very good learning systems? Or are all the children's learning systems reasonably efficient, while the linguistic environment varies greatly in its quality between children? Because empirical evidence was not available to constrain this aspect of the model, we simulated two levels of environmental variation and two levels of intrinsic variation. One of the goals of the simulation was to determine which combination gave the best fit to the past-tense empirical data.

The way in which we addressed the two challenges meant that it was important for us to clearly specify the criteria for success and failure of the simulations in accounting for SES effects in past-tense acquisition.

Criteria for evaluation the success or failure of the modelling

Because population modelling necessitated the use of a simplified base model of past-tense acquisition, we evaluated the success of the model on the qualitative fit to the empirical data rather than the quantitative fit. Ways to improve the quantitative fit are discussed later. We evaluated the success of the simulations on three criteria: (1) The qualitative fit to size of SES effects in predicting individual variability across each population, and the differential pattern across regular verbs, irregular verbs, over-generalisation errors, and proportion of finite responses. (2) The generation of novel testable predictions. We generated three such predictions, one of which we were able to test against the Bishop (2005) data set. (3) The generation of insights on candidate inferences from behaviour to mechanism, which is the particular contribution of computational modelling. What types of *modelling condition* led to what types of behavioural data? Are these the *conditions of the cognitive system* that researchers typically infer from these types of behavioural data?

Method

Architecture

Recent computational models of English past-tense acquisition have used connectionist networks to learn the association between the phonological form of the verb stem and the past-tense form. Along with the verb stem, other sources of information are provided at input, including lexical-semantic information, and information about the desired output inflection (Joanisse & Seidenberg, 1999; Karaminis & Thomas, 2010; Woollams et al., 2009). It should be noted that some researchers maintain that symbolic approaches are more appropriate for explaining past-tense acquisition, at least for regular verbs, where regularity is viewed as reflecting the operation of a rule-based mechanism (Pinker, 1999). However, these more linguistically oriented theories have not typically been realised in computational implementations of the developmental process. In the current simulations, a 3-layer, backpropagation network was used to learn to output the past-tense form of a verb from an input vector combining a phonological representation of the verb stem and lexical-semantic information. The architecture is shown in Figure 3.

====================

Insert Figures 3 about here

====================

Training set

For our training set, we used the "phone" vocabulary from Plunkett and Marchman (1991, p. 70). This comprised an artificial language set constructed to reflect many of the important structural features of English past-tense formation. There were 500 monosyllabic verbs, constructed using consonant-vowel templates and the phoneme set of English. Phonemes were represented over 19 binary articulatory features, a distributed encoding based on standard linguistic categorisations (Fromkin & Rodman, 1988). Separate banks of units were used to represent the initial, middle, and final phonemes of each monosyllable. The output layer incorporated an additional 5 features to represent the affix for regular verbs. The input layer included 500 units to encode the lexical status of each verb existing in the training set, using a localist encoding scheme (Joanisse & Seidenberg, 1999; Thomas & Karmiloff-Smith, 2003). Networks thus had 3x19+500=557 input units and 3x19+5=62 output units. There were four types of verbs in the training set: (1) regular verbs that formed their past tense by adding one of the three allomorphs of the +ed rule, conditioned by the final phoneme of the verb stem (e.g., *tame-tamed*, *wrap-wrapped*, *chat-chatted*); (2) irregular verbs whose past-tense form was identical to the verb stem (e.g., *hit-hit*); (3) irregular verbs that formed their past tenses by changing an internal vowel (e.g., *hide-hid*); (4) irregular verbs whose past-tense form bore no relation to its verb stem (e.g., *go-went*). The token frequency of this last type of irregular verb had to be higher for the network to learn them successfully (see Plunkett & Marchman 1991), as is the case in real languages. As a result, this verb type experienced three times as much training as the other verb types. The *type frequencies* were as follows. There were 410 regular verbs, and 20, 68, and 2, respectively, of each irregular verb type. Following Plunkett and Marchman (1991), the verbs were given a token frequency structure. For computational convenience, *token frequency* was implemented by mediating the weight change generated by the difference between the actual output and the target output (Plaut et al., 1996). The weight change of high frequency arbitrary verbs was multiplied by 0.9 during a given training presentation and that of low frequency arbitrary verbs by 0.6. The weight change of all other high frequency verbs (regulars, no change, and vowel change) was multiplied by 0.3 and of all other low frequency verbs by 0.1. These scalings represent a compromise between raw English frequency counts and log-transformed frequency counts. Token frequency effects were not a target of the current simulations. Finally, a separate set of novel verbs was constructed to evaluate the generalisation performance of the network. These verbs could differ depending on their similarity to items in the training set. For simplicity, we focused on 410 novel verbs each of which shared two phonemes with one of the regular verbs in the training set. Generalisation was evaluated depending on the proportion of these novel verbs that were assigned the correct allomorph of the regular past-tense rule.

Some limitations of this training set should be made clear. It represented a past-tense-like mapping task rather than the English past tense per se, since it did not use realistic verb sets in training and testing. There were only three different ways in which irregular past tenses were formed compared to over twenty in English; it did not include subtleties of phonological similarity such as the existence of rhyming verbs with different past-tense forms; and the frequency structure was also highly simplified. Nevertheless, the Plunkett and Marchman training set was sufficient to captures a number of basic empirical effects in child past-tense acquisition, including differential performance depending on the type and token frequency of verbs, as well as error patterns across development (Plunkett & Marchman, 1991, 1993, 1996). And the target empirical data, performance on the Rice-Wexler test (2001), contain relatively few high frequency regular and irregular verbs, so do not constitute a rich data set with respect to the effects of frequency and phonological similarity.

Implementing environmental variation

When Hart and Risley (1995) examined the influence of SES on English past-tense formation, they reported large differences in the number of past-tense verb forms addressed to the children, with professional families using three times as many as those on welfare. They noted there was little difference across SES in the richness of the past-tense verbs used, where richness was defined as the number of past tenses per utterance. Based on these data, we decided to begin by implementing SES as an alteration to the quantity of past-tense information available to each child. It is worth commenting that the Hart and Risley data did not investigate whether environmental variation differs across verb types, an issue that became relevant in our simulations.

In machine learning terms, we operationalised environmental variation as a time-varying function with respect to the training set. We assumed that there was a ‘perfect training set’, in this case comprising all of the verbs available in the language, along with their accepted past-tense forms. We defined the function for variations in the training set as follows:

$$\text{Training set } P_n T_t = f\,\{\textit{perfecttrainingset}, X, Y, Z\} \qquad \text{(Eq.1)}$$

The training set for Person *n* at time *t* is a function *f* of the perfect training set and three parameters: *X* = proportion of valid training trials, *Y* = proportion of invalid training trials, and *Z* = proportion of noise trials. Invalid trials have the same input as a training pattern but a different output. Noise trials have different inputs and outputs or include partial information consistent with training patterns. In principle, the function *f* could be influenced by Person *n*’s behaviour or experiences at *t-1*, creating a more complex dynamical equation. A dynamical equation would accommodate the possibility of, for instance, a reduced reward leading to reduced attention and therefore a subsequently smaller proportion of valid training trials.

We made the following assumptions in our implementation. First, once instantiated, we gave each network a *pre-conditioning* phase to produce divergent initial connection weights. This comprised a version of the training set with *X* set to 0, *Y* set to

0, and *Z* set to 1. The training set was made up of 30 random binary vectors at input and output, uniquely created for each individual, and trained for 50 epochs. This phase was intended to simulate the effect of early subjective experience prior to using the relevant learning system for (in this case) modulating phonological output forms via tense information. Next, we created a training set for the past-tense information available in each family environment. To do so, we generated a *family quotient* for each simulated child. This was our implementation of SES. The family quotient was a number between 0 and 100%. This value was used as a probability determining whether each verb in the perfect training set would be included in the family's vocabulary. In terms of Equation 1, *X*=family quotient, *Y*=0, *Z*=0. The family training set was then fixed throughout development. However, performance was always assessed against the full perfect training set (analogous to a standardised test of past-tense formation applied to all children). The family quotient manipulation corresponds to a reduction in type frequency for both regular and irregular verbs, while the token frequency of each verb (3 times greater presentation for high than low frequency) was retained.

