Modelling developmental disorders

Michael S. C. Thomas[1], Frank D. Baughman[2], Themis Karaminis[1], and Caspar Addyman[3]

[1] Developmental Neurocognition Laboratory, Birkbeck, University of London, UK

[2] School of Psychology, Curtin University, Australia

[3] Centre for Brain and Cognitive Development, Birkbeck, University of London, UK

To appear in: C. Marshall (Ed.). *Current Issues in Developmental Disorders*. Psychology Press.

Running head: Modelling developmental disorders

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
Web.: www.psyc.bbk.ac.uk/research/DNL/
Email: m.thomas@bbk.ac.uk

Chapter 4. Modelling developmental disorders

Michael S. C. Thomas, Frank D. Baughman, Themis Karaminis, and Caspar Addyman

Introduction: Why build models of developmental disorders?

A six-year-old child is given a vocabulary test where he has to name pictures of objects. He scores poorly, and makes mistakes such as naming a picture of a guitar as a 'piano'. One explanation of the child's difficulty is that he has impoverished semantic (meaning) representations. Another child has problems reading words out loud, especially new words that she has not seen before. For example, she reads the novel word 'slear' as 'sear'. One explanation of her difficulty is that she has poorly specified phonological (speech sound) representations. What do these terms 'impoverished' and 'poorly specified' mean? How did the mental representations of meaning and speech sounds get this way? If a speech and language therapist or a specialist teacher wanted to intervene to remediate these problems, what intervention would be appropriate, and at what age? Should it be a different intervention if representations are impoverished versus poorly specified?

Computational models of developmental disorders represent one technique to improve our understanding of the nature of deficits, their origin, long-term outcomes and possible pathways for remediation. The majority of the models we consider in this chapter are developmental, in that they learn abilities by exposing a developmental system to a structured learning environment. The act of building a model - specifying the nature of the developmental system, its input and output systems, and the information present in the learning environment - forces a theory to be specified in

much greater detail than would normally be the case in a verbal formulation. In the context of disorders, models allow detailed consideration of what could be affecting development - what is it that is different about the developmental system that is preventing strong learning? Is the information different in the learning environment? What exactly is different about meaning representations that are 'impoverished' or phonological representations that are 'poorly specified' and how does this lead to the kinds of error one sees in children with developmental impairments in productive vocabulary and reading, respectively?

The process of constructing a model by its nature involves simplification. The aim is to build a working system that embodies key constraints of the phenomenon under consideration, and simplifies aspects that are taken to be unimportant. For the study of behaviour in children and adults with developmental disorders, empirical data from psychology form the most frequent constraints. For example, we might find that a child with a productive vocabulary deficit also has difficulty in providing definitions of words, indicating the type of information that is missing from the representations of meaning. Models then provide several benefits. As we have seen, the act of building a model forces greater clarity on existing explanations. In addition, a working model allows researchers to test the viability of certain theoretical claims. For example, a model of vocabulary development could test the claim that representations of meaning altered in a certain way indeed lead to naming errors of the type observed in the children. Models can serve to unify a range of empirical effects via a single working implementation. Models can sometimes produce emergent effects that are unexpected consequences of the theoretical assumptions. This is particularly the case where the model system has many components and

complex interactions occur between them. Once the researcher has a working model, he or she can apply it to novel situations, to make predictions about behaviour that can then be tested empirically. With regard to development, the researcher can trace the behaviour of the model across time, for instance predicting the long-term outcome of early-observed deficits. And the researcher can carry out experiments on the model, for example, evaluating different forms of intervention to see which might best alleviate a developmental deficit.

In this chapter, we present examples of the use of computational modelling in the study of developmental disorders. In the next three sections, we illustrate three key ideas. The first section considers the use of models for testing the viability of theoretical proposals – in this case, to establish that certain kinds of deficits in the language system are sufficient to produce behavioural impairments across development, such as those found in Specific Language Impairment (SLI). This section focuses on the additional detail forced by implementation. The second section considers the role of the developmental process itself in producing the impairments, compared to building a static model and simulating deficits by breaking components of the model. This issue is considered in the context of models of reading development and dyslexia. The third section addresses the behaviour of complex cognitive architectures made up of many interacting components rather than individual systems, and explores the developmental consequences of initial limitations to individual components – do deficits subsequently spread throughout the system, or can initially normal components serve to compensate for impaired ones, so alleviating deficits later in development? Once more, this example makes reference to dyslexia in reading. These three examples employ methods drawn from two

approaches to computational modelling, artificial neural networks (sometimes called connectionism) and dynamical systems modelling. Reviews of these methods can be found in Spencer, Thomas, and McClelland (2009), Thomas and McClelland (2008), and Mareschal and Thomas (2007). Following these sections, we consider some of the latest models in the field, including those considering the effects of anomalies in reward-based learning and those modelling deficits at the population level rather than just the individual.

I. Models for evaluating the viability of theoretical proposals: examples from developmental language deficits

A number of studies have employed artificial neural network models to investigate the causes of developmental deficits in aspects of language. The studies shared a common methodology. They constructed a normal model intended to capture the profile of accuracy rates or error patterns in the target domain presented by typically developing children. A disorder model was then constructed by implementing manipulations of the parameter space of the default model in its initial untrained state. These involved altering the computational constraints or the quality of the input and output representations, or both. Examples of such manipulations are (1) the use of fewer units in the hidden layer of the network (e.g., Thomas & Redington, 2004). These are the resources over which the system develops its own internal representations to learn the target domain; (2) the use of an activation function in the processing units which rendered them less sensitive to differences in the input that they received (e.g., Thomas, 2005); (3) the addition of noise to the activation levels of units throughout or in specific parts of the network architecture (e.g., Joanisse & Seidenberg, 1999; Joanisse, 2004); (4) the weakening of the strength with which

certain types of input information were represented in the model (e.g., Hoeffner & McClelland, 1993); and, (5) the probabilistic pruning of weighted connections in the network (e.g., Joanisse & Seidenberg, 1999). When models with these altered constraints were exposed to the target language domain, they exhibited impaired developmental profiles, characterised by the errors shown by young children with language impairments.

