Modelling mechanisms of persisting and resolving delay in language development

Michael S. C. Thomas & V. C. P. Knowland

Developmental Neurocognition Lab, Birkbeck, University of London

Running head: Mechanisms of language delay

Contact author:

Prof. Michael S. C. Thomas
Developmental Neurocognition Lab
Centre for Brain and Cognitive Development
Department of Psychological Sciences
Birkbeck, University of London
Malet Street
London WC1E 7HX, UK
Tel.: +44 (0)20 7631 7386
Fax.: +44 (0)20 7631 6312
Email: m.thomas@bbk.ac.uk

Abstract

Purpose: Between a half and two thirds of cases of early-diagnosed language delay subsequently resolve to the normal range. This study employed neural network modelling to investigate the mechanistic basis of delay and to test the viability of the hypotheses that persisting delay (PD) and resolving delay (RD) lie on a mechanistic continuum with normal development.

Method: A population modelling approach was used to study individual rates of development in 1000 simulated individuals acquiring a notional language domain (here represented by English past tense). Variation was caused by differences in internal neurocomputational learning parameters, as well as the richness of the language environment, representing the effects of socioeconomic status (SES; Thomas, Ronald & Forrester, submitted). An early delay group was diagnosed and individual trajectories then traced.

Results: Quantitative variations in learning mechanisms were sufficient to produce PD and RD subgroups. SES did not predict the emergence of PD, but did predict the final ability levels of simulated individuals with RD. This novel prediction was supported by empirical data from Bishop (2005).

Conclusion: Computational modelling work suggests that persistent language delay is caused by limitations in processing capacity, while resolving delay is caused by low plasticity. Implications for language intervention are discussed. (201 words)

Keywords: Language delay, computational capacity, plasticity, socio-economic status, population modelling, artificial neural networks.

Around 15 % of children show delayed language development age 3-4 (Broomfield & Dodd, 2004); however, in a significant proportion of cases, perhaps as many as two thirds (Rannard, Lyons & Glenn, 2005), the delay resolves so that children subsequently fall in the normal range on standardised measures. Equally, many children showing early language delay go on to exhibit persisting deficits that have serious and long-term consequences for their education (Justice et al., 2009; Young et al., 2002), socialisation (see Durkin & Conti-Ramsden, 2010), mental health (Arkkila et al., 2008; Beitchman et al., 2001) and employability (Law et al., 2009). In order to target intervention to those children likely to suffer persisting deficits, it is desirable to identify early markers for language outcomes, one benefit of which is to maximise language to aid early literacy acquisition (Catts et al., 2002). However, identifying the markers of persisting deficits has proved problematic, indeed even given language measures at 18 months much of the variance in ability at 30 months is left unexplained (Henrichs et al, 2011), leaving clinicians forced to adopt a 'wait and see' policy rather than providing early intervention (Ellis & Thal, 2008). Moreover, little is understood about the developmental mechanisms that differentiate persisting from resolving language delay, with arguments made both for the importance of genetic and environmental factors. In this article, we use computational modelling methods to further a mechanistic understanding of the causes of persisting versus resolving language delay.

Dale, Price, Bishop and Plomin (2003) identified a sample of 802 2-year-old children who were at risk for language delay, based on parental reports of vocabulary, grammar, nonverbal ability, and use of language to refer to past and future events. These children were followed up at 3 and 4 years of age, again using parental measures. At 3 years of age, only 44.1% of this sample met criteria for persisting language difficulties, and at 4 years of age,

the proportion fell to 40.2%. Thus in over half of the cases, the early indicators of language delay had resolved. This pattern has been observed in a number of studies. Whitehurst and Fischel (1994) followed up a sample of 2-year-old children showing significant delay on an expressive vocabulary test and found that at 3½, 88% of the sample fell in the normal range on the same test. Rescorla, Dahlsgaard and Roberts (2000) examined the mean length of utterance of late talkers at 3 years of age and 4 years of age, and found that while 41% of the children scored above the 10th centile at 3, this figure had risen to 71% by age 4. Paul (1996) examined grammatical development in 2-3 year old children identified as late talkers and found that by age 4, 57% now fell in the normal range. The resolution of delay continues at slightly older ages. Bishop and Edmundson (1987) observed that 40% of children who showed language impairments aged 4 scored in the normal range by age 5. Bishop (2005) assessed 264 children identified as at risk for language impairment at age 4; only one third of these children met psychometric criteria for specific language impairment (SLI) at age 6. By age 7, however, greater stability is apparent in children's developmental trajectories: Conti-Ramsden et al. (in press) reported that language growth trajectories of 242 children with SLI followed longitudinally from ages 7 through 17 remained predominantly parallel to, and below, those of children in the normal range.

For their sample, Dale et al. (2003) explored whether it was possible to predict if children would fall in persisting (N=372) or resolving (N=250) delay groups based on their profiles at 2 years of age. Children whose delays would persist scored reliably lower across a number of parental rating measures, including vocabulary, grammar, reference to past and future events (displaced reference), and nonverbal skills, as well as showing reliably lower maternal education and a greater incidence of ear infection. Nevertheless, the effect sizes

were small (.01-.06) and differences were reliable only due to the large sample sizes involved. Logistic regression analyses found that children's profiles at age 2 offered only modest classification of outcome at age 4, with accuracy rates between 60-70% (where chance would be 50%); the derived function failed to detect the majority of children who would show a persisting deficit, and a substantial minority of children whose delay was predicted to resolve did not (see supplementary Table s.3). Thus although persisting and resolving groups differed marginally at diagnosis, it was difficult to predict outcome with any accuracy.

Causes of language delay

No clear picture has emerged from theoretical accounts of language delay as to why delay should resolve in some cases but not in others. Such accounts tend to differ on two dimensions: whether children with persisting and resolving delay form qualitatively or quantitatively different groups, and whether the relevant causal factors are genetic or environmental (see Bishop, Price, Dale & Plomin, 2003). For example, on the first dimension, Rice and colleagues (e.g., Rice, 2009) argue that persisting language deficits can be traced to the developmental impairment in a specialised system for acquiring morphosyntax, and therefore form a qualitatively different group. Resolving delay might then constitute the bottom of the distribution of normal variation of children without such a specific developmental impairment. This idea has gained suggestive support from genetic analyses implicating a potentially monogenetic cause of SLI (Bishop, 2005). By contrast, researchers such as Leonard (1987) and Rescorla, Dahlsgaard and Roberts (2000) have suggested that persisting and resolving delay are only quantitatively different; there is a single continuum of individual variation in rates of language development caused by the

same kinds of mechanisms, with persisting delay representing a more extreme case than resolving delay. This view is supported by the consistent finding that the strongest predictor of later language outcome, for example in vocabulary development, is earlier language performance (though even here the variance explained is usually less than 40%, e.g., Chiat & Roy, 2008; Henrichs et al., 2011). The differences found in later development may be exaggerated versions of those already present in early development.

With respect to the second nature-nurture dimension, a number of studies have implicated environmental factors in rates of language development. Nelson, Welsh, Vance Trup and Greenberg (2011) examined 336 4-year-old children living in poverty and found that a majority exhibited clinically significant language delays. To the extent that poverty is an environmental condition, this implicates environment in causing language delay. Hart and Risley (1995) observed different rates of language development in children from different socio-economic status backgrounds and linked these with large variations in the quantity of language spoken to the child. More recently, Henrichs et al. (2011) reported that SES as measured by maternal education was associated with both patterns of late-onset delay and persistent delay in population study of 3,759 toddlers. When Anushko (2008) analysed the development of language skills of 230 children at 15, 27, 37, and 72 months of age, she found that SES factors, the children's social-emotional competence, and the level of language exposure (as measured by book reading in the house) all reliably predicted rate of increase in expressive and/or receptive language (see also, Anushko, Jones & Carter, 2009). Notably, when children were initially split between low and high performing groups, those who were able to accelerate from low to high groups had significantly more exposure to and experience with language through book reading activities, compared to their peers who

remained in the low growth group across the time points. The effect of input has also been captured as an increasing SES gap over developmental time in a number of longitudinal studies (Reilly et al., 2010; Rowe, Raudenbush & Goldin-Meadow, 2012). Bishop et al. (2003) utilised behavioural genetic methods with their sample, which comprised monozygotic and dizygotic twin pairs, to explore the aetiology of persisting and resolving delay. The results indicated similar and modest heritabilities of .25 for both groups, implicating environmental factors in the cause of delay. (An analysis that split children by whether or not there was professional involvement of clinicians at aged 4 created two groups for which the heritability was 0 and .4, respectively, pointing to heterogeneity in the sample and cases of stronger or weaker environmental causes).