Our final decision was how to sample the family quotient values. Inspection of SES distribution in Hart et al.'s (2007) maternal education data indicated a normal distribution with a large standard deviation but a positive skew. The distribution from Bishop (2005) for the composite SES measure had a similar normal distribution with large standard deviation but with a negative skew. We selected a uniform distribution, which slightly exaggerated the incidence of high and low SES. We selected two ranges of environmental variation, a *narrow range* with reasonably high quality, sampling family quotient values between 60% and 100%; and a *wide range* that accommodated potentially very poor quality environments, sampling quotient values between 0 and 100%. These manipulations led to similar variations in information for regular and irregular verbs.

Due to simulation results with lower rates of over-generalisation errors than those found in the empirical data, we considered two further manipulations in which irregular verb information was poorer than that for regular verbs. These were exploratory conditions that were not constrained by existing empirical data and evaluated the hypothesis that the rate of over-generalisation might in part be driven by impoverished irregular verb information received from the environment. In variant 1, two family quotient values were independently sampled for each simulated child. Regular verbs were sampled between 60% and 100% (mean 80%), while irregular verbs were sampled in the range 40-80% (mean 60%). This manipulation had *identical range* but a lower *absolute level* for irregular verbs. In variant 2, a lower absolute level for irregulars was achieved by *widening the range*: irregular verbs were sampled between 20 and 100% (mean 60%).

Implementing intrinsic variation

Connectionist networks contain a range of parameters that increase or decrease their ability to learn a given training set. Parameters such as learning rate, momentum, and number of hidden units feature in most published simulations. In models of normal/average development, such parameters are usually optimised to achieve best learning (usually in the presence of the perfect training set). Certain parameters have been proposed as candidates to explain individual differences, search as the learning rate (as a proxy for neuroplasticity; Garlick, 2002). However, a given parameter may have differential effects across the parts of a problem domain. For example, Thomas (2005) demonstrated how the ‘temperature’ or steepness of the sigmoid activation function in the artificial neurons had more effect on regular than irregular verbs in past-tense acquisition. In order to remain neutral with regard to which parametric variations were responsible for intrinsic variation in the learning system, we simultaneously varied a number of parameters across individuals. As with environmental variation, we considered two different ranges of intrinsic variation, either *narrow* or *wide*.

Fourteen computational parameters were allowed to vary between individuals, serving to alter the learning capacity of each network. The parameter settings allowed for over 2000 billion unique individuals. The parameters, split by their role, were as follows: *Network construction*: Architecture, number of hidden units, range for initial connection weight randomisation, and sparseness of initial connectivity between layers. *Network activation*: unit threshold function, processing noise, and response accuracy threshold. *Network adaptation*: backpropagation error metric, learning rate, and momentum. *Network maintenance*: weight decay, pruning onset, pruning probability, and pruning threshold. The parameter ranges for narrow and wide intrinsic variation can be found in Appendix 1.

Design

For each population, 1000 sets of the 14 computational parameter values were generated. These were instantiated as 1000 connectionist networks, which were then trained as follows. The individual was trained initially on a unique *pre-conditioning training set* to produce unique and divergent starting weights. A family quotient value was then generated from the appropriate range and used to create the *family training set*. Following pre-conditioning, each network was trained for 1000 epochs on its family training set. At each epoch, performance was measured on the perfect training set. Performance was assessed on regular verbs, irregular verbs, over-generalisation errors, and on generalisation of the past-tense rule to novel forms. Performance was measured in via nearest-neighbour accuracy levels (% correct). Four populations were run in a 2x2 design, of narrow or wide environmental variation and narrow or wide intrinsic variation, to assess which combination of variability would best explain the Bishop data.

Results

Mean levels

The Rice-Wexler (2001) test contains primarily vowel-change irregular past tenses. The simulated populations were benchmarked to the point in training where the mean accuracy of irregular vowel-change verbs was equivalent to the Bishop (2005) sample. Results for the four populations in the 2x2 design are included in Table 1. It is immediately obvious that, when matched on irregular verb performance, both regular verb performance and the rate of over-generalisation errors were lower in all populations than in the children. There are four potential reasons why this disparity occurred. First it could be due to the simplifications in our base model. Recent models of past-tense formation include aspects of their design and training that support regular past-tense formation. These include training on the verb stem at output, which is the main component of a regular past-tense form, and the provision of a 'past-tense unit' at input, which forms a strong association to the past-tense inflection at output (Karaminis & Thomas, 2010; Woollams et al., 2009).[1] Second, model performance was tested on the full training set, including high and low frequency items, whereas the empirical data were for a restricted set of regular and irregular verbs; and as we saw earlier, the influence of regularity in these data was higher for the Rice-Wexler data than that found in other past-tense elicitation tasks (with admittedly much smaller samples). Third, because the children in the Bishop (2005) study were drawn from a group over-sampled for risk of language disorder, it is possible that the children had slightly atypical systems, or the language of the parents was atypical in a way not modelled here. Finally, it may be that the greater influence of regularity in the Bishop data represents an environmental effect, namely, that for the children, irregular verb information is poorer than regular information. Across a range of SES environments, perhaps children are raised in a linguistic environment where parents and peers make more errors on irregular than regular verbs. Alternatively, in some dialects co-varying with SES, some irregular forms may be regularised. As a consequence, for a given level of irregular performance, regular performance would be higher, along with more over-generalisations.

We tested this last idea by running two additional conditions where irregular information was poorer on average than regular information. In variant 1, we lowered the absolute level of the family quotient value for irregular verbs but kept the range the same. In variant 2, we lower the absolute level but also widened the range of family quotient values for irregulars. These were run for a population with wide intrinsic variation. Table 1, rightmost columns, shows that such a manipulation indeed raised the relative level of regular verb performance compared to irregulars, and increased the rate of over-generalisation errors. However, the population means still demonstrated a shortfall in the influence of regularity compared to the full sample. We now shift to our main focus, factors predicting individual differences in the populations.

Predictive power of SES

The family quotient value was used as a proxy for SES, and used to predict individual differences in past-tense performance for the 2x2 design. Table 2 compares simulation results for the 4 populations, and Figure 4 demonstrates the scatter plots for comparison with Figure 1.

============================

Insert Table 2 and Figure 4 about here

============================

We refer to the populations by whether the intrinsic variation (I) was narrow (N) or wide (W), and whether the environmental variation (E) was narrow or wide. For the simulations, all conditions demonstrated a reliable regularity effect, where SES was a stronger predictor of irregular verb performance than regular verb performance. The regularity effect was larger in those conditions with narrow environmental variation ($F(1,3992)=37.53$, $p<.001$). When the environment varied widely in quality, this drove performance on regular and irregular verbs to a very similar extent. The regularity effect was not modulated according to whether intrinsic variation was narrow or wide ($F(1,3992)=.00$, $p=.974$). Only when intrinsic variation was wide did the negligible predictive power of SES on regular verbs emerge (IW-EN). This condition therefore represents the best qualitative fit to the empirical data. The condition did, however, over-estimate the predictive power of SES on over-generalisation errors. Of the two variant conditions, there was again a strong regularity effect, but variant 2 (where the range of variation of irregular information was widened compared to regular verbs) displayed much higher SES predictive power for irregulars and over-generalisation errors. Notably, despite this difference in the influence of the environment, the two variants had identical mean levels of performance (Table 1).

Some caution is necessary in directly mapping between data and model, because the model contains no measurement error either in the family quotient or in past-tense performance. In reality, SES is measured by variables such as parental income and education, which can only give an estimate of actual causal factors. In addition, we know that the Rice-Wexler test has some measurement error, indicated by its test-retest reliability of 0.8 (Rice & Wexler, 2001). With measurement error added to the simulations, the predictive power of our SES proxy would be smaller than the numbers in Table 2.

Novel predictions

1. SES and delay versus giftedness

We used the model to predict the extent to which variation in the information available in the environment could predict whether individuals would fall in the bottom or top 10% of the population, equivalent to developmental delay or giftedness. This analysis was carried out independently for the four dependent measures, at the point in development when the models were matched to the performance level of the Bishop (2005) sample on irregular verb performance. For over-generalisation errors, better performance was scored as fewer errors. The results are shown in Table 3.

====================

Insert Table 3 about here

=====================

From the previous section, the conditions with narrow environmental variation were closer to fitting the predictive power of SES for all the six-year-old children. Here, these conditions suggested that SES should be able to predict whether children were gifted but barely (if at all) able to predict whether children were delayed. To our knowledge, this novel prediction has not been made by any other theory of individual differences. The prediction could be tested by the Bishop (2005) data set, and the results are shown in the bottom half of Table 3. The data confirm that SES reliably predicted whether children would be in the top 10% of the population for regular verb performance, irregular verb performance, and the rate of over-generalisation. SES did not predict delay for any of the dependent measures. This provides powerful support for the model.