The decision of which parameter to manipulate in the normal model was theoretically driven, i.e., linked to aetiological accounts of the language disorder addressed in each study (cf., Thomas & Karmiloff-Smith, 2003). For example, the model of Hoeffner and McClelland (1993) investigating the possible origin of deficits in children with SLI in inflectional morphology. Inflectional morphology is a domain of language that concerns how words change their form to indicate their grammatical status in a sentence (for instance, verbs may be in the present or past tense and nouns may be in the singular or plural). The model evaluated the effect of weakening phonological representations of words, in line with theoretical accounts positing that a low-level perceptual deficit results in the emergence of behavioural impairments at higher levels within the language system (e.g., Tallal & Piercy, 1973a, 1973b). By contrast, Thomas (2005) demonstrated that a similar deficit could be simulated in a learning system that had processing units with reduced sensitivity to variations in the incoming signal. This condition corresponds to theoretical views suggesting that SLI is caused by general processing limitations (e.g., Bishop, 1994; Kail, 1994).

In the following examples drawn from the domain of inflectional morphology, the aim of connectionist studies of language disorders was to illustrate that when applied

to normal models, certain computational constraints were sufficient to alter the acquisition of behaviour to capture the linguistic profile of the language disorders. The models thereby established the viability of the related theoretical account to explain the cognitive profile of the disorder in the target domain. Additionally, the implemented models offered more detailed mechanistic explanations for the application of the general principles of the theoretical proposals in the specific linguistic domains considered. We consider each study with respect to three factors, scope, implementation, and implications. The scope assesses the range of empirical phenomena the model simulated, the implementation addresses which computational conditions were used to simulate the deficit, and the implications consider the wider theoretical consequences for the field of psycholinguistics. The key emphasis throughout this section will be on how the models established the viability of theoretical accounts and added to their detail.

Hoeffner & McClelland (1993): Verbal morphology in SLI

Hoeffner and McClelland (1993) addressed a wide range of deficits of children with SLI in verbal morphology, including those found in the production of base, 3rd person singular, progressive, past tense, and past participle forms of verbs. The normal model was a connectionist attractor network, in which a phonological and a semantic layer were connected bi-directionally. This network was required to learn mappings between distributed phonological and semantic representations of base and inflected verb forms from an artificial language of monosyllabic verb stems design to parallel features of English. An important assumption of the model was that in normal development, certain speech sounds, including word-final stops and fricatives, whether morphemic (e.g., the *-d* in changed) or non-morphemic (e.g., the *-d* in need),

are characterised by lower phonetic saliency for language learners. Therefore, less strong phonological representations were used to represent these phonemes. The reduction of the strength for these representations was implemented in the following manner. Phonemes were represented by activation of the articulatory features that described a given phoneme. For high-salience phonemes, a value of 1 was used for active features. For low-salience phonemes, a value of 0.3 was employed.

The impaired model of Hoeffner and McClelland (1993) evaluated a theoretical proposal that a perceptual deficit may be the root cause of SLI (Tallal & Piercy, 1973a, 1973b). To implement this deficit, the atypical model was given weaker phonological representations for *all phonemes*; for low-salience phonemes, this meant that they were weaker still: thus word-final stops and fricatives were now represented by activation values of 0.1.

With this manipulation present throughout training, the model of Hoeffner and McClelland (1993) successfully simulated a range of morphological deficits presented by children with SLI. Notably, these deficits had been considered strong evidence for an underlying failure of the rule-based system of language by other researchers (e.g., Gopnik, 1990; Gopnik & Crago, 1991). By contrast, the model employed associative learning and did not contain rule-based representations. The model captured the differential degrees of impairment observed in SLI across the different inflections – more pronounced in the 3rd singular, the past tense and past participle, and less severe in the progressive and in verb stems with non-morphemic word-final stops and fricatives (cf., Leonard, 1999). Impaired performance was associated with an increased percentage of inflectional suffix omission errors, also observed in children

with SLI (Rice, Wexler, & Cleave, 1995). Finally, within the semi-regular domain of English past tense, the model simulated a greater degree of impairment in regular than in irregular inflection reported by Gopnik (1990; see also Gopnik & Crago, 1991). It should be noted, however, that the subsequent literature has not confirmed this pattern as characteristic of SLI – rather, lower levels of performance are observed for both regular and irregular verbs, with a residual small advantage for regular verbs (e.g., (van der Lely & Ullman, 2001; see Joanisse, 2004, for discussion). For Hoeffner and McClelland (1993), the model supported the viability of the claim that a general perceptual deficit within the learning system could reproduce a profile of a rule-based impairment, similar to that suggested by language-specific accounts of SLI.

Joanisse (2004): Past tense in SLI

Joanisse (2004) implemented a model for the learning of past tense based on an attractor network architecture trained on mappings between distributed phonological and localist semantics (lexical semantics) representations of English verbs. In a similar fashion to Hoeffner and McClelland (1993), this model also considered an underlying phonological deficit for SLI (based on Leonard, 1999; Tallal, Miller, & Fitch, 1993). However, here the deficit was implemented by the addition of small amounts of random noise to the phonological representations.

The impaired model was weaker in learning both regular and irregular past-tense forms, while performance in generalising the 'past-tense rule' to novel verbs was very low. Joanisse (2004) argued that his model demonstrated that children's representation of phonology is important for all aspects of the acquisition of the past tense. Moreover, he suggested that SLI could not be solely a rule-learning deficit,

since irregular inflection was also affected, both in the model and empirical data (e.g., van der Lely & Ullman, 2001).

Thomas & Karmiloff-Smith (2003): Past tense in Williams Syndrome

Thomas and Karmiloff-Smith (2003) investigated the acquisition of past tense in Williams syndrome (WS), a rare genetic disorder in which language is a relative strength against a background of learning disability. The authors employed a three-layered feedforward connectionist network in which the input layer contained distributed representations of phonology and localist representations of semantics of base forms of verbs, while the output layer was required to produce the phonological form of the past tense of the verb. The training set was an artificial language with monosyllabic stems constructed so as to represent the domain of the English past tense. Thomas and Karmiloff-Smith (2003) considered a wide range of theoretically-driven manipulations of the parameter space of their normal model to contrast five hypotheses for the underlying mechanisms of atypical language development in WS. In particular, they considered atypical conditions that corresponded to (i) a delay, (ii) a hyper-phonological morphological system, (iii) atypically structured phonological representations, (iv) lexical-semantic anomalies, or (v) an integration deficit. Multiple possibilities of atypical conditions were examined for each hypothesis. These ranged from alterations in the number of hidden units or the sensitivity of the activation function to incoming activation, to changes in the architecture or the representational schemes for the different types of information in the model. Thomas and Karmiloff-Smith (2003) showed that different low-level constraints in the computational system could lead to different atypical developmental trajectories, comparable to behavioural data observed in WS (Thomas et al., 2001).