By contrast, other researchers have argued that environmental factors play little role in language delay. In a sample of 1766 children, Zubrick et al. (2007) found that SES, family, and maternal characteristics did not predict language delay at age 2. The strongest predictors were family history for late language, male gender, and early neurological problems. In a smaller longitudinal study, Rice, Wexler and Hershberger (1998) did not find any predictive power of maternal education, a marker for SES, on the growth of inflectional morphology in children with SLI or typically developing controls. Moreover, when Dale et al. (2003) added level of maternal education to their logistic regression analysis, it failed to improve their ability to predict whether children's language delays would persist or resolve. Overall, while children diagnosed with early language delay are clearly heterogeneous, the dimensions defining the heterogeneity remain unclear (Desmarais et al., 2008).

Part of the challenge stems from the current limited understanding of the mechanisms that might cause delay. Delay is most often used descriptively rather than

mechanistically, to recognise the fact that the behaviour in some target group resembles that of younger, typically developing children (Thomas et al., 2009). At a mechanistic level, maturational views of delay have been most clearly articulated. These characterise language development as analogous to biological growth, and variations in rates of growth as reflecting differences in (putative) genetically controlled timing mechanisms (e.g., Rice, 2009, for such an account in the context of the development of morphosyntax). Experience-dependent views of delay are less frequently articulated, but these presumably entail either a language system that receives fewer learning experiences, or a learning system that is less malleable, such that more experience is required to effect a change in behaviour. To demonstrate the preliminary nature of the current mechanistic understanding of delay, Thomas, Karaminis and Knowland (2010) recently formulated six predictions based on a theory of 'slow' development[1] and argued that these predictions are rarely articulated or tested, and in some cases are most probably false.

Computational modelling to investigate mechanism

One way to address the superficial consideration given to mechanisms of delay is through the use of computational modelling. The computational modelling of developmental systems can potentially serve a number of roles in this context (Mareschal & Thomas, 2007). By virtue of implementation, modelling can advance the detail with which theoretical accounts are specified. Models that embody theoretical proposals can then test the viability of those proposals to account for the empirical data. Models can show how a single mechanistic account can unify a range of previously disparate empirical phenomena. And models can generate novel predictions that can then be evaluated against empirical data.

In this paper, we consider a computational model addressed to the differences between persisting and resolving language delay. The model takes advantage of a new approach called *population modelling* (Thomas et al., 2013). Whereas most previous models of cognitive and language development have been applied to capturing the development of the average child (and sometimes to capturing the profile of groups of children with particular deficits), in population modelling, the aim is to simulate a large population of individuals undergoing a developmental process. In this population, multiple intrinsic and extrinsic properties are varied across individuals (where intrinsic properties refer to the computational abilities of each system, and extrinsic properties refer to the quality of the learning environment to which it is exposed). In combination, these factors produce a distribution of performance as the population acquires the target behaviour. Potentially, atypical conditions can be applied to individuals against this background of variation. The framework permits both the study of individual differences in the rates of development and also the investigation of how cases of qualitatively atypical development may differ from the normal population. For example, Thomas, Ronald and Forrester (submitted) used a population modelling approach to investigate causal pathways by which SES may modulate rates of language development; Thomas, Knowland and Karmiloff-Smith (2011) used the approach to consider candidate neurocomputational mechanisms that might cause developmental regression in autism.

In the current work, we employed the modelling framework of Thomas, Ronald and Forrester to pursue the following four aims: (1) to establish whether a quantitative account of developmental variations in a population is sufficient to generate subgroups that demonstrate persisting delay and resolving delay, or whether qualitative differences are

necessary; (2) to evaluate whether there are differences in the behavioural profiles of these subgroups when delay is first diagnosed, which can predict subsequent developmental outcomes; (3) to assess the role of environmental variation (in this case, a proxy for SES) in causing developmental delays or aiding their resolution; (4) in implemented simulations, to investigate the mechanisms responsible for producing cases of persisting versus resolving delay.

Method

Thomas, Ronald and Forrester (submitted) employed a connectionist model of development to explore the adequacy of manipulations of *environmental information* to simulate SES effects in English past-tense acquisition, in a data set provided by Bishop (2005). The simulations sought to implement both environmental and genetic/intrinsic sources of individual differences; and in doing so succeeded in capturing both the qualitative patterns of regularity effects in population performance and the predictive power of SES observed in the empirical data. SES was successfully implemented as variability in the size of the training set available to each network. The model architecture is depicted in Figure 1.

In the current context, the modelling framework was utilised in a more illustrative setting, as an example of a developmental system applied to the problem of extracting the latent structure of a language domain through exposure to a training environment (although a prediction of the model was subsequently tested against past-tense data from the Bishop [2005] sample). The model was used in a more illustrative setting because individual differences in the early phases of acquisition in the model, and their subsequent developmental outcomes, were compared to the qualitative patterns of data identified in the Introduction. Most often, however, the measures of early language delay in children are

based on parental reports of vocabulary acquisition, while the later assessments of language performance span a wider range of measures.

==========
Insert Figure 1 about here
==========

Simulation overview

A population of 1000 artificial neural networks were exposed to the language domain (English past tense) and their developmental trajectories were analysed. Two sources of variation caused individual differences in their rates of development. *Extrinsic variation* was encoded by altering the quality of the learning environment, and specifically the amount of information available in the input. This can be thought of as equivalent to the effects of variation in SES on language input (Hart & Risley, 1995). *Intrinsic variation* was encoded by altering the quality of the learning mechanism. This was implementing by variations in 14 neurocomputational parameters controlling the artificial neural network construction, activation, adaptation, and maintenance. These parameters can be viewed as serving different types of processing role within the network. Parameters affect the network's *learning capacity*, *plasticity*, *quality of signal*, as well as possible *regressive events* (although some parameters contribute to more than one role). In the results, we interpret the contributions of neurocomputational parameters to types of delay in terms of these roles, together with the effect of the *learning environment*. From the population of 1000 simulated individuals, early performance on regular verb acquisition was used to define a delay group and their subsequent progress was traced with reference to the population normal range to identify different possible outcomes. A detailed description of simulation methods can be found in the supplementary materials.

Results

Defining delay

Five time points were defined in the development of the population, when the population accuracy for regular verb production was 40, 50, 60, 65, and 78%. These occurred at 31, 49, 84, 127, and 500 epochs of training, respectively. Figure 2(a) shows the distribution of regular performance for time 1, at which point individuals were identified as exhibiting developmental delay if their performance fell more than 1 standard deviation below the population mean. This corresponded to 28.7% of the population. Family quotient (henceforth FQ) served as the proxy for the operation of SES effects on acquisition. In that sense, the time 1 delay group comprised 76 individuals from the lowest SES quartile, 74 and 79 from the middle quartiles, and 58 from the upper quartile. Although highest SES quartile had the fewest delayed individuals, the distribution was not reliably different from chance ($\chi^2(3)=3.69$, $p=.297$). Developmental delay was then re-diagnosed at each subsequent time point, shown in Figure 2(b) to (e). Figure 3 shows the proportion of the population diagnosed with delay at each time point. A small number of those delayed at the final time point (18) were not delayed at time 1. Late onset delay is also a prominent pattern reported in the literature (e.g., Henrichs et al., 2011; Ukoumunne et al., 2011). Of those delayed at the first time point, 118 showed a delay that persisted through to the final time point, while the delay resolved in 169 (or 58.9%) of the cases.