2. SES effects across development

The empirical data set provide a snapshot of the predictive power of SES at a single point in time. The simulations allow us to predict where the snapshot would fall within a developmental trajectory. That is, for our 2x2 design, we can assess whether SES effects should rise or fall with age. Figure 5 displays these data for the predictive power of the family quotient variable on the four dependent measures at an early point in training (50 epochs), mid point in training (100 epochs) and a late point in training (750 epochs). The surprising feature of this figure is that the model predicts that SES effects should rise across development. This is most marked for the conditions with wide environmental variation. The reason is that for these conditions, the environment becomes the *limiting factor* on the best performance that an individual can achieve. The pathway to this endpoint, by contrast, is influenced by the computational properties of the learning system. That is, when environmental variability is wide, intrinsic factors may alter rate of development but environmental factors will be a strong predictor of ceiling performance. This prediction remains to be tested against longitudinal data for past-tense acquisition.

===================

Insert Figure 5 about here

====================

3. Generalisation to diagnose the locus of environmental influence on development

Some tests of past-tense acquisition elicit novel past tenses from children, for words they have not heard before. The Rice-Wexler test used by Bishop (2005) did not include this condition. The inflection of novel verbs necessarily tests the generalisation ability of the system, rather than its storage of past-tense forms that it has previously encountered. Generalisation sometimes serves as a better index of the computational properties of a learning system than its ability to memorise knowledge. For the model, we compared the predictive power of the family quotient variable on individual performance for information in the training set (regular and irregular verbs) to generalisation performance on a set of novel verbs. To get a robust view of the relationship of the verb types, the percentage variance explained was calculated across a tranche of development, from epoch 50 to epoch 250 in each population. The mean of the 200 R^2 values for each verb type are shown in Table 4.

===================

Insert Table 4 about here

====================

When there was wide variability of information in the environment, family quotient predicted slightly more variability in regular than irregular verbs, and a substantial but lesser amount of variability in generalisation. When there was narrow variability of information in the environment and the computational properties of the learning system were more to the fore, family quotient predicted more variability in irregular than regular verbs, and negligible amounts of variability in generalisation. When the computational properties of the learning system were the most prominent factor (wide intrinsic variability, narrow information variability), the predictive ability of family quotient disappeared for generalisation but not for regular verbs. (These interactions were all highly reliable when analysed with a mixed analysis of variance).

The relevance of this finding is that, as we saw in the Introduction, SES may in principle affect either the information in the environment or the computational properties of the learning system (e.g., via the effects of nutrition or stress) – though in the current simulations, we considered only the former possibility. Generalisation is a closer index of the computation properties of a learning system than performance on the training set. Therefore, it may serve to isolate the locus of environmental effects: the simulations suggest that a comparison of the ability of SES to predict variability in performance on children's knowledge (training set) versus extension of that knowledge to novel forms (generalisation) may be one way to untangle whether the SES effects operate via the information content of the environment or via influencing the learning system itself.

Discussion

To our knowledge, these simulations represent the first attempt to capture the effects of SES in an implemented computational model of language development. We proposed to evaluate the simulation results on three criteria: (1) qualitative fit to the empirical data, (2) novel testable predictions, and (3) insight into inferences from behaviour to mechanism. How did the model do?

Qualitative fit to data

We elected to evaluate the model on a qualitative rather than quantitative fit to the empirical data because one key assumption of the model was unconstrained by the empirical data: the relative range of variation of environmental versus intrinsic sources of individual differences. The key feature of the empirical data was stronger performance on regular verbs than irregular verbs, small amounts of variance predicted by SES (between 0.7 and 4.7% of variance across the four dependent measures of past-tense performance assessed), and stronger predictive power of SES on irregular than irregular verbs. Qualitatively, the modelling condition best able to capture this pattern was one that combined wide intrinsic variation in the power of the computational learning systems with narrower variation in (and reasonably high quality of) the information content of the environment to which those learning systems were exposed. The regularity effect in population means and in the predictive power of SES emerged despite no differential treatment of regular and irregular verbs in either the architecture or the learning environment. Regular verbs were less sensitive to variation in the environment solely because of their systematic structure: learning one regular past tense helps in producing another, more than learning one irregular past tense helps in producing another. To some extent, systematicity in the structure of a problem domain serves to liberate the learning system from variations in the information available in the environment. Differential effects of SES across parts of language have been found elsewhere, such as in syntax, e.g., for complex but not simple sentences (Vasilyeva, Waterfall and Huttenlocher, 2008).

Despite including the main architectural assumptions of current models of past-tense acquisition, the quantitative fit of the model to the data was poor in two respects: in terms of the population means, there was insufficiently high performance on regular verbs and too few over-generalisation errors. There are a number of ways the quantitative fit to mean performance could be improved. One could increase the complexity of the base model of past-tense acquisition, by scaling up the model to learn multiple inflectional classes across multiple grammatical classes, and one could employ a more realistic training corpus (see Karaminis & Thomas, 2010; Woollams et al., 2009). Recent models include architectural constraints and training phases that improve performance on regular verbs and generalisation (e.g., see Note 1). We could improve the quantitative fit of the level of variance predicted by the SES proxy by post hoc adjustments to the relative ranges of intrinsic variation and environmental variation, but since these ranges are not empirically constrained, this would constitute data fitting.

It is worth noting that state-of-the-art models of inflectional morphology are addressed to capturing the development of *the average* child. Their parameters, in terms of training set, training regime and learning system, are optimised to simulate that profile. By contrast, the simulation results we report for performance on past-tense production are arithmetic means of a large population; if there is ‘an’ average individual, in our simulations this network will be *sub-optimal*, both in its training set and its computational parameters. This distinction reflects the different focus of the two modelling enterprises: simulating development alone, versus simulating both development and individual differences at the population level.

The mismatch between all our initial modelling conditions and the level of over-generalisation errors observed in the Bishop sample did motivate us to consider two exploratory variants of the model. We considered the possibility that knowledge of irregular verbs might be poorer because the information content of the environment was poorer with respect to irregular verbs than regular verbs. Hart and Risley (1995) did not distinguish past tenses based on regularity, so data cannot yet directly constrain the possibility of *regularity x SES* interactions on child-directed past-tense information. However, it is possible that children may experience less reliable information on irregular verbs, especially if some of this information is from peers who are themselves making over-generalisation errors in production. Moreover, some dialects characteristic of different SES levels differ with respect to their past tenses. Notably, when we sought to implement the idea of a *regularity x SES* interaction, it became apparent that there is more than one way to achieve this result. Irregular verb information may have a lower absolute level of quality but the same range of variability; or it may have a lower absolute level as well as a wider range of variability. Our simulations demonstrated that either was sufficient to increase the level of over-generalisation errors. Moreover, these variants did not maximise the potential for such errors since irregular verbs were only omitted from the training set, rather than added in their regularised form. It is of interest that the two variants exhibited equivalent population mean levels of accuracy (Table 1) but differed in the extent to which measures of the variation in environmental quality could predict individual differences in irregular past-tense performance (Table 2). This is a demonstration that the characteristics of a population may dissociate with respect to development and individual differences.

While empirical data cannot yet directly speak to the existence of a *regularity x SES* interaction, there are suggestive data. Bishop's original data set comprised 224 MZ and 218 DZ twins (Bishop, 2005). For that data set, the correlation between regular past-tense performance of MZ twins was .67 and for DZ twins was .12; for irregulars, the MZ correlation was .45 and the DZ .42 (Thomas, Forrester, & Ronald, in preparation). A behavioural genetic analysis points to a greater contribution of environmental variation to individual differences in irregular verb performance than regular verb performance. One way to produce such a difference would be if the range of variation in the environmental input were wider for irregular verbs than regular verbs. That said, as we have seen, systematicity in regular verbs means that environmental variation has less effect on their development: systematicity may exaggerate estimates of heritability.

Novel predictions

The model produced three predictions, one of which was testable against the Bishop (2005) data set. The first prediction was that SES would reliably predict whether a child was performing in the top 10% of the population, but not whether a child was performing in the bottom 10% of the population. To our knowledge, this is prediction has not been made by any existing theory of individual differences. The Bishop data set confirmed this prediction, providing a powerful validation of the model even in its qualitative form.