Thomas (2005): Past tense in SLI

In the study of Thomas (2005), the normal architecture of the Thomas and Karmiloff-Smith (2003) model was used to evaluate a proposal of Ullman and Pierpont (2005) with respect to SLI. The Procedural Deficit Hypothesis argues that the language impairments in SLI stem from a deficit in the procedural memory system, in this case in the brain structures involved in the learning of rule-based aspects of language. Importantly, a compensation mechanism from the complementary declarative memory system, which supports lexicon-based inflections, was also proposed. This would explain why aspects of inflectional morphology in SLI sometimes indicate residual knowledge of rule-based inflection, such as over-regularisation errors in irregular verbs (e.g., 'thinked') or generalisation of the past-tense rule to novel forms (e.g., 'wugged') (Leonard, Bortolini, Caselli, McGregor, & Sabbadini, 1992). In Ullman and Pierpont's proposal, the residual knowledge stems from the operation of compensatory declarative mechanisms rather than the procedural system.

Simulations in Thomas (2005) showed that when processing units had activation functions of low sensitivity, so that the units were not good at discriminating small differences in input activations, the model exhibited a qualitative fit to the developmental profile of SLI in past-tense production. However, these impaired units were not part of a system dedicated to regular inflections, as postulated by the Procedural Deficit Hypothesis. Instead, the manipulation was to a low-level constraint in a general processing channel; the constraint happened to be more important for the learning of regular verbs than irregular verbs (that is, it was *domain-relevant*, cf., Karmiloff-Smith, 1998). Importantly, changing the computational properties of this

shared processing channel also changed the balance between how the model used its two information sources: it came to rely more heavily on lexical semantic input, and less heavily on phonological input, to drive its residual behaviour. In this sense, the model reflected the compensatory character of Ullman and Pierpont's proposal. Overall, the study demonstrated the importance of implementation for specifying the nature of compensatory processes in atypical language development.

Karaminis and Thomas (2010): Noun, verb, and adjective morphology in English and Modern Greek SLI

Models of language acquisition need to be general in two ways. They need to account for both typical and atypical language acquisition, and at the same time, they should be able to address language development across languages with different typological characteristics. Karaminis and Thomas (2010) addressed how SLI might emerge in two languages with such different characteristics: English and Modern Greek. Their normal model, called the Multiple Inflection Generator (MIG), combined elements of previous connectionist models of morphology (e.g., Hoeffner & McClelland, 1993; Joanisse & Seidenberg, 1999; Plunkett & Marchman, 1991; Plunkett & Juola, 1999). The architecture is shown in Figure 1. The MIG implemented a more general process of producing inflected forms that would encompass multiple grammatical classes and multiple inflections within a grammatical class. While this adds somewhat to the complexity of the learning in English morphology, for the much richer inflectional paradigms found in Modern Greek, it presents a formidable challenge. The architecture considered was a three-layered feed-forward neural network, which learned to integrate multiple cues presented in the input layer (input phonology, lexical semantics, grammatical class, and target inflection) to output the phonological

form of a word that would be appropriate to the grammatical sentential context in which the word was to be produced.

==========

Insert Figure 1 about here

==========

The same architecture was used to learn mappings from artificial languages incorporating characteristics of the inflectional systems of English and Modern Greek. In the latter case, the training corpus consisted of a notably greater number of mappings, reflecting the complexity and the fusional character of the Modern Greek system of inflectional morphology (Stephany, 1997). In both languages, the default version of the MIG simulated a wide range of empirical phenomena in morphological acquisition in typical development.

The impaired version of the MIG (Karaminis, 2011) combined the use of fewer hidden units in the hidden layer with an implementation of the weaker phonological representations utilised by Hoeffner and McClelland (1993). The same constraint was considered for the English and the Modern Greek version of the MIG. In both cases, the model simulated morphological deficits of children with SLI (e.g., English: van der Lely & Ullman, 2001; Modern Greek: Stavrakaki, Koutsandreas & Clahsen, in press). Importantly, English-speaking and Greek-speaking children with SLI show subtly different patterns of deficit, which the model was able to capture. For example, in past-tense elicitation tasks (e.g., van der Lely & Ullman, 2001), English-speaking children with SLI produced a greater proportion of forms that were not marked for tense than typically developing children. The formation of the perfective past tense of

verbs in Modern Greek requires fusing the stem with morphological features marking the past tense and those marking the perfective aspect. Tense indicates the time when an event happened (past, present, future) while aspect indicates its state at that point in time (imperfective = ongoing, or perfective = a simple whole event). Stavrakaki, Koutsandreas and Clahsen (in press), who considered a perfective past-tense production task, found that the deficits of Greek-speaking children with SLI were more pronounced in the marking of aspect (perfective) than in the marking of tense (past). For the English and Modern Greek versions of the MIG, the same atypical processing constraints produced the increase in unmarked forms in the English case, and the greater deficit in the marking of aspect than tense in the Greek case.

With regards to the aetiology of SLI, the model demonstrated the viability of the idea that weaker representations and processing limitations could provide a unified account of the impairment across different linguistic domains and across languages, with the manifestation of behavioural deficits in each domain and each language depending on an interaction between atypical processing constraints and the structure of the problem domain.

II. The importance of the developmental process as an explanation of developmental deficits: the example of reading and dyslexia

In this section we examine models of developmental dyslexia. We compare the two computational models of reading, the Dual Route Cascade (DRC) model of Coltheart and colleagues (Coltheart et al, 1993; Coltheart et al., 2001) and the triangle model of Seidenberg and McClelland (Plaut et al. 1996; Seidenberg & McClelland, 1989). The models crucially differ with respect to the role of development. The former is an

explicit, cognitive model, with hand-wired structures while the latter is a connectionist learning model. While both were implemented as models of skilled adult reading, here we consider how well they account for the varieties of developmental dyslexia.

Reading is a hard won skill and some children find it much harder than others. Developmental dyslexia is a behaviourally defined disorder associated with poor reading. It is diagnosed when there are severe problems with reading against a background of otherwise normal sensory acuity and cognitive ability, and where the deficit could not be wholly attributable to inadequate instruction, opportunity or motivation to learn. Yet dyslexia is also a developmental disorder where early detection and intervention show remarkable remedial success (e.g., Kujala et al., 2001). Nevertheless, without a properly developmental and computational account it is hard to account for the success (and failure) of remedial programs.