==========
Insert Figures 2 and 3 about here
==========

In terms of intrinsic and extrinsic influences on development, simulated individuals differed from each other only quantitatively, yet these quantitative differences were

sufficient to produce persisting and resolving delay groups. Indeed, the proportion of cases that resolved was broadly similar to the rates observed in empirical studies (Bishop & Edmundson, 1987; Bishop, 2005; Dale, Price, Bishop & Plomin, 2003; Paul, 1996; Rescorla, Dahlsgaard & Roberts, 2000; it was, however, less than that observed by Whitehurst & Fischel, 1994).

If persisting and resolving delay lie on a strict developmental continuum, one might expect cases of resolving delay to lie towards the bottom of the normal range. That is, resolving cases would slip into the normal range but would still perform relatively poorly, and not therefore represent complete resolution of the delay. The population rank orders of individuals in the resolving group were examined to evaluate this idea. Of those showing resolving delay, 80% (136 individuals) indeed remained in the bottom 500 of the population. However, in some individuals, performance at time 5 was somewhat better: 17% (28 individuals) had a rank order in the top 500, and a few (3%, 5 individuals) even finished in the top 200. The outcome of resolving delay was therefore variable. Good final outcomes were possible, suggesting that in some cases, delay could completely resolve. Examples of individual developmental trajectories for the five groups, typically developing, persisting delay, resolving delay with low outcome, resolving delay with good outcome, and resolving delay with very good outcome, are depicted in Figure 4. Each plot also contains the five time points, and the mean developmental trajectory for the population as a whole.

==========

Insert Figure 4 about here

==========

Predicting persisting vs. resolving delay from time 1 behavioural profiles (and SES)

The behavioural profile of simulated individuals in the persisting and resolving delay groups were compared at time 1. The profile initially included 9 measures of various aspects of past

tense performance on training and generalisation sets. Behaviour on 3 of these measures summarised the pattern: regular verbs, vowel-change irregular verbs, and regularisation of novel verbs. At time 1, the persisting delay group performed reliably worse on regular verbs and novel verbs than the resolving group, but there was no difference on irregular verbs (MANOVA for overall profile difference: $F(3,283)=12.22$, $p<.001$, $\eta_p^2=.115$; individual measures: regular: $F(1, 285)=36.84$, $p<.001$, $\eta_p^2=.114$; irregular: $F(1, 285)=2.29$, $p=.131$, $\eta_p^2=.008$; novel: $F(1, 285)=28.00$, $p<.001$, $\eta_p^2=.098$). Although the differences were highly reliable, they were of small effect size.

The time 1 behavioural differences between persisting and resolving groups were small, but were they sufficient to reliably predict outcome group? Logistic regression analyses were used to predict outcome group based on the profile of performance on regular, irregular, and novel verbs. The results are contained in Table 1, which also includes the results of Dale et al.'s (2003) analyses that sought to predict delay outcome based on time 1 measures of verbal ability, displaced reference, nonverbal ability, and maternal education (as a marker for SES). Dale et al. found that small initial behavioural differences between groups were not sufficient for accurate classification of delay outcome; and that the addition of maternal education did not markedly increase predictive power. Similarly, the simulation data indicate that the small behavioural differences at time 1 did not produce accurate prediction of outcome; and addition of the SES proxy as a predictor produced no marked improvement. (The prediction equations for empirical data and model were poor in different ways, with the data equation over-predicting the resolution of delay and the model equation over-predicting the persistence of delay).

The bottom row of Table 1 shows the predictive power when each simulated individual's neurocomputational parameters were added to the logistic regression: accuracy of classification was now much higher, but still some way short of 100%. This illustrates the operation of stochastic factors in the model, and the non-linear interactions that occur between parameters in determining learning ability. Notably, addition of the neurocomputational parameters to the analysis improved the ability to predict which networks would have resolving delay from a position of previously over-predicting persisting delay. This suggests we may find neurocomputational parameters that are markers for the resolution of delay.

Finally, it is worth noting that even within a single computational learning system and using noise-free performance measures, the correlation between population performance at different time points becomes smaller as the time points become more remote. Time 1 performance on regular verbs predicted only 37% of the variance in time 5 performance. The result is consistent with the empirical finding that early language performance is the best predictor of later language performance, but the variance explained can be relatively modest (e.g., Chiat & Roy, 2008; Henrichs et al., 2011).

==========

Insert Table 1 about here

==========

Mechanistic explanations for persisting vs. resolving delay

To explore the mechanistic basis of the distinction between persisting and resolving delay groups, we carried out two complementary sets of analyses, using neurocomputational parameter values to predict outcome group either using multivariate analysis of variance or multinomial logistic regression. The supplementary materials contain three tables of

statistical results, incorporating a statistical comparison of the mean neurocomputational parameter values for simulated typically developing, persisting delay and resolving delay groups; equivalent results for a comparison of the resolving delay group, split by whether the final outcome was low (bottom 500 of population), good (top 500), or very good (top 200); the mean parameter values per group; and three case studies of individual parameter sets, which demonstrate the extent to which individuals conform or diverge from group averaged effects.

Both delay groups differed from the typically developing group across a range of neurocomputational parameters (Table s.1). The strongest effect size for the difference between typical and persisting delay was the power of the learning algorithm. The strongest effect size for the difference between typical and resolving delay was the learning rate. Individually, the delay groups did not differ from the typically developing group on the SES proxy. However, when combined, there was a small, marginally significant difference, with delay groups showing lower family quotient values ($t(998)=1.93$, $p=.054$, Cohen's $d=.136$; see Figure 5).

==========

Insert Figure 5 about here

==========

The persisting delay group differed from the resolving delay group over a smaller number of parameters. Individuals showing persistent delay tended to have fewer hidden units, a higher pruning threshold (leaving the network at greater risk of connection loss across development), a less powerful learning algorithm, and higher processing noise. In terms of processing roles, the more salient cause of delay in this group was lower computational *capacity* and poorer *signal*. By contrast, individuals showing resolving delay had a shallower unit activation function and a lower learning rate in the semantic pathway.

In terms of processing roles, the more salient cause of delay in this group was lower *plasticity*.

Table s.2 indicates which parameters predicted the final outcome for individuals showing resolving delay. Outcome depended on the two previously identified plasticity parameters, semantic pathway learning rate[2] and unit threshold function. A lower semantic-pathway learning rate was associated with poorer final outcome, while a shallower unit threshold function was associated with better final outcome. As causes of resolving delay, these parameters had differential effect on the potential final level that could be achieved. Most notably, however, the final level of performance was associated with the SES proxy, the family quotient parameter. The richer the environment, the higher the final level that could be achieved. This pattern emerged despite the relatively weak contribution of the SES proxy in explaining individual differences in the population as a whole (e.g., at time 5, family quotient predicted only 2.2% of the variance in regular verb performance in the full population).

In summary, a consideration of mechanisms suggests the following picture: both persisting and resolving delay are caused by a combination of suboptimal learning parameters. Most salient in persisting delay is a limit on the computational capacity of the learning system, which places a ceiling on the highest level that can be achieved. Most salient in the resolving delay is lower plasticity, which reduces the rate of learning but does not place the same ceiling on the highest level that can be achieved. To some extent, the final level is then determined by the richness of the environment in which the learning system is embedded. By contrast, the richness of the learning environment is much less

relevant to learning in reduced capacity systems. While early on, the delay groups are conflated, capacity places a limit on subsequent learning in a way that plasticity does not.