Next, the model predicted that where environment is the limiting factor on performance, SES effects should increase across development. This remains to be tested against longitudinal past-tense data. Such data would need to ensure the widest range of environmental variation possible, and exclude children with heritable language disorders, to narrow genetic range. A few studies have taken longitudinal measures of other measures of language and cognition. In some cases, increases in SES effects are observed over development. When Petrill et al. (2004) tested expressive vocabulary and grammatical complexity via parental report in a sample of 6000-8000 twins, SES predicted 3.2% of the variance at age 3 and 3.6% at age 4. In a recent longitudinal study, Tucker-Drob et al. (2011) reported that in a sample of 750 twin pairs, SES was not related to mental ability at 10 months but was present at 2 years of age. By contrast, when Hart et al. (2007) tested 287 pairs of twins of elementary school age on a test of general cognitive ability at two time points two years apart, SES predicted 5.8% of the variance at time 1 and 2.9% of the variance at time 2.

Last, the model predicted that variation in the computational properties of the learning system might be more closely linked to generalisation than to performance on the training set. A comparison of the ability of SES to predict variability in performance on children's knowledge (training set) versus extension of that knowledge to novel forms (generalisation) may be one way to untangle whether the SES effects operate via the information content of the environment or via influencing the properties of the learning system itself.

Insight into inferences from behaviour to mechanism

Using a sample of 20 children, Rice, Wexler and Hershberger (1998) reported that SES (as measured by maternal education) did not reliably predict past-tense performance, explaining less than 1% of the variance. They inferred that the development of this aspect of morphosyntax was best explained in terms of maturational mechanisms, where changes in behaviour over time are due to the aging process rather than experience dependent learning. The logic was that if development is not sensitive to variations in the environment, then developmental mechanisms cannot be relying on the environment. For the Bishop data set, SES similarly predicted only around 1% of the variance in regular past-tense performance, supporting the Rice et al. result. In our simulations, the condition that combined narrow variation in the information content of the environment with wide variation in the computational properties of the learning system also reproduced the result that the SES proxy predicted around 1% of the variance in past-tense performance. Crucially, however, the simulations demonstrate that Rice et al.'s inference about developmental mechanisms is unsafe. This is because development in the connectionist model was *entirely experience dependent* and not at all maturational. Without exposure to the training set, no network would have learned anything. The lesson from the model is that caution must be exercised in drawing inferences about developmental mechanisms based on data from individual differences. The failure of measures of environmental variation to predict individual differences does not legitimise conclusions about the role of the environment in the developmental process.

As we saw in the novel predictions above, both empirical data and model indicated that there was no statistical relation between SES and whether a child fell in the bottom 10% of the population on past-tense formation. Because we know how the model functions, we can see that in fact, this result is a fairly curious one. We know that for the model, a poor environment *does cause* poor acquisition: the same model exposed to a poorer training set will learn less. We have, here, a divergence between statistical relations and causal relations. However, unlike the more familiar case of correlation without causation, here we have causation (poor environment causes poor development) without correlation (poor environment does not predict poor development). How can this be? The reason for the divergence is that the relationship of cause to effect was *many-to-one*. In addition to an impoverished environment, there were many other possible causes of poor acquisition, resulting from the settings of computational parameters. The many-to-one causal relationship diluted the strength of the statistical association of any one cause. By contrast, both parameters and environment had to be good for a network to feature in the top 10%; whether the environment was good or not then had stronger predictive power. In a population with largely adequate computational parameters, the nature of the environment should be a more symmetrical predictor of success and failure, and this was confirmed by simulation of such a population (Table 3, *IN-EW* condition). The asymmetry was therefore predicted to be sensitive to the sampling. This is a further indication that in the actual data, variations in learning ability were a stronger determinant of individual differences than environmental variation in information. The model, then, demonstrates that many-to-one causal relations may compromise the window that statistical associations offer on causal mechanism.

Future challenges for investigating SES through computational modelling

We identified several routes through which SES effects might operate at the level of mechanism, and explored the adequacy of one of these, the amount of information available in the environment, to explain observed SES effects on behaviour. Future work should be addressed to considering other possible avenues: variations in the quality of the information available, variations in the subjective environment for the child (through variation in attention, perhaps induced by reward and punishment schedules), and influences on the computational properties of the learning system via the effects of stress and/or diet associated with poverty (Hackman & Farah, 2009). In Equation 1, we included parameters to capture the possibility of invalid training trials and time varying training sets.

One complication of studying SES is that causal pathways through which its effects operate may alter across development, and indeed may depend on the range of SES under consideration. As we saw earlier, Aikens and Barbarin (2008) found that *family characteristics* predicted more of the SES-linked variability in initial reading ability in kindergarten children but *school and neighbourhood conditions* explained more in subsequent improvements in reading through to third grade. In the first world, even in the face of relative poverty, there is some minimal provision for the healthy upbringing of children. From the point of view of cognitive development, environmental variation may impact more on the information available, and on the particular schedules of reward and punishment experienced by a child. However, in the third world, the environmental range is much wider. Nutritional deficits during child development can be severe enough to cause stunting in growth and a statistically associated incidence of poor cognitive development (Grantham-McGregor et al., 2007). In this wider range, environmental variation may impact much more on biological aspects of neural function and therefore its computational properties.

Finally, we reiterate the focus of this paper on the causal mechanisms by which SES effects operate. Research in this field is challenged by the many confounded factors associated with SES. They may all play a causal role, or some may be non-causal correlations. Computational modelling permits consideration of the adequacy of specific factors to explain behavioural data, but of course does not demonstrate that these mechanisms are truly responsible. For this, intervention studies are required. The potential reward of understanding causal pathways is that although the confounded factors may be many, if the causal pathways are few, then alleviating the effects of poverty on cognitive development may be easier than the alleviating poverty itself.

Footnotes

1. For example, we trained Karaminis and Thomas's (2010) Multiple Inflection Generator model on just the English past tense, or simultaneously on multiple inflections for English verbs, nouns, and adjectives. At a point in training matched on vowel-change irregular verbs at 40% accuracy, the multiple-inflection-model had regular verb accuracy levels 20% higher than the past-tense-only model.

Acknowledgements

This research was supported by ESRC grant RES-062-23-2721, MRC grant G0300188, and a Leverhulme Study Abroad Fellowship. We thank Dorothy Bishop for allowing access to her data set, and for her help and encouragement throughout this project.

References

Aikens, N. L. & Barbarin, O. (2008). Socioeconomic differences in reading trajectories: The contribution of family, neighborhood, and school contexts. *Journal of Educational Psychology*, *100(2)*, 235-251.

Bishop, D. V. M. (2005). DeFries-Fulker analysis of twin data with skewed distributions: Cautions and recommendations from a study of children's use of verb inflections. *Behavior Genetics*, *35(4)*, 479-490.

Coltheart, M. (1981). The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology*, *33A*, 497-505.

D'Angiulli, A., Herdman, A., Stapells, D., & Hertzman, C. (2008). Children's event-related potentials of auditory selective attention vary with their socioeconomic status. *Neuropsychology*, *22(3)*, 293-300.

Farah, M. J., Shera, D. M., Savage, J. H., Betancourt, L., Giannetta, J. M., Brodsky, N. L., Malmud, E. K., & Hurt, H. (2006). Childhood poverty: specific associations with neurocognitive development. *Brain Research*, *1110*, 166-174.

Fromkin, V. & Rodman, R. (1988). *An introduction to language (4th Ed.)*. Holt, Rinehart and Winston, Inc.: London.

Garlick, D. (2002). Understanding the nature of the general factor of intelligence: The role of individual differences in plasticity as an explanatory mechanism. *Psychological Review*, *109(1)*, 116-136.

Ginsborg, J. (2006). The effects of socio-economic status on children's language acquisition and use. In J. Glegg & J. Ginsborg (Eds.), *Language and social disadvantage: Theory into practice*. John Wiley & Sons Ltd.

Grantham-McGregor, S., Cheung, Y. B., Cueto, S., Glewwe, P., Richter, L. et al. (2007). Child development in developing countries 1: Developmental potential in the first 5 years for children in developing countries. *The Lancet*, *369*, 60-70.

Hackman, D. A., & Farah, M. J. (2009). Socioeconomic status and the developing brain. *Trends in Cognitive Sciences*, *13(2)*, 65-73.