From an evolutionary perspective, reading is a recent cultural invention where evolution is unlikely to have had any direct influence on its acquisition. Therefore it is a cognitive ability that requires a mechanistic account. This is not to say that biology is irrelevant. Learning to read recruits and reshapes pre-existing systems. In particular, it reorganises areas of the brain associated with visual object recognition and is constrained by the computational abilities of this area of the cortex (Dehaene, 2009). What is going wrong to cause dyslexia? Are there general deficits, which manifest themselves only in this particular cognitive domain? Or might dyslexia be attributable to specific deficits in particular brain regions or as a consequence of their connectivity? What does the developmental process itself contribute to poor reading?

There is one straightforward way in which dyslexia is developmental. Children with difficulties reading spend less time reading books. Gabrieli (2009) cites the surprising statistic that "outside of school in 5th grade, a good reader may read as many words in two days as a poor reader does in an entire year" (Gabrieli, 2009, p. 280). With this reduced input and less opportunity to practice, their ability will lag even further behind their peers. Likewise, interventions that increase a child's sensitivity to speech sounds (their 'phonological awareness') may have greatest effect if provided at critical early stages. (For a review of early intervention studies see Torgesen, 2004.) Nevertheless, while environmental factors such as access to reading material and appropriate education can change the outcomes for poor readers, there is a strong genetic component to dyslexia. DeFries and Alarcon (1996) found a 68 percent concordance in identical (monozygotic) twins compared to 38 percent in fraternal (dizygotic) twins. This puts the heritability for dyslexia in the range 54-75 percent (Pennington, 1999). High heritability does not implicate a specific gene for dyslexia. It could be that multiple variants are present in the normal population each of which adds a small risk for reading deficits (see Newbury, this volume). Any genetic account, however, raises the question of how a specific deficit for reading can be heritable when reading is (in evolutionary terms) a recent cultural invention (Mareschal et al., 2007)

Two main types of developmental dyslexia have been described with respect to English (Castles & Coltheart, 1993). *Phonological* dyslexics have most difficulty with regular new words and pseudowords (i.e., word-like nonwords, like HEAN or STARN); *surface* dyslexics have difficulty reading irregular words (e.g., YACHT,

HAVE); some children exhibit a mixed pattern with difficulty on both types of words. The dissociation between reading novel words and irregular words was instrumental in motivating models which posited two processing routes between print and speech, one based on a lexicon of whole words, the other based on links between particular letters / letter clusters and speech sounds.

Coltheart et al. (2001): The Dual Route Cascade model

The Dual Route Cascade model (Coltheart et al, 1993; Coltheart et al., 2001), shown in Figure 2(a), was a hand-coded computational model that fitted a wide range of data from laboratory tasks with skilled adult readers in English. The model was conceived as a comprehensive model of the fully formed adult reading system and conformed to a highly modular view of the reading system.

==========

Insert Figure 2 about here

==========

The Dual Route Cascade (DRC) model has two separate mechanisms for reading words out loud. It has a lexical route, which features an orthographic and a phonological node for each monosyllabic word in English, allowing for their recognition and pronunciation. It also maintains a set of grapheme-to-phoneme conversion rules that apply to all regular pronunciations in English. (A grapheme is a written letter or letter combination that corresponds to a single speech sound or phoneme). When a word is encountered, it activates its representation in the lexical route and simultaneously it starts sequentially activating the conversion rules from left to right, grapheme by grapheme. The activation from both these pathways cascades

through the network until a complete set of phonemes pass a pronunciation threshold. The model does not implement a further semantic route although this is thought to be important for discriminating between words that sound the same (homophones; e.g., HERE and HEAR).

The model provided a good fit to a wide range of adult behavioural data from non-word and pseudoword naming, capturing effects from frequency and regularity and neighbourhood size (that is, how many similar words are there are to a given word in the lexicon.) However, the parameters for the model were all hand-coded in order to fit the empirical data. The lexical route was trained to eliminate mistakes in its performance on tasks with non-word material. "Our odyssey through parameter space thus consisted of running exception word / nonword pairs" (Coltheart et al., 2001, p.219). This is problematic because it means the focus of the fine-tuning of the model was based on non-word materials that would not normally be encountered by a child. Similarly the grapheme-to-phoneme conversion (GPC) rules were selected in advance to be optimal across all possible regular pronunciations in English.

Coltheart et al. (2001) were able to demonstrate that lesioning their model produced many of the characteristics of acquired dyslexia following brain damage. The model did not directly address developmental dyslexia but attributed surface and phonological dyslexia, respectively, to a failure of the lexical route or GPC route to develop properly. Because the DRC architecture was not a learning model, it was unable to show how the developmental deficits could arise from the outcome of an atypical developmental process. Nevertheless, Coltheart et al. (2001) speculated on how the DRC model could inform these issues. They took a highly modular approach

suggesting that "[a]n impairment in learning to read could [correspond to] an impairment in acquiring any one component of this architecture" (ibid, p. 246). A deficit in just the GPC route could lead to a specific difficulty in reading nonwords and could potentially account for phonological dyslexia. A deficit in some part of the lexical route would be manifested as a selective deficit with irregular words, as found in surface dyslexia. Coltheart, Dufty and Bates (2000) (cited in Coltheart et al., 2001) demonstrated that different post-hoc parameterisations of the DRC model could capture the performance of typically developing children and children with dyslexia between the ages of 7 to 15 years.

Despite its success in capturing a range of data, there are three limitations of the DRC model as applied to developmental dyslexia. First, research has demonstrated that the identical damage applied to the initial state of a developmental system can have a quite different effect to that applied to the end state. For instance, noise in processing is more damaging to a developing system than a trained system, while loss of resources is more damaging to a trained system than a developing system. Acquired and developmental deficits cannot be directly analogous (Thomas & Karmiloff-Smith, 2002). Second, the idea of developmental damage to one route of a multi-route system fails to consider the possibility of compensatory changes in initially undamaged routes (see next section). Third, the absence of a developmental process prevents the model from providing a means to investigate ways to intervene to improve a system that has begun to develop atypically. To the extent that the model captures patterns of developmental deficits, it may be doing so for the wrong reasons.

Seidenberg & McClelland (1989); Plaut et al. (1996): The Triangle model

The triangle model of reading was initially described by Seidenberg and McClelland (1989) and its architecture is show in Figure 2(b). Here we focus on the version described in a later article (Plaut et al., 1996). The direct link between the written and spoken forms of words involves only a single route, a connectionist network which learns to associate a grapheme-encoded input with the appropriate phonemically-encoded outputs. The phonetic and orthographic representations also connect to an (unimplemented) semantic representation, and it is this three-way connectivity that gives the model its name. In Plaut et al. (1996), the inputs were 105 units encoding graphemes; these were fully connected to 100 hidden units which fed-forward to 61 fully connected phoneme units. In some simulations there were also recurrent connections in the output layer that served to clean up the selection. Activation passed through the network and training was performed by a standard backpropagation of error algorithm.