Late onset delay

A small number of simulated individuals showed late onset delay (N=18), falling within the normal range at time 1 but below the normal range at time 5. This pattern has been observed empirically (e.g., Henrichs et al., 2011; Ukoumunne et al., 2011). While it is not the focus of this paper, a brief analysis is merited. Late-onset delay appeared to comprise two sub-groups. In one group, late performance was limited by the poverty of the information in the environment. In the other group, late performance was impacted by regressive events that caused excessive pruning to network connections. Regressive events within the current modelling framework are analysed in Thomas, Knowland and Karmiloff-Smith (2011). The two groups were identified based on their family quotient value, where group 1 fell below the population mean of 0.8 (N=6) and group 2 fell above (N=12), with respective family quotient means of 0.75 and 0.90. The two groups then differed in the parameter that determined late onset connectivity pruning, with group 2 showing reliably more aggressive pruning than group 1 (and the rest of the population), implicating late onset pruning as the cause; and group 1 showing no difference in the pruning parameter compared to the rest of the population, implicating the limits of a deprived environment as the cause of late onset delay [group 1 vs. group 2 pruning parameter: $t(16)=3.37$, $p=.004$, Cohen's $d=1.79$; group 1 vs. population: $t(986)=.029$, $p=.977$, $d=.01$; group 2 vs. population: $t(992)=8.72$, $p<.001$, $d=2.53$]. The model therefore predicts that late onset delay is a heterogeneous group with both intrinsic and extrinsic causes.

Testing a novel prediction of the model

As indicated above, the model generated a novel prediction that the quality of the environment, as assessed by SES, should reliably predict outcome in the resolving delay group but not the persisting delay group. In this section, we test this prediction using a data set from Bishop (2005)[2]. Bishop (2005) analysed data from the large British sample of twins considered in Dale et al. (2003) and Bishop et al. (2003). Bishop (2005) identified a sample of the twins who exhibited language delay risk at 4 years of age. These children, along with a sample of twins not identified as at risk, were tested at 6 years of age on a test of English past tense production. At 6 years of age, around one third of the early language impairment risk group then met psychometric criteria for SLI, compared to one in ten of those not identified as at risk (Bishop et al., 2006). From these children, three groups could be identified: 94 6-year-old children both exhibiting language impairment risk at 4 years and meeting psychometric criteria for SLI at 6 years (persisting delay); 104 6-year-old children exhibiting language impairment risk at 4 years of age but not meeting psychometric criteria for SLI at 6 years (resolving delay), and 166 children exhibiting neither language impairment at 4 nor SLI at 6 (typical development). Crucially, SES data were also available for these children's families (see supplementary materials for further details). Mean SES values for the groups are shown in Figure 5.

The key comparison was whether there was a relationship between SES and performance in the resolving group but not the persisting group for this empirical data set. Table 2 compares the results of linear regressions between SES and past-tense performance for the Bishop (2005) sample. The persisting group shows no sign of a relationship, while for the resolving group, there is a weak trend: the higher the SES value, the better the regular past tense performance. A comparison of these two relationships yielded a reliable

interaction between persisting and resolving groups, SES, and regular verb performance, whereby there was a reliably stronger relationship between SES and performance in the resolving group than the persisting group ($F(1,194)=4.015$, $p=.047$, $n_p^2=.020$). Finally, addition of irregular verb performance from the Rice-Wexler (2001) test increased sensitivity, with fewer ceiling scores. In this case, both typically developing and resolving delay groups demonstrated a reliable relationship between SES and past-tense performance, while the persisting group did not.

==========

Insert Table 2 about here

==========

In sum, a test of the novel prediction of the model through available empirical data produced support for the model. Despite the weak overall predictive power of SES on performance (see, e.g., Rice, Wexler & Hershberger, 1998), the resolving group showed a stronger relationship with SES in their past-tense performance than the persisting group.

Discussion

The current computational model was successful in demonstrating that in a population of developing systems, which varied along continua of intrinsic and extrinsic parameters, early-diagnosed delay resolved in some individuals but persisted in others when exposed to a language domain. The proportion of resolving cases was similar to that observed in empirical studies of early language delay (Bishop & Edmundson, 1987; Bishop, 2005; Dale, Price, Bishop & Plomin, 2003; Paul, 1996; Rescorla, Dahlsgaard & Roberts, 2000; it was, however, less than that observed by Whitehurst & Fischel, 1994). The model captured this pattern through quantitative variations in learning parameters between individuals, demonstrating the viability of the proposal of Leonard (1987) and Rescorla, Dahlsgaard and

Roberts (2000) that resolving and persisting delay lie on a continuum of individual variation in rates of language development.

The model qualitatively accorded with six further empirical findings. First, as per Dale et al. (2003), there were small differences in the behavioural profiles of persisting delay and resolving delay groups when delay was first diagnosed. In the model, these were differences in extracting the latent structure of the language domain (the regular past tense 'rule'). Second, also as per Dale et al. (2003), these small behavioural differences were not particularly effective in predicting individual outcomes. Third, the model included a manipulation equivalent to variations in SES (Thomas, Ronald & Forrester, submitted); this manipulation only accounted for a small amount of the variance in past-tense formation, similar to the findings of Rice, Wexler, and Hershberger (1998), and in the larger dataset of Bishop (2005). Fourth, as with Dale et al. (2003), addition of this SES measure to early behavioural differences did not improve the ability to predict delay outcomes. Similar to the data of Bishop (2005), delay groups showed slightly lower SES scores than the typically developing group, implicating environmental factors (to some minor extent) in the cause of delay. Fifth, the model simulated cases of late onset delay, which have also been observed in the literature (Henrichs et al., 2011; Ukoumunne et al., 2011). Lastly, despite the apparently small influence of the SES manipulation on the simulated population, the model generated a novel prediction that SES should reliably predict the outcome of individuals with resolving delay but not those with persisting delay; this novel prediction was subsequently supported by the empirical data of Bishop (2005). The model indicated that while the majority of individuals with resolving delay finished in the low normal range, some individuals in a rich environment finished in the top half or top fifth of the population. While resolving delay fell

between persisting delay and typical development on a mechanistic continuum, some cases of early delay could truly completely resolve.

The advantage of a computational model is that one can examine the mechanisms responsible for (re)producing behavioural patterns. Variations in rates of development were generated by small differences in a relatively large number of neurocomputational parameters (14) within an artificial neural network learning system. The majority of these parameters were implicated in causing delay. Variations in these parameters overlapped between persisting and resolving delay groups, but broad differences could be discerned between delay types. Parameters were identified by computational roles, including those of *capacity*, *plasticity*, *signal*, and *environment*. Persisting delay was more strongly associated with limits in capacity, as well as noise in the processing signal. For these networks, acquisition of the problem domain was restricted as a result of reduced processing resources that put a limit on the amount and complexity of information that could be learned. Resolving delay was more strongly associated with low plasticity, that is, the speed of learning and the responsiveness of the system to inputs; and with signal limitations whereby precise outputs were required to drive responses, which were not achieved until later in learning. The environment influenced resolving delay outcomes because low plasticity eventually allowed the learning system to take advantage of richer information available in the environment, while for persisting delay, capacity limitations made networks insensitive to the information available in richer environments.

The broader aim of the modelling work was to advance a mechanistic understanding of delay, which hitherto has been ascribed either to unspecified maturational processes or remained as a largely descriptive notion. Population modelling was used to address

individual differences within a developmental framework. Individual variation in rates of development was caused by a combination of intrinsic influences (neurocomputational parameters) and extrinsic influences (the richness of the learning environment). The result was a population exhibiting a normal distribution of performance, which shifted across development and which was skewed by early floor effects and late ceiling effects (see Figure 2). All parameters, as well as the learning environment, were modelled as varying independently. Delay was caused by an accumulation of suboptimal learning parameters in unlucky individuals. Through non-linear interactions between parameters, different constellations of poor parameters could lead to subtly different behavioural profiles: delay was heterogeneous in detail (as illustrated by the persisting versus resolving patterns), despite its quantitative origins. Appendix A outlines three cases studies showing that even individuals within persisting and resolving groups could show minor differences to the overall group patterns.