Hall, W. S., Nagy, W. E., & Linn, R. (1984). *Spoken words: Effects of situation and social group on oral word usage and frequency*. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, Maryland: Paul H. Brookes Publishing Co.

Hart, S. A., Petrill, S. A., Deckard, K. D. & Thompson, L. A. (2007). SES and CHAOS as environmental mediators of cognitive ability: A longitudinal genetic analysis. *Intelligence*, *35*, 233-242.

Huttenlocher, J., Vasilyeva, J., Cymerman, E., & Levine, S. (2002). Language input and child syntax. *Cognitive Psychology*, *45*, 337-374.

Joanisse, M. F., & Seidenberg, M. S. (1999). Impairments in verb morphology after brain injury: A connectionist model. *Proceedings of the National Academy of Sciences of the United States of America*, *96*, 7592-7597.

Karaminis, T. N., & Thomas, M. S. C. (2010). A cross-linguistic model of the acquisition of inflectional morphology in English and Modern Greek. In: S. Ohlsson & R. Catrambone (Eds.), *Proceedings of 32nd Annual Conference of the Cognitive Science Society*, August 11-14, 2010, Portland, Oregon, USA.

Klibanoff, R. S., Levine, S. C., Huttenlocher, J., Vasilyeva, M., & Hedges, L. V. (2006). Preschool children's mathematical knowledge: The effect of teacher "Math Talk". *Developmental Psychology*, *42(1)*, 59-69.

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence: Brown University Press.

Mareschal, D. & Thomas M. S. C. (2007) Computational modeling in developmental psychology. *IEEE Transactions on Evolutionary Computation (Special Issue on Autonomous Mental Development)*, *11(2)*, 137-150.

Petrill, S. A., Pike, A., Price, T. & Plomin, R. (2004). Chaos in the home and socioeconomic status are associated with cognitive development in early childhood: Environmental mediators identified in a genetic design. *Intelligence*, *32*, 445-460.

Pinker, S. (1999). *Words and rules: The ingredients of language*. Basic Books, New York.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. E. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, *103*, 56-115.

Plomin, R., DeFries, J. C., McClearn, G. E., & McGuffin, P. (2008). *Behavioral genetics (5th Edition)*. New York: Worth Publishers.

Plunkett, K. & Marchman, V. (1991). U-shaped learning and frequency effects in a multilayered perceptron: Implications for child language acquisition. *Cognition*, *38*, 1-60.

Plunkett, K. & Marchman, V. (1993). From rote learning to system building: acquiring verb morphology in children and connectionist nets. *Cognition*, *48*, 21-69.

Plunkett, K. & Marchman, V. (1996). Learning from a connectionist model of the English past tense. *Cognition*, *61*, 299-308.

Raizada, R. D. S. & Kishiyama, M. (2010) Effects of socioeconomic status on brain development, and how cognitive neuroscience may contribute to leveling the playing field. *Frontiers in Human Neuroscience*, *4*. doi:10.3389/neuro.09.003.2010

Raizada, R. D. S., Richards, T. L., Meltzoff, A., & Kuhl, P. K. (2008). Socioeconomic status predicts hemispheric specialisation of the left inferior frontal gyrus in young children. *NeuroImage*, *40*, 1392-1401.

Rice, M. L., & Wexler, J. (2001). *Rice/Wexler Test of Early Grammatical Impairment*. San Antonio: Psychological Corporation.

Rice, M. L., Wexler, K., & Cleave, P. (1995). Specific language impairment as a period of extended optional infinitive. *Journal of Speech, Language and Hearing Research*, *38*, 850-863.

Rice, M. L., Wexler, K., & Hershberger, S. (1998). Tense over time: The longitudinal course of tense acquisition in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *41*, 1412-1431.

Richardson, F. M., Baughman, F. D., Forrester, N., & Thomas, M.S.C. (2006b). Computational modelling of variability in the balance scale task. In *Proceedings of the 7th International Conference of Cognitive Modeling.* (pp 256-261). Trieste, Italy: Edizioni Goliardiche.

Richardson, F. M., Forrester, N. A., Baughman, F. D., & Thomas, M. S. C. (2006a). Computational modeling of variability in the conservation task. In *Proceedings of*

The 28th Annual Conference of the Cognitive Science Society (p. 2010-2015), July 26-29, Vancouver, BC, Canada.

Rumelhart, D. E. & McClelland, J. L. (1986). On learning the past tense of English verbs. In J. L. McClelland, D. E. Rumelhart, & the PDP Research Group (Eds.) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 2: Psychological and Biological Models* (pp. 216-271). Cambridge, MA: MIT Press.

Smith, S. D. (2007). Genes, language development, and language disorders. *Mental Retardation and Developmental Disabilities Research Reviews*, *13(1)*, 96-105.

Stevens, C., Lauinger, B. and Neville, H. (2009). Differences in the neural mechanisms of selective attention in children from different socioeconomic backgrounds: An even-related brain potential study. *Developmental Science*, *12(4)*, 634-646.

Thomas, M. S. C. (2005). Characterising compensation. *Cortex*, *41(3)*, 434-442.

Thomas, M. S. C. & Karmiloff-Smith, A. (2003a). Modelling language acquisition in atypical phenotypes. *Psychological Review*, *110(4)*, 647-682.

Thomas, M. S. C. & Karmiloff-Smith, A. (2003b). Connectionist models of development, developmental disorders and individual differences. In R. J. Sternberg, J. Lautrey, & T. Lubart (Eds.), *Models of Intelligence: International Perspectives*, (p. 133-150). American Psychological Association.

Thomas, M. S. C., & McClelland, J. L. (2008). Connectionist models of cognition. In R. Sun (Ed.), *Cambridge handbook of computational cognitive modelling* (pp. 23-58). Cambridge: Cambridge University Press.

Thomas, M. S. C., Forrester, N. A., & Ronald, A. (in prep.). *A connectionist model of the heritability of English past tense formation in six-year-old children*. Manuscript in preparation.

Thomas, M. S. C., Grant, J., Barham, Z., Gsodl, M., Laing, E., Lakusta, L., Tyler, L. K., Grice, S., Paterson, S. & Karmiloff-Smith, A. (2001). Past tense formation in Williams syndrome. *Language and Cognitive Processes*, *16 (2/3)*, 143-176.

Tucker-Drob, E. M., Rhemtulla, M., Harden, K. P., Turkheimer, E., & Fask, D. (2011). Emergence of a gene x socioeconomic status interaction on infant mental ability between 10 months and 2 years. *Psychological Science*, *22(1)*, 125-133.

van der Lely, H. K. J., & Ullman, M. (2001). Past tense morphology in specifically language impaired children and normally developing children. *Language and Cognitive Processes*, *16*, 177-217.

Vasilyeva, M., Huttenlocher, J., & Waterfall, H. (2006). Effects of language intervention on syntactic skill levels in preschoolers. *Developmental Psychology*, *42(1)*, 164-174.

Vasilyeva, M., Waterfall, H., & Huttenlocher, J. (2008). Emergence of syntax: commonalities and differences across children. Developmental Science, 11(1), 84-97.

Wexler, K. (1994). Optional infinitives, head movement and the economy of derivations. In D. Lightfoot and N. Hornstein (Ed.), *Verb movement*. Cambridge, England: Cambridge University Press.

Wexler, K. (1996). The development of inflection in a biologically based theory of language acquisition. In M. L. Rice (Ed.), *Towards a genetics of language*. Mahwah, NJ: Lawrence Erlbaum Associates.

Woollams, A., Joanisse, M., & Patterson, K. (2009). Past-tense generation from form versus meaning: Behavioural data and simulation evidence. *Journal of Memory and Language*, *61*, 55-76.

Appendix

Computational parameter	*Narrow range*	*Wide range*
Architecture	3-layer only	2-layer, 3-layer, or fully-connected
Hidden units	15-35	6-500
Sparseness	80-100%	30-100%
Weight variance	±0.01 to ±1	±0.005 to ±2.25
Processing noise	0-0.75	0-5
Sigmoid activation function temperature	0.25-1.75	0.0625-4
Connection weight decay	0 to 9.8 x10^{-6} per pattern presentation	0 to 9.8 x 10^{-5} per pattern presentation
Pruning onset	50-200 epochs	0-1000 epochs
Pruning threshold	0.1-0.75	0.1-1.5
Pruning probability	0.01-0.75	0.1-1.5
Learning algorithm error measure	Cross-entropy	Euclidean distance or Cross-entropy
Learning rate	0.05-0.2	0.005-0.5
Momentum	0-0.25	0-0.75
Nearest neighbour threshold	0.01-0.2	0.005-0.5

Tables

Table 1. English past-tense performance of the Bishop (2005) sample of 270 6-year-old children on the Rice-Wexler test (2001), for regular verbs, irregular verbs, over-generalisation (OG) errors, and a composite finiteness measure; and simulation data for the 2x2 design with different ranges of intrinsic (I) and environmental (E) variation (W = wide; N = narrow). Also shown are empirical data for a sub-sample of children removing those who had regular verb but not irregular verb performance at ceiling; and two simulation conditions where the environmental information for irregular verbs was poorer than for regular verbs (see Methods). Values show mean accuracy and (standard deviation).