As with the DRC model, the triangle model focused on reading monosyllabic words and could produce a similarly high proportion of correct pronunciations. It also captured the influence of word frequency and regularity (and their interaction), as well as exhibiting similar performance to human pronunciation of nonwords with consistent or inconsistent neighbourhoods (Glushko, 1979). However, the triangle model did not *a priori* divide the reading problem into regular and irregular words. In an analysis of hidden units, Plaut et al. (1996) showed that units could not be partitioned according to which type of word they responded. The contribution to the solution was distributed across all hidden nodes in both cases.

Harm and Seidenberg (1999) adapted this model to look at the early development of reading and dyslexia. Children come to reading with an extensive knowledge of the phonology of their native language. So Harm and Seidenberg first had their model acquire phonological representations before learning to map visual word forms onto phonological output. They were then able to investigate how impaired phonological representations affected learning. Mild impairments only affected non-words while severe impairments produced a mixed deficit. Reducing the computational capacity of the network (by removing hidden units before training) produced a pattern similar to surface dyslexia. Harm, McCandliss and Seidenberg (2003) extended this work utilising the model to demonstrate why giving poor readers remedial training in spelling–sound correspondences is more effective than phonological awareness training (McCandliss et al., 2003). In line with the literature, they showed in their model that improvements due to phonological awareness training are only effective in an early sensitive period. Importantly, by virtue of its developmental process, the model was able to shed light on the role of timing on intervention: the quality of phonological representations needs to be improved before links are learned to orthography. Links between orthography and poor phonological representations are hard to unlearn.

Of course, both models have limitations. As in the previous section, models of language processing need to be general not just across typical and atypical development, but also across languages. Both models remain at the cognitive level, with few established links to the neural substrates that may underlie successful and unsuccessful acquisition of reading skills. Lastly, neither model addresses why a heritable disorder should be specific to reading. The suspicion is that whatever

properties are atypical in these reading models, they must be properties that are more general in the language or visual systems, but less easily detected when they are awry outside the realm of reading.

III: Deficit spread versus compensation across development in complex cognitive architectures

In this section we outline a form of computational modelling called *dynamical systems modelling*, which has recently been applied to the study of developmental disorders (see Spencer, Thomas & McClelland, 2009, for a general introduction to the approach). We begin by offering a brief background to the approach and a description of some of its core aims. We then step through a more detailed example in order to demonstrate a key virtue often extolled of dynamical systems: *that complex behaviours of a system can emerge as a consequence of dynamic interactions during development between a number of relatively simple component processes*. We describe recent work in which dynamical systems models addressed issues concerning the specificity of impairments in developmental disorders (Baughman & Thomas, 2008). Following on from the previous section, we use developmental dyslexia as our focus and examine key issues related to the degree of specificity in the reading disorder, and the neurocomputational conditions that may deliver the observed behavioural deficits.

Dynamical systems models are one of a number of mathematical modelling approaches that study change over time. These approaches are derived from *dynamical systems theory* (Thelen & Smith, 1994) and they include, for example, *dynamic field theory*, *growth modelling, catastrophe theory,* and *population*

dynamics. The phenomena targeted by these approaches vary, as does the time course over which change is observed. For example, within the study of infant sensory and motor development, dynamic field theory has been used to explain changes over the millisecond and second range in children's ability to reach and control their movements for objects (Spencer, Perone, & Johnson, 2009; Thelen, Schöner, Scheier, & Smith, 2001). Within the study of the development of language, growth modelling has been used to study how children acquire vocabulary, over a period of days, weeks, months and years (van Geert, 1991). Catastrophe theory has been used to explore changes over hours, days and weeks, in children's ability to reason (van der Maas & Molenaar, 1992). Population dynamics has been used to study the effects of biological and environmental variables on changes to population numbers, over years, decades and centuries (Hofbauer & Sigmund, 1998).

Though these examples of dynamical systems are varied, a common purpose unifies them within the context of human learning and development – this is to understand the *mechanics* that underlie change in complex systems and which allow those systems to alter their behaviour (i.e., to learn and produce new behaviours). To illustrate what we mean by 'mechanics', let us briefly consider the example mentioned earlier, of van Geert's use of growth modelling. Van Geert (1991) was interested in identifying the key influences on language development. He began by simplifying the process of the development of vocabulary to a component process, defined by a single growth curve. The growth of this process was constrained by a number of variables, or parameters that related to: (1) a given, initial level of linguistic proficiency, (2) the rate of linguistic growth, (3) the level of resources in the environment, and (4) the level of environmental feedback. By testing the effects of

small manipulations to these parameters, van Geert showed how in a single model (representing a single learner) the trajectory of development could be dramatically altered. Additionally, van Geert demonstrated how the interactions *between* two models (i.e., two learners) also had markedly different effects on the two models' developmental outcomes. In particular, he found that a combination of supportive and competitive interactions between models best simulated language growth. The mechanics underlying dynamical systems may thus simply be thought of as comprising two aspects, firstly the variables (both biological and environmental) that influence the development of component processes, and secondly the way that these processes are organised and interact (henceforth referred to as the system or network *architecture*).

Dynamical systems approaches offer an advantageous framework for researchers interested in studying cognitive development because at a behavioural level, models exhibit several features that closely resemble change observed in human development. Dynamical systems models can exhibit profiles of change that are non-linear, with the emergence of new abilities often being preceded by periods of marked instability. Around these times of instability, behaviour is influenced both by previously learned, latent knowledge and newer, active representations of knowledge. Additionally, significant changes in the performance of dynamical models often occur suddenly, giving the appearance of stage-like transitions. However, analysis reveals that increases in ability are due neither to the emergence of new, more advanced underlying processes, nor the restructuring of existing processes. Rather, they are the result of continuous change within the interactions of the underlying properties of the system.