The population modelling approach presented here contrasts with previous computational models of atypical development, which have simulated disorders by the manipulation of single parameters, while other learning parameters were held constant. For example, Joanisse (2004) simulated deficits associated with SLI in inflectional morphology by the addition of processing noise to phonological representations in a connectionist network, while Thomas (2005) captured similar empirical data by altering the unit threshold function in a past-tense model (see Karaminis, 2011, for a more general consideration of how neurocomputational processing limitations can lead to behavioural symptoms of SLI in inflectional morphology, syntax comprehension and syntax comprehension, in a cross-linguistic context.) The current simulations are more consistent with the quantitative trait

loci (QTL) approach within behavioural genetics, which argues that normal and abnormal behaviour lie on the same continuum of genetic variation; and that, with the exception of known genetic mutations of large effect, many behaviourally defined disorders represent the chance accumulation of a large number of common gene variants each carrying a small risk for disorder, accumulations that will inevitably occur in large populations (Plomin, DeFries, McClearn & McGuffin, 2008; Kovas, Haworth, Dale & Plomin, 2007). In the remaining paragraphs, we focus on the implications of the model, should it turn out to be correct. Strengths and weaknesses of the model are considered in detail in the supplementary materials.

Implications

SES was simulated by a manipulation of the information content of the learning environment (Thomas, Ronald & Forrester, submitted). SES effects on delay were weak, in line with their limited effects in the population overall. Nevertheless, simulated individuals showing resolving delay who finished with good final outcomes were associated with richer learning environments. This accords with Anushko's (2008) finding that in a sample of 230 children whose language trajectories were followed between 15 months and 6 years of age, those children accelerating from the lower performing group to the higher performing group had significantly more exposure to and experience with language through book reading activities, compared to peers in the low growth group. The model suggests a clearer framing of the role of environmental input: it is *not* the cause of early language delay; delay is the result of the intrinsic property of low plasticity; but where low plasticity is the cause of delay, greater experience with language can maximise subsequent outcomes.

As with the findings of Dale et al. (2003), the modelling results did not generate optimism that early differences in behavioural profiles could predict the outcome of early diagnosed language delay. Reliable differences were found, with individuals whose delays resolved being better able to extract the latent structure of the language domain to which they were exposed, but as with Dale et al., these early differences were not sufficient to usefully predict individual outcomes.

The model suggested that the causes of language delay were limitations in intrinsic neurocomputational processing properties (low capacity for persisting, low plasticity for resolving delay). How can this be reconciled with the twin-study findings of Bishop et al. (2003) of lower heritability in language scores for cases of resolving delay, higher heritability for language scores for cases of persisting delay, and significant environmental involvement in both cases?

So far, the sole extrinsic factor in the model was the proxy for SES, that is, the richness of the language environment. However, extrinsic factors could be aligned with the properties of the model in a different way. In particular, it could be that some neurocomputational properties of the learning system are influenced by environmental conditions, rather than being purely inherited dimensions of individual variability. The current results could be reconciled with the findings of Bishop et al. (2003) if shared environment factors have a direct effect on some intrinsic factors but not others; and specifically, if shared environment factors were to have a direct effect on plasticity (through factors such as attention, motivation, reward, stress, nutrition), while computational capacity were determined more by heritable factors involving prenatal phases of brain development (for mothers who have adequate nutrition, no toxins, and no viral infections, reducing

environmental factors in these phases). This idea gains some suggestive support by findings from Rowe, Raudenbush and Goldin-Meadow (2012) that growth trajectories in vocabulary development are more strongly predictive of language outcomes for children from low compared to high SES backgrounds. One possibility is that this reflects opportunities for higher SES children to develop linguistic competence through changes in factors related to plasticity, whilst the greater predictive power for low SES background children reflects more developmentally consistent capacity.

It is important to note, here, that SES may bear a different relationship to causal factors depending on the absolute level that is being considered. It may be that from lower middle class upwards, the influence of SES on language development operates mainly via the information content of the environment, but as poverty and deprivation increase, SES operates via an influence on neurocomputational processing properties; such that the most fundamental constraint is primary. This would explain why Nelson et al. (2011) identified so many cases of clinical language delay in the sample of children in poverty that he considered, while at the same time Zubrick et al. (2007) determined that persisting delay was best predicted by family history of late language, male gender, early neurological growth, rather than environmental factors.

Two key issues emerge in attempting to integrate the disparate findings in this field. First, how do the measures of SES that are typically collected and entered into analyses relate to the causal pathways by which environmental factors operate? Does a measure like maternal education or book reading better index the information content of the environment (as captured here), while a measure like income or free school meals better index nutrition and stress, and thereby influences on neurocomputation? Second, how is the operation of

environmental effects on rates of language development determined by the *absolute* level of SES under consideration? Perhaps the same measure indexes different causal pathways at different absolute levels.

Finally, three future directions are desirable to extend the current work. First, modelling should be expanded to consider multiple systems within language, thereby allowing early delay to be diagnosed on different behaviours than those considered later in development; and for the effects of SES on different levels of language to be explored (e.g., see Noble, Norman & Farah, 2006). Second, in cases of persisting delay, modelling work should investigate whether and how tailored learning environments might enhance performance of learning systems that are limited by virtue of certain learning parameters, such as reduced capacity. Persisting delay may require exposure to altered learning environments in order to maximise performance given internal constraints. This might involve working to simplify linguistic environments to match the constraints of the system. By contrast, resolving delay would be best addressed by exposure to a rich (but otherwise normal) language-learning environment. In this case, environmental alterations might work to increase the quantity of both simple and complex language available to the child. Both these approaches could be adopted within traditional speech and language therapy frameworks but might shift the focus away from attempts to alter properties of the system, that is, work with the child, and towards attempts to alter the language environment to either stimulate the system or match it. Third, with respect to intervention, a consideration of neurocomputational mechanisms suggests a potential point of contact with developmental cognitive neuroscience. Even if behavioural profiles are not sufficiently predictive of delay outcome, perhaps markers of capacity limitations may be discerned from those of low

plasticity via brain imaging methods, for example by the use of electrophysiology and event-related potentials. This would allow the targeting of language intervention to children with limited processing capacity, while cases of low plasticity could be exposed to richest language environments to achieve the best developmental outcomes for these children.

Acknowledgements

This research was supported by ESRC grant RES-062-23-2721 and a Leverhulme Study Abroad Fellowship held at the University of Chicago.

Notes

1. The six predictions were: (1) for any domain with a ceiling performance level, individuals with delay should eventually catch up; (2) in those sensory domains with sensitive periods, the periods should be extended (for example, specialisation to the phonemic contrasts of one's own language); (3) in domains where there is specialisation of function, this specialisation should also emerge later (for example, in face recognition, for faces presented in an upright orientation); (4) there should be identical quality of processing when individuals are matched for performance level, where quality is assessed by the effect of implicit variables such as frequency, imageability, similarity, and so forth; (5) if delay is argued to be widespread across the cognitive profile, the reduction in developmental rate should be the same across all cognitive domains, since the same mechanism cannot obviously explain many different delays (other than *post-hoc*); (6) if delay is argued to be focal, under no developmental theory should any other cognitive system rely on the affected component for its own successful development.
2. We are grateful to Dorothy Bishop for making the raw data available to us.

References

Anushko, A. E. (2008). *Multi-domain predictors of trajectories of language development in early childhood*. Unpublished PhD Thesis, Fordham University, New York.

Anushko, A. E., Jones, S. M., & Carter, A. S. (2009). Multi-domain predictors of trajectories of language development in early childhood. Poster presented at *Society for Research in Child Development*, April 2-4 2009, Denver, USA.

Arkkila, E., Rasanen, P., Roine, R. P., & Vilkman, E. (2008). Specific language impairment in childhood is associated with impaired mental health and social well-being in adulthood. *Logoped Phoniatr Vocol, 33* (4), 179-189.

Beitchman, J. H., Wilson, B., Johnson, C. J., Atkinson, L., Young, A., Adlaf, E., Escobar, M., & Douglas, L. (2001). Fourteen-year follow-up of speech/ language impaired and control children: psychiatric outcome. *Journal of the American Academy of Child and Adolescent Psychiatry, 40* (1), 75-82.

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.

Bishop, D. V. M., & Edmundson, A. (1987). Language-impaired four-year-olds: Distinguishing transient from persistent impairment. *Journal of Speech and Hearing Disorders*, *52*, 156-173.