Measure	*Data*		*Simulation condition*					
	N=270	N=64	*IN-EN*	*IN-EW*	*IW-EN*	*IW-EW*	*IW-EN*	*IW-EN*
			N=1000 per population				*Variant1*	*Variant2*
Regular	96 (12)	83 (20)	79 (19)	66 (30)	66 (31)	62 (31)	79 (26)	80 (25)
Irregular	42 (25)	39 (32)	42 (25)	42 (28)	42 (30)	40 (29)	42 (24)	42 (29)
OG errors	46 (46)	38 (26)	16 (12)	17 (13)	9 (10)	12 (12)	21 (15)	23 (20)
Finiteness	91 (12)	80 (20)	68 (20)	62 (29)	59 (30)	57 (30)	71 (26)	72 (26)

Table 2. Predictive power of SES on English past-tense performance of the Bishop (2005) sample of 270 6-year-old children on the Rice-Wexler test (2001), for regular verbs, irregular verbs, over-generalisation errors, and a composite finiteness measure; and simulation data for the 2x2 design with different ranges of intrinsic (I) and environmental (E) variation (W = wide; N = narrow), along with two variant conditions. In the variant conditions, where family quotient differed for regular and irregular verbs, the mean value was used to predict performance. Values in *italics* are 95% confidence intervals around the R^2 value.

Measure	*Data*		*Simulation condition*					
	% variance explained	p-value	*IN-EN*	*IN-EW*	*IW-EN*	*IW-EW*	*IW-EN + Variant 1*	*IW-EN Variant 2*
			% variance explained					
Regular	1.3	0.060	6.2	70.0	0.8	44.2	2.7	0.0
	0.0 – 4.4		*4.0 – 8.8*	*67.3 – 72.5*	*0.1 – 2.0*	*40.3 – 48.0*	*1.3 – 4.6*	*0.0 – 0.0*
Irregular	4.7	0.000	7.6	69.9	2.6	43.7	6.7	25.2
	1.4 – 9.6		*5.1 – 10.4*	*67.2 – 72.4*	*1.2 -4.5*	*39.8 -47.5*	*4.4 – 9.4*	*21.3 – 29.1*
OG errors	3.8	0.001	12.1	3.3	11.3	7.0	3.2	32.1
	0.9 – 8.3		*9.1 – 15.4*	*1.7 – 5.4*	*8.4 -14.5*	*4.6 – 9.7*	*1.6 – 5.2*	*28.1 – 36.1*
Finiteness	0.7	0.170	3.6	65.8	0.5	37.1	2.3	0.4
	0 – 3.3		*1.9 – 5.7*	*62.8 – 68.6*	*0.0 – 1.5*	*33.1 – 41.0*	*1.0 – 4.1*	*0.0 – 1.3*

Table 3. Role of the environment in predicting performance in the tails (delayed=bottom 10%, gifted = top 10% of population). For simulations, environmental index was *family quotient value*; for empirical data, it was *SES* (Bishop, 2005; Petrill et al., 2004). Values in cells show percentage of variance explained. Values in *italics* show 95% confidence intervals around the R^2 value.

Simulation						
			Environmental variation			
			Narrow		Wide	
			Delayed	Gifted	Delayed	Gifted
Intrinsic variation	Narrow	Regular	1.2	11.1	23.2	21.5
			0.3 – 2.6	*8.2 – 14.3*	*19.4 – 27.1*	*17.8 – 25.3*
		Irregular	0.6	8.7	15.3	21.9
			0.1 – 1.7	*6.1 – 11.7*	*12.0 – 18.9*	*18.1 – 25.8*
		OG errors	1.6	4.8	1.9	3.9
			0.6 – 3.1	*2.9 – 7.2*	*0.7 – 3.6*	*2.2 – 6.1*
		Finiteness	0.7	5.6	22.3	19.0
			0.1 – 1.8	*3.5 – 8.1*	*18.5 – 26.2*	*15.4 – 22.7*
	Wide	Regular	0.0	11.5	7.1	19.2
			0.0 – 0.4	*8.5 – 14.8*	*4.7 – 9.9*	*15.6 – 30.0*
		Irregular	0.0	11.8	3.3	19.0
			0.0 – 0.3	*8.8 – 15.1*	*1.7 – 5.4*	*15.4 -22.7*
		OG errors	1.3	2.3	3.1	1.2
			0.3 – 2.7	*1.0 – 4.1*	*1.6 – 5.1*	*0.3 – 2.6*
		Finiteness	0.0	6.8	7.0	15.6
			0.0 – 0.5	*4.5 – 9.5*	*4.6 – 9.7*	*12.2 – 19.2*

Data

	Delayed		Gifted	
	% explained	p-value	% explained	p-value
Regular	0.9	.117	3.5	.002
	0.0 – 3.7		*0.8 – 7.9*	
Irregular	0.2	.493	3.6	.002
	0.0 – 2.1		*0.8 – 8.1*	
OG errors	0.8	.138	2.2	.016
	0.0 – 3.5		*0.2 – 6.0*	
Finiteness	0.6	.204	0.1	.595
	0.0 – 3.1		*0.0 – 1.7*	

Table 4. Predictive power of variation in the information available in the environment for learning the training set (regular and irregular verbs) versus generalisation (novel verbs) for the simulations. Values in cells show percentage of variance explained by the environmental index, *family quotient value*, averaged over epochs 50-250 for each modelling condition. The value in brackets depicts the standard deviation of these 200 R^2 values.

		Variability of environmental information			
		Wide		Narrow	
		Variability of intrinsic computational properties			
		Wide	Narrow	Wide	Narrow
Training set	Regular verbs	34.3 (4.1)	68.8 (3.5)	0.8 (0.1)	11.9 (3.0)
	Irregular verbs	29.3 (5.9)	67.1 (10.5)	2.3 (0.6)	15.8 (4.8)
Generalisation	Novel verbs	24.7 (1.9)	52.2 (1.5)	0.0 (0.0)	2.7 (0.0)

Figure captions

Figure 1: Scatter diagrams relating SES (Petrill et al., 2004) to past-tense performance on the Rice-Wexler (2001) test for 270 6-year-old children (Bishop, 2005). Data show the accuracy of production of regular past tenses and irregular past tenses, the proportion of over-generalisation errors for irregular verbs, and the proportion of finite responses. (N=270 per plot).

Figure 2: Data from Figure 1, removing children who were at ceiling on regular but not irregular past-tense production (N=64 per plot).

Figure 3: Architecture of the past-tense model.

Figure 4: Scatter diagrams relating family quotient values (the SES proxy; x-axis) to past tense performance for the model (y-axis). Simulated data are for the 2x2 design with a narrow or wide range of intrinsic variation and a narrow or wide range of environmental variation, for regular verbs (R), irregular verbs (I), over-generalisation errors (OG), and finite responses (F) (see Figure 1). (N=1000 per plot).

Figure 5: The amount of variance in simulated past-tense performance explained by family quotient values (the SES proxy) at three different points in training, early (50 epochs), mid (100 epochs), and late (750 epochs), for regular verbs (Regular), irregular verbs (Irreg), over-generalisation errors (OG) and proportion of finite responses (FIN).

Data are shown for the 2x2 design of intrinsic (I) and environmental (E) variation. Error bars show 95% confidence intervals around R^2 values.