Dynamical systems models applied to the study of developmental disorders

Dynamical systems models provide a useful framework for exploring how the process of development contributes to the emergence of abilities and disabilities by virtue of their distinction between the component processes of a system, the variables that influence their growth, and the nature of the interactions between those processes. To explore the possible source of developmental disorders such as developmental dyslexia, we must establish what these mechanics might be. To guide this search, we turn to the literature pertaining to cognitive theories of developmental disorders. Here, we find that one of the common assumptions is that by adulthood, cognition is organised largely in a *modular* manner. That is, the range of cognitive abilities that humans come to develop is the product of a number of functionally specialised cognitive components. While we described the DRC model in these terms, it was also true of the triangle model of reading, in its distinction between orthographic, semantic, and phonological representations of knowledge.

Evidence for claims of functional specialization are often derived from studies of adult patients who, following brain damage or disease, have been shown to exhibit dissociations in their cognitive abilities. That such cases appear to show that the cognitive system may become 'fractionated' has been used as the basis for developing models of the normal adult cognitive architecture (Shallice, 1988). The use of modular architectures to explain the causes of developmental disorders is an issue of contention, inasmuch as they risk de-emphasising or even ignoring the developmental process, as we saw in the previous section (see Temple, 1997; Thomas & Karmiloff-Smith, 2002; Thomas, 2006). Within modular accounts of disorders, one finds that the

root cause of a disorder is often explained in terms of either a 'delay' or, a 'deficit' to the functioning of a single cognitive module, or process (illustrated in the application of the DRC model to developmental dyslexia). Yet, such modular explanations rest on two assumptions: (1) that the cognitive system of the child is also modular, and (2) that during the process of development, module-specific deficits can persist without compensation by or spread to other causally linked cognitive processes. Both these assumptions have been challenged (e.g., Filippi & Karmiloff-Smith, this volume; Karmiloff-Smith, 1998). Whilst uneven profiles of cognitive abilities are often found during childhood, current debates concern precisely how deficits emerge and the true extent of specificity of a deficit in a developmental disorder.

The answers to these questions very much depend on the nature of the cognitive architecture present in children. For example, distributed theories (of the sort inspired by McClelland & Rumelhart, 1988) lead to doubts that any deficit, however domain specific to begin with, could remain so across development. In such theories, cognitive processes are graded and interactive, relying on the contribution of many different components. Evidence from the neurosciences supports the view that the brain is highly interactive and capable of compensation following some forms of early damage (Thomas & Karmiloff-Smith, 2002; Thomas, 2003). Between the extremes of fully modular and fully distributed theories lie various positions that propose more limited degrees of cognitive differentiation. For example, hemispheric specialisation may be important even if functions are interactive within each hemisphere, as evidenced by the emergence of laterality effects after unilateral brain damage in the domains of language (Bates & Roe, 2001) and spatial cognition (Stiles, 2001). Some accounts focus on the importance of a central executive (see e.g., Baddeley, 1996),

while others emphasise hierarchical organisation in cognition (e.g., Anderson & Lebiere, 1998).

Computational modelling once more provides the opportunity to test the viability of different theoretical accounts. By taking a modelling approach, we can explicitly assess the consequences of assuming a given architecture for the development of an impairment. In the following illustrative example, we simplified the simulation of developmental processes at the level of individual components in order to focus on the implications of their interactivity in five large-scale architectures. The simplification involved assuming that the development of a cognitive process can be captured by a growth curve defined by a small number of parameters, including its onset, rate of growth, and final asymptotic value. Variations in these parameters were then used to depict heterogeneous underlying mechanisms and domains. By postulating different global architectures (fully distributed, hemispheric, central processor, hierarchical, and modular, shown in Figure 3), we then examined the consequences on development of damage that initially occurred to a single process – these are the conditions that modular theories propose to be responsible for apparently domain-specific developmental deficits like developmental dyslexia.

==========

Insert Figure 3 about here

==========

A dynamical systems model

The dynamical systems model we used was based upon a framework developed by van der Maas and colleagues. Van der Maas et al. (2006) proposed a dynamical model

of the development of intelligence. It simulated cognitive development for a number of different components via non-linear growth curves in a fully connected system (depicted in Figure 4).

==========

Insert Figure 4 about here

==========

A fundamental feature of the model is that all of the processes within the system co-operate throughout development. Unique parameters help guide the development of individual processes, but development is also influenced dynamically by the performance of all other processes. These interactions result in mutually beneficial and positive influences over development. Hence, the model was referred to as the 'mutualism' model. The following coupled differential equation specifies the dynamics of the mutualism model.

$$\frac{dx_i}{dt} = a_i x_i (1 - x_i / K_i) + a_i \sum_{\substack{j=1 \\ j \neq i}}^{w} M_{ij} x_j x_i / K_i$$

The mutualism equation was derived from population dynamics and the *Lotka-Volterra* equation. The equation states that at each point in time (t) the change in the performance level x of a given process i (dx_i) is a product of the sum of the interaction weights of each process j with which it is functionally connected ($M_{ij}x_jx_i$), multiplied by the rate of growth of process i (a_i) times the current level of performance of process x_i, divided by the asymptote level for that process (K_i). For each process, changes in x_i at each time step are constrained by the performance (and thus the individual properties) of all other processes to which it is connected. Because

the parameters that influence the model's behaviour are relatively few (i.e., a, K and x) and because the functional architecture can be explicitly specified via a matrix of functional connectivity (***M***), we considered the model to be a useful framework for investigating issues surrounding specific developmental impairments under various architectures.

Due to the fact that the model necessarily sits at a fairly high level of abstraction, one consequence is that it becomes more difficult to elucidate what each of the model parameters relates to, in terms of *specific* biological, or environmental factors. At this level of simplification, the model parameters reflect largely a blend of influences from both. For instance, the growth (a) of a given cognitive process may likely be influenced by both biological and environmental factors. On the other hand, the initial level (x) of a process (which may be initially constrained by growth) may be primarily dependent on environmental input. The capacity of a process (K) may be influenced both by environmental and biological factors, and the degree of interconnectivity (M) may initially be largely dependent on biological factors, but susceptible to effects in the environment.