Bishop, D. V. M., Laws, G., Adams, C., & Norbury, C. F. (2006). High heritability of speech and language impairments in 6-year-old twins demonstrated using parent and teacher report. *Behavior Genetics*, *36(2)*, 173-184.

Bishop, D. V. M., North, T., & Donlan, C. (1995). Genetic basis of specific language impairment: Evidence from a twin study. *Developmental Medicine and Child Neurology*, *37*, 56-71.

Bishop, D. V. M., Price, T. S., Dale, P. S., & Plomin, R. (2003). Outcomes of early language delay: II. Etiology of transient and persistent language difficulties. *Journal of Speech, Language, and Hearing Research*, *46*, 561-575.

Broomfield, J. & Dodd, B. (2004). Children with speech and language disability: caseload characteristics. *International Journal of Language and Communication Disorders, 39*(3): 303-324.

Catts, H. W., Fey, M. E., Tomblin, J. B., & Zhang, X. (2002). A longitudinal investigation of reading outcomes in children with language impairments. *Journal of Speech, Language and Hearing research, 45*, 1142-1157.

Chiat, S., & Roy, R. (2008). Early phonological and sociocognitive skills as predictors of later language and social communication outcomes. *Journal of Child Psychology and Psychiatry*, *49*, 635-645.

Conti-Ramsden, G., St Clair, M. C., Pickles, A., & Durkin, K. (in press). Developmental Trajectories of Verbal and Nonverbal Skills in Individuals with a History of SLI: From Childhood to Adolescence. *Journal of Speech, Language, and Hearing Research*.

Dale, P. S., Price, T. S., Bishop, D. V. M., & Plomin, R. (2003). Outcomes of early language delay: I. Predicting persistent and transient language difficulties at 3 and 4 years. *Journal of Speech, Language, and Hearing Research*, *46*, 544-560.

Desmarais, C., Sylvestre, A., Francois, M., Bairati, I., & Rouleau, N. (2008). Systematic review of the literature on characteristics of late-talking toddlers. *International Journal of Language and Communication Disorders*, *43(4)*, 361-389.

Durkin, K., & Conti-Ramsden, G. (2010). Young people with specific language impairment: A review of social and emotional functioning in adolescence. *Child Language Teaching and Therapy*, *26* (2), 105-121.

Ellis, E. M., & Thal, D. J. (2008). Early language delay and risk for language impairment. *Perspectives on Language Learning and Education, 15* (3), 93-100.

Felsenfeld, S., Broen, P. A., & McGue, M. (1994). A 28-year follow-up of adults with a history of moderation phonological disorder: educational and occupational results. *Journal of Speech and Hearing Research*, *37(6)*, 1341-1353.

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

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

Henrichs, J., Rescorla, L., Schenk, J. J., Schmidt, H. G., Jaddoe, V. W. V., Hofman, A., Raat, H., Verhulst, F. C., & Tiemeier, H. (2011). Examining continuity of early expressive vocabulary development: The Generation R study. *Journal of Speech, Language and Hearing Research*, *54*, 854-869.

Hoeffner, J. H., & McClelland, J. L. (1993). Can a perceptual processing deficit explain the impairment of inflectional morphology in developmental dysphasia? A computational investigation. In E. V. Clark (Ed.), *Proceedings of the 25th Child Language Research Forum* (p. 38-49).

Joanisse, M. F. (2004). Specific language impairments in children: Phonology, semantics and the English past tense. *Current Directions in Psychological Science*, *13(4)*, 156-160.

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.

Justice, L. M., Bowles, R. P., Pence Tumbell, K. L., & Skibbe, L. E. (2009). School readiness among children with varying histories of language difficulties. *Developmental Psychology, 45* (2), 460-476.

Karaminis, T. N. (2011). *Connectionist modelling of morphosyntax in typical and atypical development for English and Modern Greek*. Unpublished PhD Thesis, University of London.

Kovas, Y., Haworth, C.M.A., Dale, P.S., & Plomin, R. (2007). The genetic and environmental origins of learning abilities and disabilities in the early school years. *Monographs of the*

Society for Research in Child Development, *Volume 72*, whole number 3, Serial No. 188, pp. 1-144.

Law, J., Rush, R., Schoon, I., Parsons, S., (2009). Modeling developmental language difficulties from school entry into adulthood: literacy, mental health and employment outcomes. *Journal of Speech, Language, and Hearing Research***,** 52, 1401-1416.

Leonard, L. B. (1987). Is specific language impairment a useful construct? In S. Rosenberg (Ed.), *Advances in applied psycholinguistics* (Vol. 1, pp. 1-39). Hillsdale, NJ: Erlbaum.

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.

Nelson, K. E., Welsh, J. A., Vance Trup, E. M., & Greenberg, M. T. (2011). Language delays of impoverished preschool children in relation to early academic and emotion recognition skills. *First Language*, *31(2)*, 164-194.

Noble, K. G., Norman, M. F., & Farah, M. J. (2006). Neurocognitive correlates of socioeconomic status in kindergarten children. *Developmental Science, 8*(1), 74-87.

Paul, R. (1996). Clinical implications of the natural history of slow expressive language development. *American Journal of Speech-Language Pathology*, *5(2)*, 5-30.

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.

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.

Rannard, A., Lyons, C. & Glenn, S. (2005). Parent concerns and professional responses: the case of specific language impairment. *British Journal of General Practice, 55*(518): 710-714.

Reilly, S., Wake, M., Ukoumunne, O. C., Bavin, E., Prior, M., Cini, E., Conway, L., Eadie, P., & Bretherton, L. (2010). Predicting language outcomes at 4 years of age: Findings from Early Language in Victoria study. *Paediatrics, 126* (6), e1530-e1537.

Rescorla, L., Dahlsgaard, K., & Roberts, J. (2000). Late-talking toddlers: MLU and IPSyn outcomes at 3;0 and 4;0. *Journal of Child Language*, *27*, 643-664.

Rice, M. L. (2009). Language acquisition lessons from children with specific language impairment: Revisiting the discovery of latent structures. In V. C. Gathercole (Ed.), *Routes to language*. Psychology press: Hove, East Sussex, UK.

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

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.

Rowe, M. L., Raudenbush, S. W., & Goldin-Meadow, S. (2012). The pace of vocabulary growth helps predict later vocabulary skill. *Child Development, 83* (2), 508-525.

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

Thomas, M. S. C., Annaz, D., Ansari, D., Serif, G., Jarrold, C., & Karmiloff-Smith, A. (2009). Using developmental trajectories to understand developmental disorders. *Journal of Speech, Language, and Hearing Research*, *52*, 336-358.

Thomas, M. S. C., Baughman, F. D., Karaminis, T., & Addyman, C. (2013). Modelling development disorders. In C. Marshall (Ed.), *Current Issues in Developmental Disorders*. Psychology press: Hove, East Sussex, UK.

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

Thomas, M. S. C., Karaminis, T. N., & Knowland, V. P. (2010). What is typical language development? *Language Learning & Development*, *6*, 162-169.

Thomas, M. S. C., Knowland, V. C. P., & Karmiloff-Smith, A. (2011). Mechanisms of developmental regression in autism and the broader phenotype: A neural network modeling approach. *Psychological Review*, *118(4)*, 637-654.

Thomas, M. S. C., Ronald, A., & Forrester, N. A. (submitted). *Modelling socio-economic status effects on language development*. Manuscript submitted for publication.

Ukoumunne, O. C., Wake, M., Carlin, J., Bavin, E. L., Lum, J., Skeat, J., Williams, J., Conway, L., Cini, E., & Reilly, S. (2011). Profiles of language development in pre-school children: a longitudinal latent class analysis of data from the Early Language in Victoria study. *Child: Care, Health and Development*. doi:10.1111/j.1365-2214.2011.01234.x

Whitehurst, G. J., & Fischel, J. E. (1994). Early developmental language delay: What if anything should the clinician do about it? *Journal of Child Psychology and Psychiatry*, *35*, 613-648.