Figures

Figure 1

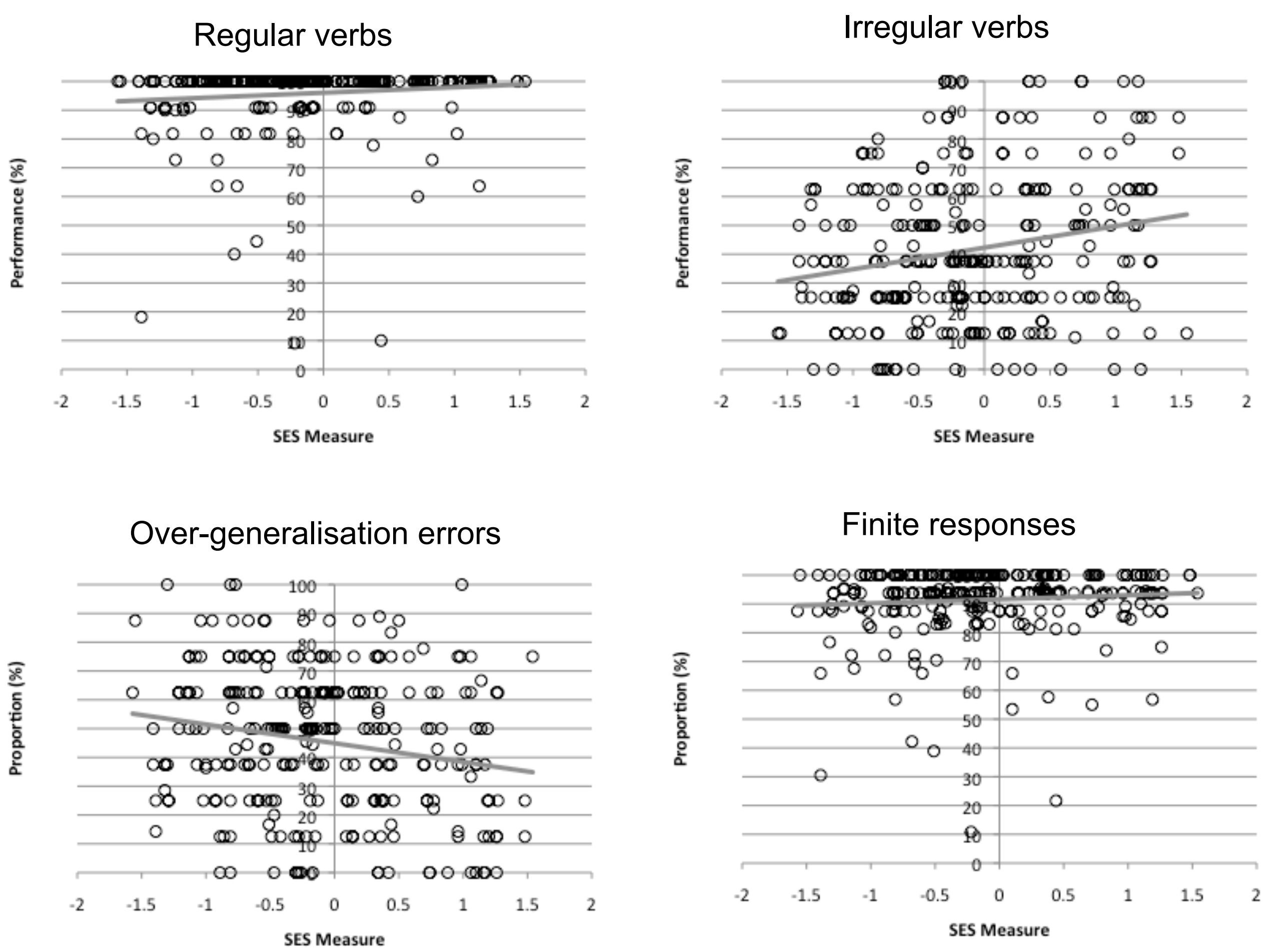


Figure 2

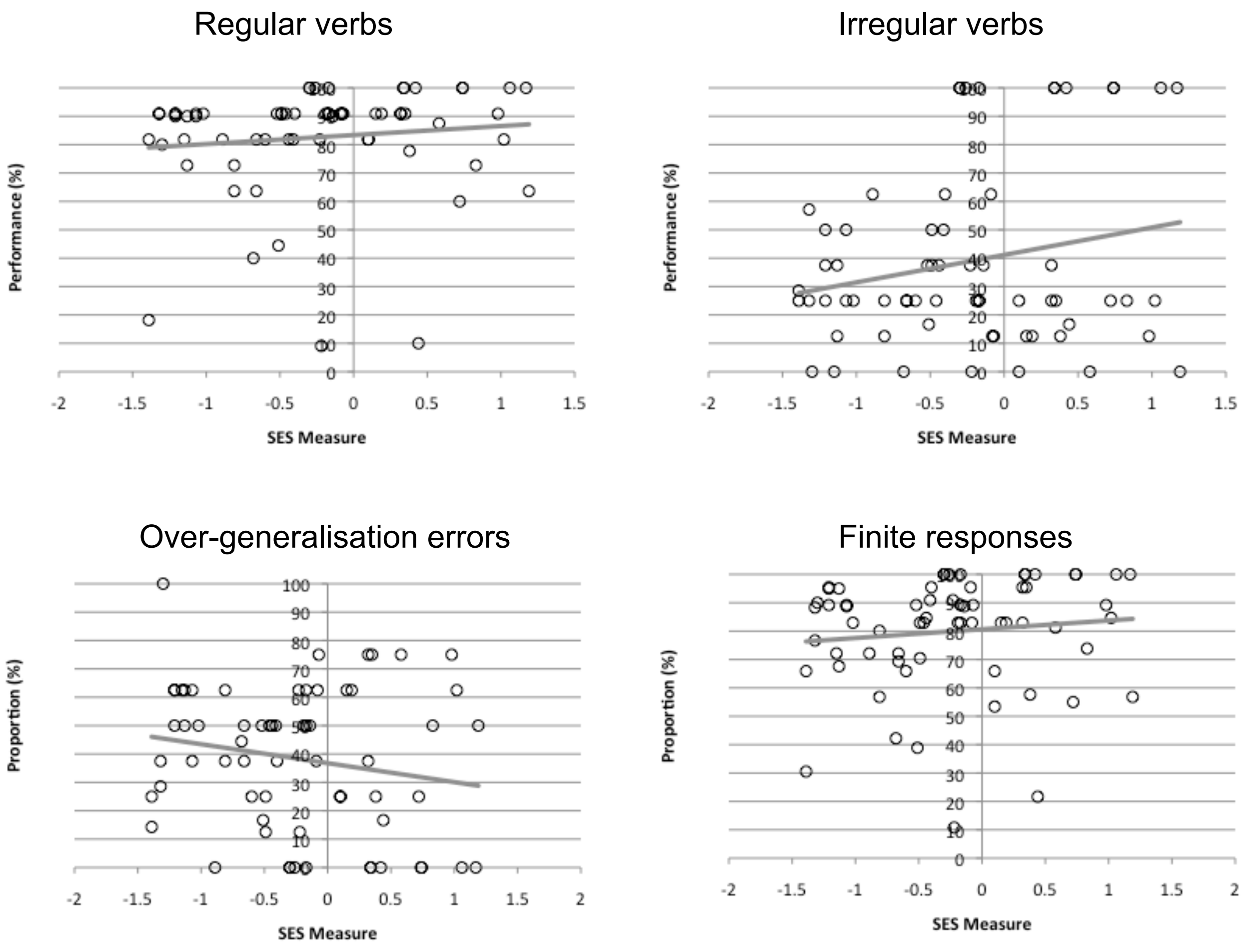

Regular verbs
Performance (%)
SES Measure
Irregular verbs
Performance (%)
SES Measure
Over-generalisation errors
Proportion (%)
SES Measure
Finite responses
Proportion (%)
SES Measure