Once a cognitive architecture has been specified, one must decide where it is appropriate to apply an initial deficit. Within each architecture, processes differ in their interconnectivity, and thus are likely to differ in the amount of influence they exert on the development of the system as a whole. To illustrate, take the following comparison between the fully distributed model and the central processor model, shown in Figure 3(a) and 3(c), respectively. In the fully distributed model one can see that all processes share the same degree of connectivity. Therefore, the effect of

damage to one process should be equivalent to the effect of damage to any other. In contrast, in the central processor model, the degree of connectivity differs between processes. Whereas the central process has the greatest number of connections (it is connected to all processes), the connectivity of any other process is more limited (each other process is connected only to the processes within the same cluster). Within each architecture, therefore, the consequences of an initial deficit should vary, depending on whether it is applied to a *peripheral* process (one with relatively fewer connections) versus a *key* process (one with a relatively greater number of connections). For each of the architectures given in Figure 3, we applied an initial focal deficit to a single component, either to its onset, growth rate, final asymptote, or combinations of these three. We then traced the effects of the deficit separately on both peripheral and key processes, over the full architecture as development proceeded. Deficits were applied to the start state of a population of simulated individuals, who had minor variations in the initial values of their onsets, growth rates, and asymptotes, but all of whom shared a common architecture.

An illustrative example of deficit spread versus compensation

With respect to developmental dyslexia, the most pertinent result concerns conditions where simulations produced lasting deficits for a single process (corresponding, say, to the GPC component in the DRC model). We found that a large impairment (e.g., a 75% reduction of the normal level) to just one parameter (the *K* parameter, asymptote level) was sufficient to produce this outcome across the range of models. We assessed two additional properties. *Compensation* was assessed based on whether the performance of the initially damaged process was reliably different to if it had

developed in isolation, unconnected to any other process. *Deficit spread* was assessed based on whether the performances of the initially undamaged processes were reliably different to those same processes in the normal model. Figure 5 depicts the developmental profiles for each of the individual processes for each architecture. Blue lines show the trajectories of each process in the normally developing models and red lines show the trajectories of each process in developmentally disordered models. Horizontal dashed lines depict the level of performance that is predicted for the level of damage, were the process to develop in isolation, and against which the action of compensation was gauged.

==========

Insert Figure 5 about here

==========

Unsurprisingly, in the modular architecture, early selective damage to a single process resulted in a dramatic drop in the performance level for that process. Also of no surprise was the fact that in the modular model the initially unaffected processes developed normally. Due to the lack of interconnectivity, the modular network exhibited no spread of deficit and equally, offered no compensation to the damaged process. In this case, an early deficit would result in a truly specific impairment. The pattern was different in the other four architectures. As the process of development unfolded, the effects of early damage to a single process were not isolated. Figure 5 illustrates the extent of deficit spread for each of the architectures (shown via the lower-than-normal developmental trajectories for initially unaffected processes) and compensation (where performance for the damaged process was above the level predicted).

Notably, while the performance of the damaged component in the latter architectures was significantly lower than the normal model, there was no reliable difference in performance between the initially undamaged processes and their normally developing counterparts, *against the background of variability in the population as a whole*. That is, the spread of the deficit over development was masked by the fact that the performance was *within the normal range* for the population, even though it was below the level it would have been if the starting conditions in each system were normal (see Figure 6). Superficially, the behavioural profile suggests that specific impairments, of the sort reported in cases like developmental dyslexia, are possible under a variety of neurocomputational conditions, including those that did not specify a modular functional architecture; in fact, in non-modular, interactive cognitive architectures, the effects of the damage were never truly specific but instead were widespread and subtle, with the system's dynamics determining the degree of deficit spread and the amount of compensation following early forms of damage. The implication of these findings is that if the functional architecture of cognition in the child is not modular, then a range of other cognitive domains outside the primary deficit may show subtle deficits, even under conditions where that the initial deficit began as more restricted (see also, Williams and Lind, this volume).

==========

Insert Figure 6 about here

==========

Computational simulations of this kind have the potential to reconcile views of the apparently specific nature of behavioural impairments in disorders such as

developmental dyslexia with those that posit the highly distributed nature of cognition. The task of narrowing down the range of candidate architectures requires the combined efforts of empirical (both behavioural and neurosciences) and computational approaches. For example, if it were the case that current assessment techniques do not detect the subtle developmental effects of a deficit in one domain on other cognitive domains, then a good starting point would be to utilise more sensitive behavioural measures of the apparently normally functioning domains. Indeed, studies aimed at refining the methods for assessing children's cognitive abilities are underway (see e.g., Bornstein, 2011; Rezazadeh, Wilding, & Cornish, 2011). By more accurately profiling the abilities of children, it may be possible to eliminate some architectures from enquiry. Converging evidence from the neurosciences will be invaluable in this matter. For example, studies targeting whole-brain patterns of activity are beginning to identify the causal, functional relations between cognitive domains (see e.g., Bressler & Menon, 2010; Hu et al., 2011; Jolles, van Buchem, Crone, & Rombouts, 2011; Menon, 2010).

Indeed, some of this work has suggested that the functional architecture of cognition may be organised according to the properties of *small-world networks*. Small-world networks offer another example of a dynamical system in which a network consists of a number of component processes, between which are varying amounts of connectivity. While each process is causally related to each other process, their influence can be exerted via shorter or longer pathways. Interactions between any two processes can take place either via direct connections or via pathways employing variable numbers of intermediate nodes. In the case of a *regular* small-world network shown in Figure 7(a), where a regular relationship exists for all processes,

connections are limited to processes that are near each other. Longer-range interactions will require many intermediate nodes. In other cases, shown in Figure 7(b) and 7(c), small-world networks exhibit additional *random* connections. The effect of these random connections is a shortening of paths, and thus more direct influence between processes from diverse areas of the system.

==========

Insert Figure 7 about here

==========

The dynamics of small-world networks has been studied in a variety of contexts (e.g., social networks and in the spread of disease in populations). However, their relevance here comes from the use of brain-imaging techniques, which have demonstrated that distributed patterns of activity resembling small-world networks underlie a range of cognitive activities (Boersma et al., 2011; Ferrarini et al., 2009; Fransson, Aden, Blennow, & Lagercrantz, 2011; van den Heuvel & Pol, 2010). In these cases, the view that is emerging is that cognition is comprised of regions of highly connected processes or 'cortical hubs' (Achard, Salvador, Whitcher, Suckling, & Bullmore, 2006) and regions where connectivity between processes is more diffuse. The functional differences in the properties of these networks have been examined in disorders such as Alzheimer's disease (Stam et al., 2009), schizophrenia (Liu et al., 2008) and attention deficit hyperactivity disorder (ADHD) (Wang et al., 2009). Our current simulation work is extending dynamical systems modelling of developmental deficits to small-world scenarios.

Despite these advances, the challenge of modelling at the level of cognitive architectures is to understand for a given disorder exactly how widespread the neurocomputational differences are in the atypically developing brain, and how the effects of these differences unfold through the intricacies of the developmental process (see also, Filippi & Karmiloff-Smith, this volume).