Young, A. R., Beitchman, J. H., Johnson, C., Douglas, L., Atkinson, L., Escobar, M., & Wilson, B., (2002). Young adult academic outcomes in a longitudinal sample of early identified language impaired and control children. *Journal of Child Psychology and Psychiatry, 43* (5), 635-645.

Zubrick, S. R., Taylor, C. L., Rice, M. L., & Slegers, D. W. (2007). Late language emergence at 24 months: An epidemiological study of prevalence, predictors, and covariates. *Journal of Speech, Language, and Hearing Research*, *50*, 1562-1592.

Table 1. Results from logistic regression analyses, predicting delay group (persisting versus resolving). (a) Empirical data from Dale et al. (2003, Table 6), predicting resolving delay at 4 years of age based on parental report measures of vocabulary, displaced reference, and nonverbal ability at 2 years of age; (b) Empirical data from Dale et al. (2003), adding in gender and a measure of mother's education; (c) Simulation data predicting Time 5 delay group based on Time 1 measures of regular verb, irregular verb, and novel verb performance; (d) Simulation data adding in each individual's family quotient parameter, a proxy for SES; (e) Simulation data, adding in the full set of neurocomputational parameters for each individual.

Predictors	Fit[a] *df*, χ^2	% classified correctly	Sensitivity[1] (%)	Specificity[2] (%)	Positive predictive value[3] (%)	Negative predictive value[4] (%)
Empirical data from Dale et al. (2003)						
(a) Vocabulary, displaced reference, nonverbal	3, 67.0	65.8	44.6	80.5	61.4	67.7
(b) Add in gender and maternal education	5, 91.1	68.5	51.5	80.0	63.8	70.7
Simulation data						
(c) Time 1 behavioural markers	3, 39.6	65.5	79.7	55.6	55.6	80.0
(d) Add SES proxy	4. 39.7	65.2	79.7	55.0	55.3	79.5
(e) Add neurocomputational parameter set	21, 113.9	79.1	72.0	84.0	75.9	81.1

[a] All chi-square values significant at p<.001
[1] Sensitivity = Proportion of PD whose persisting delay was correctly predicted
[2] Specificity = Proportion of RD whose resolving delay was correctly predicted
[3] Positive predictive value = Proportion of predicted PD who had persisting delay
[4] Negative predictive value = Proportion of predicted RD who had resolving delay

Table 2. Effect size (R^2) of relationship between SES and performance, per group, for the Bishop (2005) sample in English past tense production. Typically developing N=166; persisting delay N=94; resolving delay N=104.

Group	Measure			
	Regular verbs		Regular+Irregular verbs	
	R^2	p	R^2	p
Typically developing	.002	.540	.044	.007
Persisting delay	.017	.209	.003	.611
Resolving delay	.028	.087	.043	.035

Regular verbs, interaction of group (PD vs. RD) x SES: $F(1,194)=4.015$, $p=.047$, $n_p^2=.020$.
Regular and irregular verbs, interaction of group X SES: $F(1,194)=2.83$, $p=.094$, $n_p^2=.014$.

Figure captions

Figure 1. Schematic of the population simulations [reproduced with permission from Thomas, Ronald & Forrester, submitted]

Figure 2. Performance distribution on regular verbs at each time point, along with the cut-off for defining developmental delay. μ is the mean and σ is the standard deviation at each time point.

Figure 3. Proportion of simulated population exhibiting language delay at each time point, where delay was defined as falling more than 1 standard deviation below the population mean at that time point.

Figure 4. Sample developmental trajectories for regular verbs, for each group: (a) typical development, (b) persisting delay, (c) resolving delay with low outcome, (d) resolving delay with good outcome, (e) resolving delay with very good outcome. Trajectories are shown for the first 600 epochs. The final time point to determine outcome was 500 epochs of training.

Figure 5. Mean SES values for typically developing (TD), persisting delay (PD), and resolving delay (RD) groups. (a) Empirical data from Bishop (2005), for N=166 TD, 94 PD, 104 RD; (b) Simulation data for N= 713 TD, 118 PD, 169 RD. SES measures have been rescaled to a common range (1=lowest mean, 2=highest mean).

Figures

Figure 1.

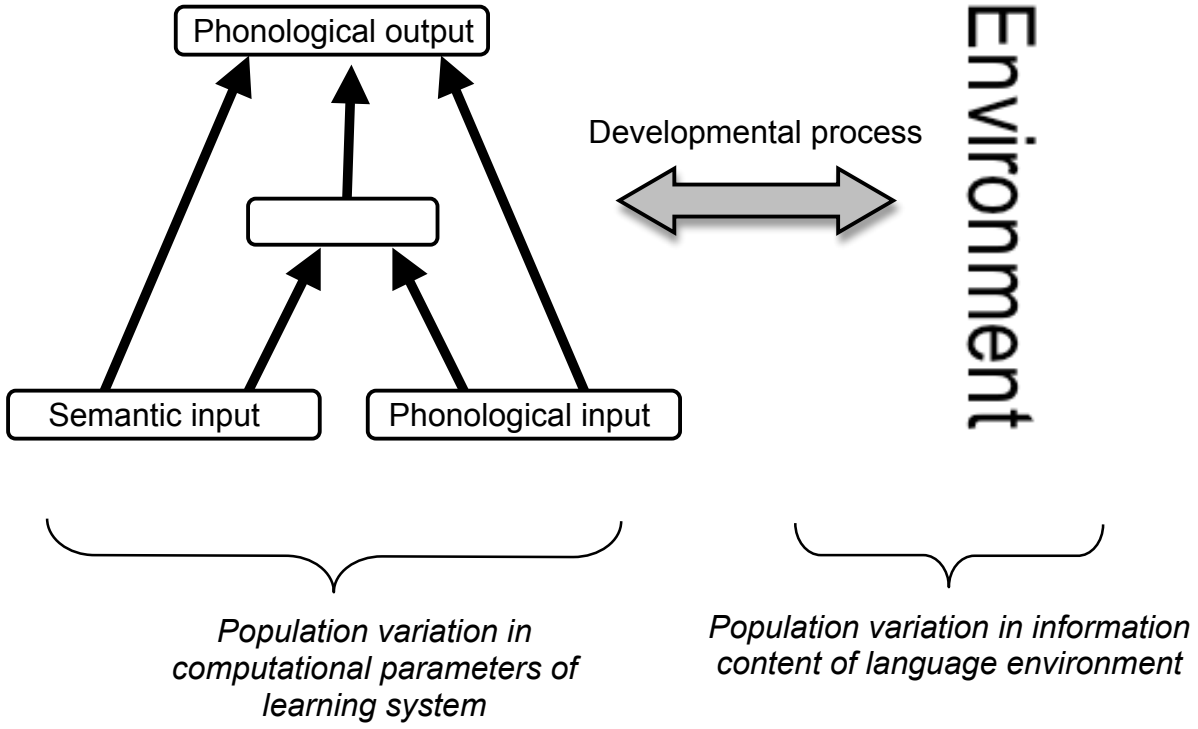


Figure 2.