Figure 3

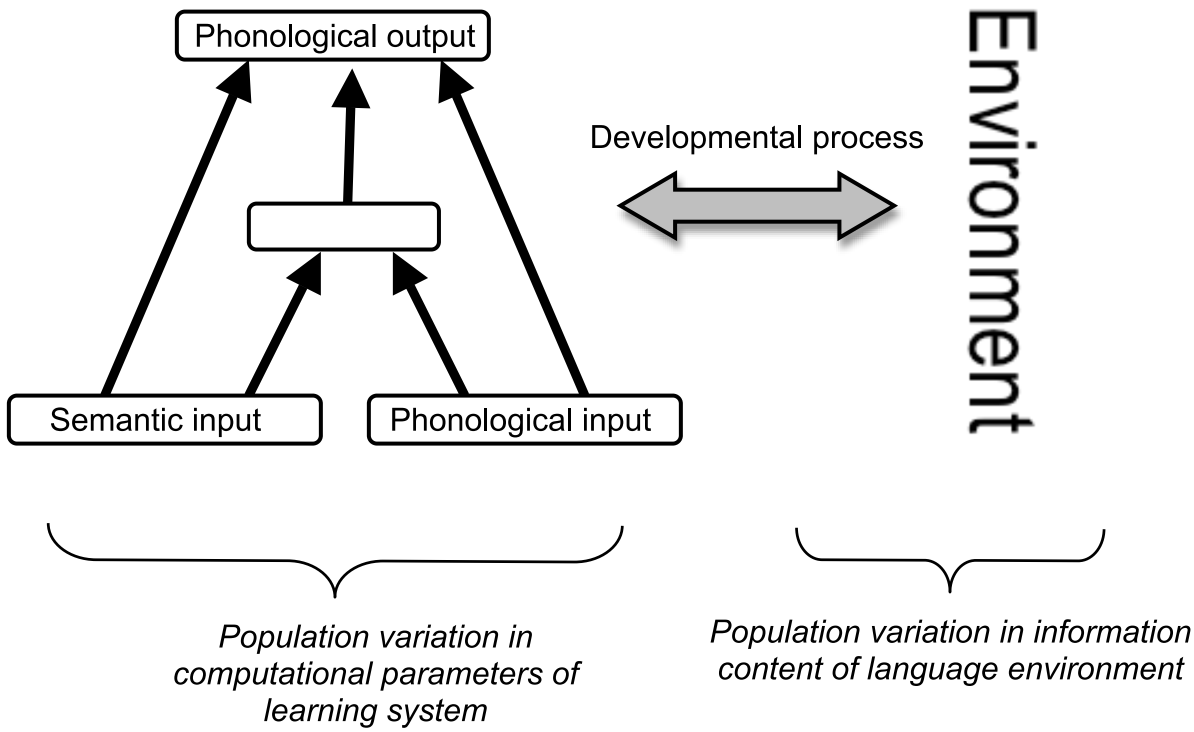

Phonological output
Developmental process
Environment
Semantic input
Phonological input
Population variation in computational parameters of learning system
Population variation in information content of language environment

Figure 4

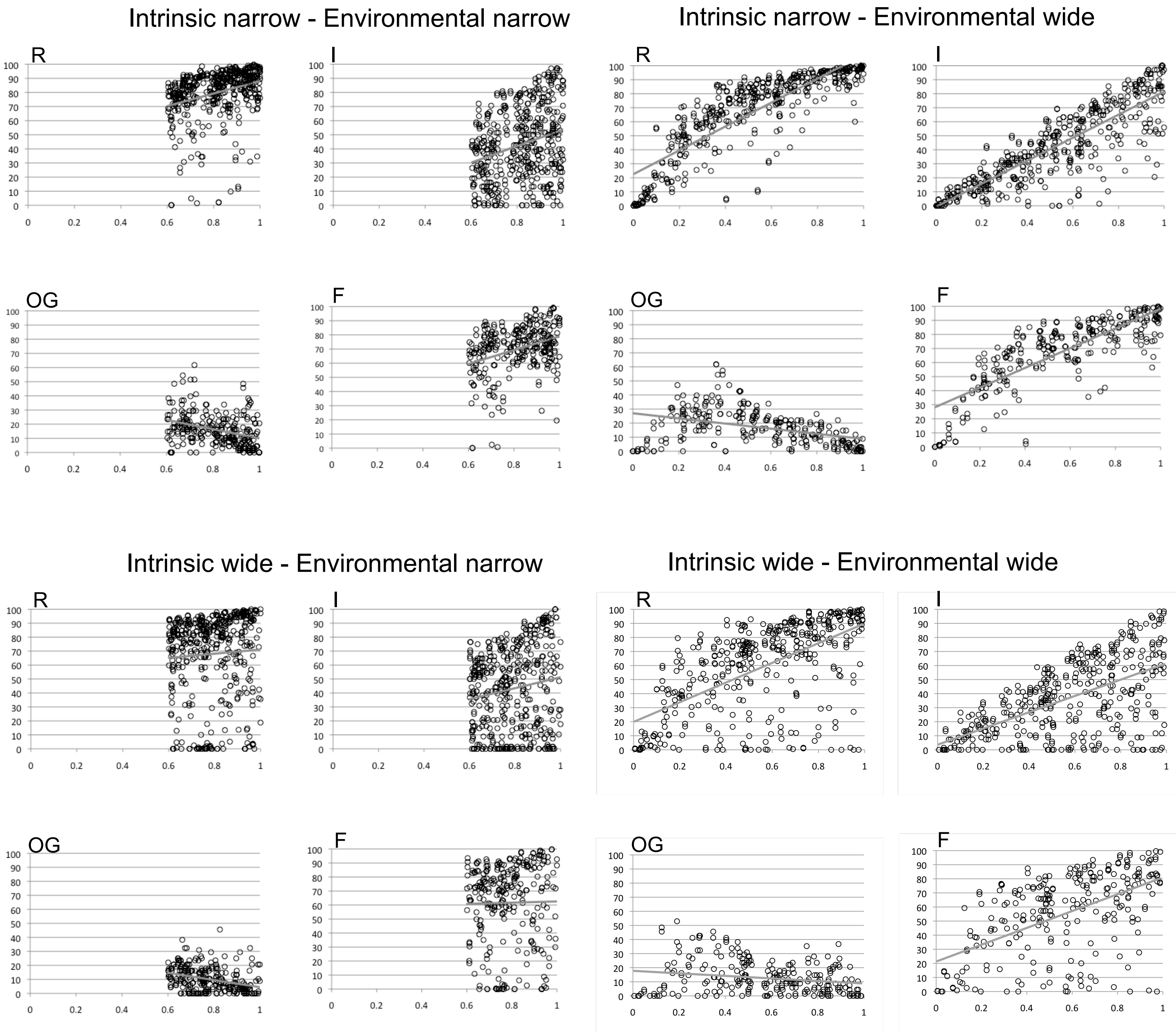
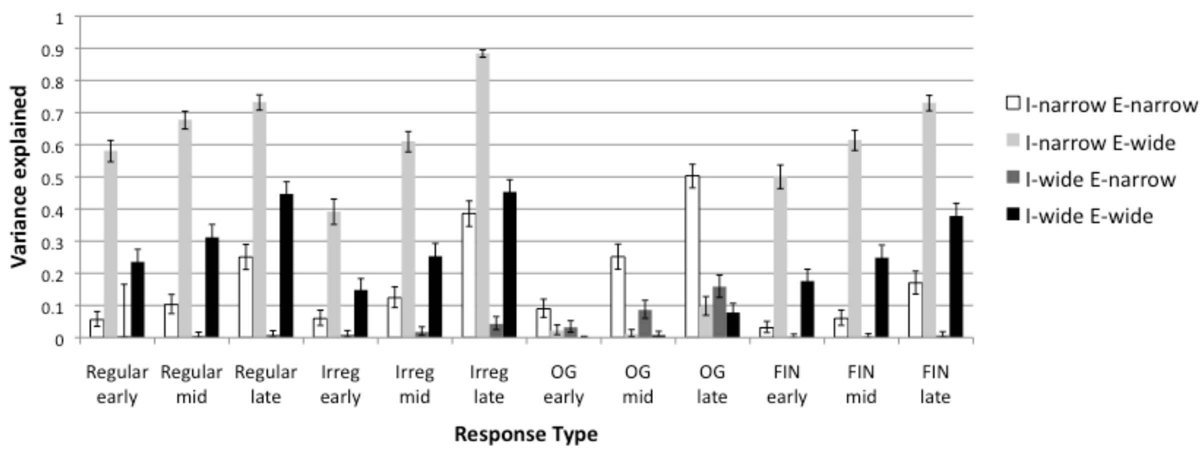

Intrinsic narrow - Environmental narrow
Intrinsic narrow - Environmental wide
R
I
R
I
OG
F
OG
F
Intrinsic wide - Environmental narrow
Intrinsic wide - Environmental wide
R
I
R
I
OG
F
OG
F

Figure 5


Variance explained
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Regular early
Regular mid
Regular late
Irreg early
Irreg mid
Irreg late
OG early
OG mid
OG late
FIN early
FIN mid
FIN late
Response Type
I-narrow E-narrow
I-narrow E-wide
I-wide E-narrow
I-wide E-wide