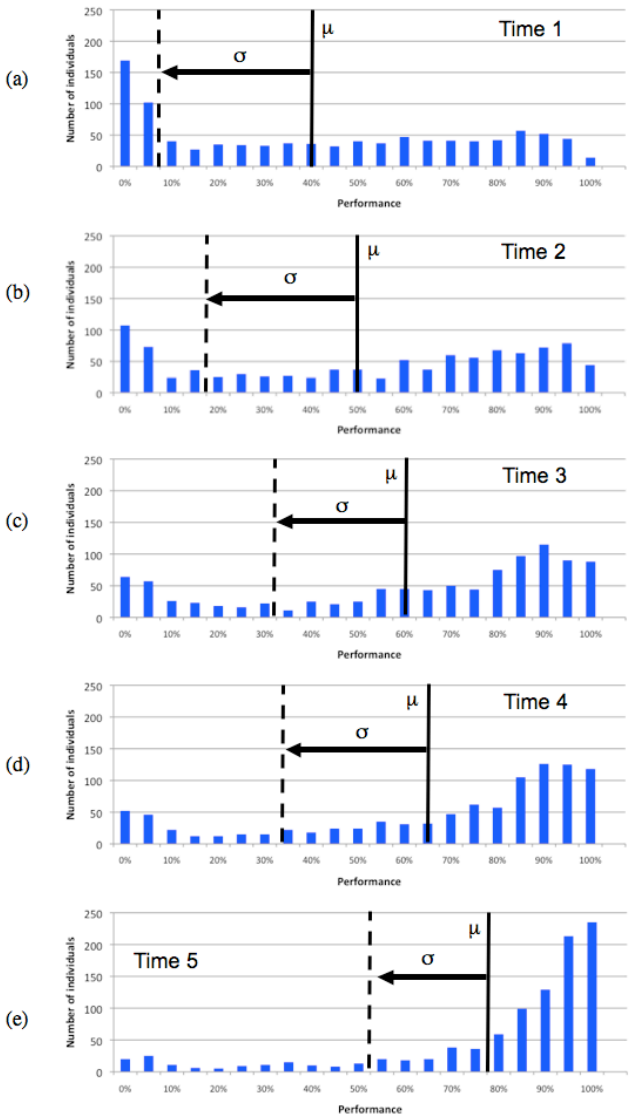


Figure 3.

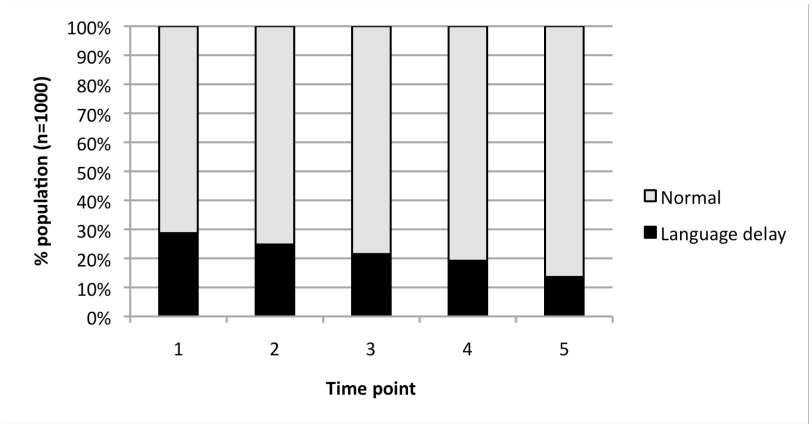


Figure 4.

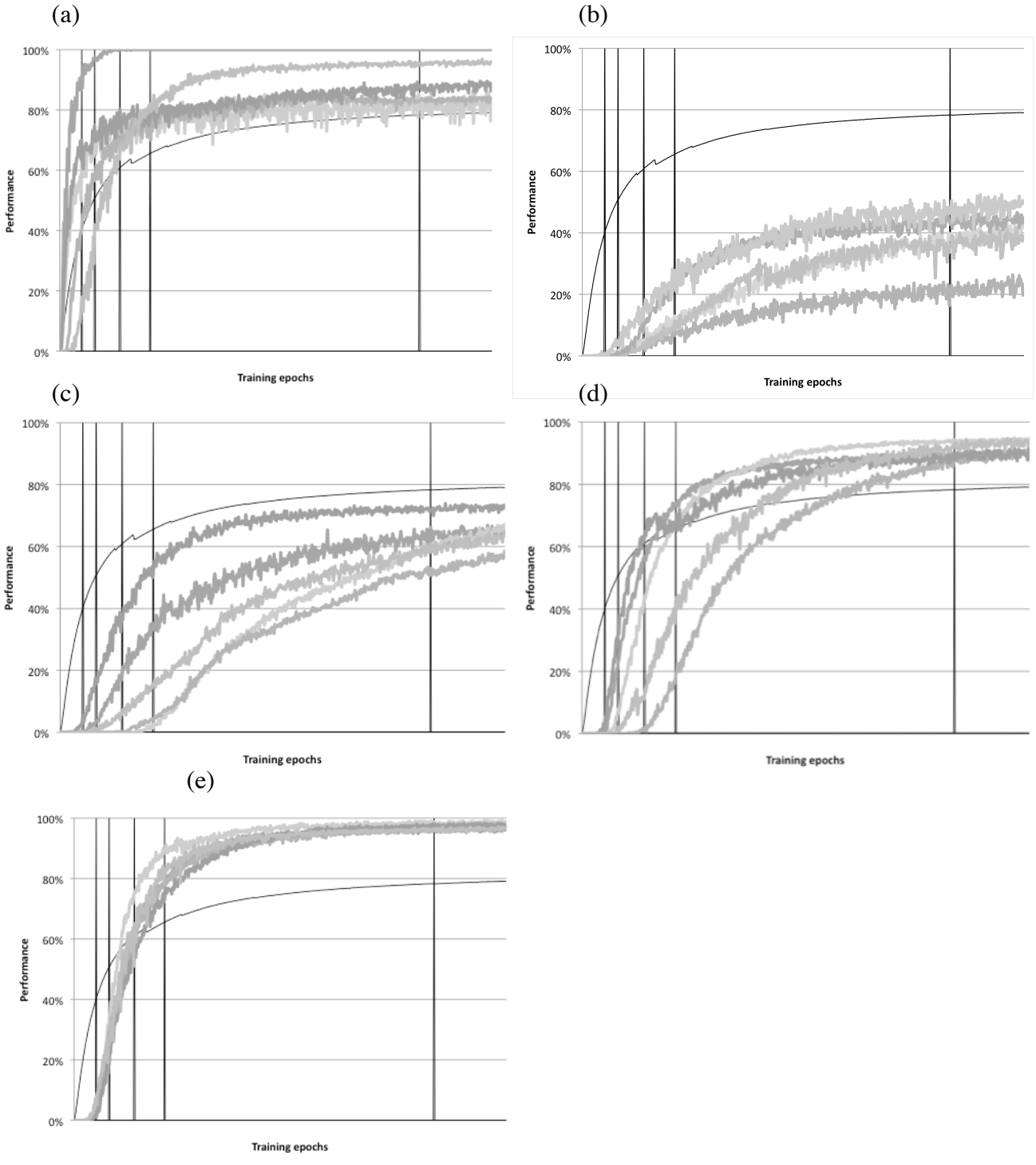
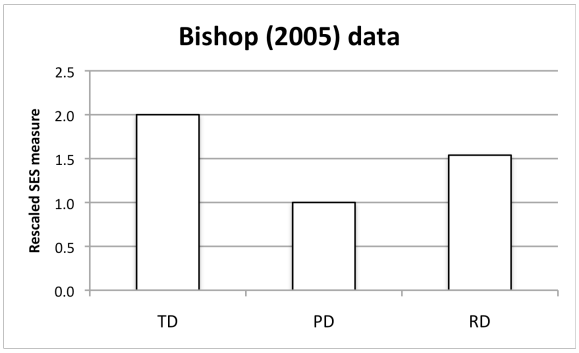
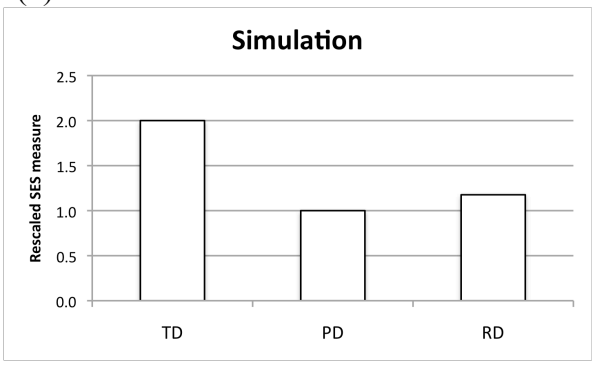


Figure 5.

(a)


Bishop (2005) data
Rescaled SES measure
2.5
2.0
1.5
1.0
0.5
0.0
TD
PD
RD

(b)


Simulation
Rescaled SES measure
2.5
2.0
1.5
1.0
0.5
0.0
TD
PD
RD