Forthcoming work in the modelling of developmental disorders

The previous models have captured developmental deficits in terms of neurocomputational limitations to the representations or processing within associative learning systems. However, there are other types of learning which provide alternative candidate pathways for developmental deficits. Reinforcement learning involves learning cognitive operations, actions, or sequences of actions that maximise rewards. It is possible that individuals with disorders find different aspects of their environments rewarding compared to typically developing children. This in turn may change the way children with disorders interact with and attend to their environments and indeed, the subjective nature of the environments to which they are exposed. One example comes from the domain of eye gaze behaviour. Typically developing infants learn to use the direction of their caregiver's gaze to predict where to find interesting objects in the immediate environment. Infants and children with autism tend to avoid looking at caregivers' eyes, leading to disruptions in the development of dyadic (two-person) interactions. By contrast, in WS, infants and children seem captivated by the faces of caregivers, yet they show deficits in triadic interaction, where there is failure to establish shared attention between child and adult on an object. Triesch et al. (2006) constructed a computational model of the development of infant eye-gaze following based on reinforcement learning. In this model, the simulated infant learned

that if she looked at her caregiver's direction of eye gaze, this might serve as a predictive cue of where in the environment interesting objects might be found, which the infant could then fixate. Through a sequence of exploratory behaviour in the simulated environment, the infant came to maximise the reward she gained from fixating her caregiver and from using the direction of her caregiver's gaze to look at rewarding objects around her. Triesch et al. then simulated two conditions of atypical development, building in constraints from autism, where faces are hypothesised to be intrinsically less rewarding, and WS, where faces are hypothesised to be more rewarding than normal. In both cases, the simulated infants showed developmental deficits in gaze following behaviour, where faces were either avoided, so attenuating caregiver eye gaze direction as a predictive cue, or fixated for longer than normal, so failing to move on to fixate objects in the environment. In both cases, the atypical reward conditions led to emergent deficits in the development of gaze following. Notably, these two atypical models for autism and WS, distinguished only by the reward value attached to faces, looked very similar in the early stages of development. However, the small difference in the start state led to a radical divergence between the systems across development, until they exhibited very different behaviour. (See Richardson & Thomas, 2006; Williams & Dayan, 2004, for related work modelling reward learning in ADHD).

Kriete and Noelle (submitted) recently postulated that problems in reward-based learning might contribute to deficits in executive functioning observed in adolescents and adults with autism. The authors used as their model of normal development the Cross-Task Generalisation (XT) model of pre-frontal cortex task control (Rougier et al., 2005). The objective was to capture two pieces of empirical evidence regarding

executive dysfunction in older individuals with autism: perseverative errors on the Wisconsin Card Sort Task (WCST) and normal performance on the Stroop task. In the WCST, subjects must show cognitive flexibility in altering the dimensions over which they sort cards (e.g., by colour, by shape). In the Stroop task, subjects must respond to a single dimension of a stimulus and ignore another potentially more salient dimension. The original XT model combined both reinforcement learning and associative learning to simulate performance on the WCST and Stroop tasks, in adults and in individuals suffering acquired frontal brain damage.

Kriete and Noelle pursued the hypothesis that dopamine may be reduced in autism, thereby dysregulating the interaction of the mesocortical dopamine system with the prefrontal cortex (PFC). The dopaminergic neurotransmitter system implements reward-based learning in PFC. The postulated deficit in the reward signal in the XT model affected a gating mechanism that destabilised short-term PFC representations supporting task performance. Destabilisation was key for flexibility in behaviour, since it opened the PFC to change its task configuration. A change of task configuration is crucial in WCST but is not required in the Stroop task. As a result, the reduction in destabilisation in the autism condition led to perseverative behaviour in a simulation of WCST (i.e., continuing to sort the cards by a dimension that was no longer relevant) but did not alter performance on the Stroop task. In addition, Kriete and Noelle found that the WCST impairment was late emerging in the development of the autistic model, because the PFC component only adopted the role of supporting cognitive flexibility once associative mechanisms (modelling posterior cortex) had acquired relevant abilities. The model therefore provides a novel causal explanation of why behavioural deficits in executive function may be late emerging in autism,

even though their primary cause (a reduction in dopamine) is in place throughout development.

As the Kriete and Noelle paper illustrates, one advantage of using computational models based on principles of neurocomputation is the opportunity to make links to evidence from neuroscience. Ultimately, the causal explanation of a disorder will span many levels of description. The known genetic basis of some disorders (such as the genetic mutations in disorders like Down syndrome and WS) and the high heritability observed in behaviourally-defined disorders (such as autism and dyslexia) implies that the lowest level of description will be genetic. Yet there are puzzles that arise from genetic accounts of disorders. One of these is that genetic mutations and gene variants are often only probabilistically associated with behavioural outcomes observed in disorders. There must be other risk and protective factors that modulate the relationship between a given genetic cause and the behavioural phenotype observed in a disorder. Moreover, some common gene variants have been associated with more than one disorder (e.g., developmental language impairment and autism; Vernes et al., 2008).

A new simulation approach based on population modelling has begun to investigate the probabilistic nature of the causes of developmental deficits. Population modelling involves simulating large numbers of individuals undergoing both typical and atypical development. Apart from the hypothesised cause of the disorder, the framework includes the possibility of population-wide variation in the neurocomputational properties of all children, as well as variations in the quality of the environment to which all children are exposed. One model utilising this approach, by Thomas,

Knowland and Karmiloff-Smith (2011), evaluated the hypothesis that autism may be caused by disruptions in connectivity occurring during synaptic pruning. During development, the brain initially produces exuberant connectivity, which is subsequently pruned back in childhood. This gives the brain greater plasticity in early development, to adapt to the environment in which it finds itself, while saving on metabolic resources later in development. However, if the pruning process is too aggressive, rather than just removing spare computational resources, it can comprise the neurocomputational properties of the system or even lead to regression in behaviour. Notably, Thomas et al. found that the cause of the disorder in their networks (over-aggressive synaptic pruning) interacted with other dimensions that varied in the general population, such as the amount of computational resources, the rate of learning, and the richness of the learning environment to which the individual was exposed. These risk and protective factors led to a probabilistic relationship between the (in the model, known) cause of the disorder and its manifestation in behavioural deficits. Moreover, the authors demonstrated how a direct cause of one disorder (e.g., slow development) could be a risk factor for another (e.g., slow development makes the effects of aggressive synaptic pruning worse). This would explain why there should be shared causal factors (such as gene variants) between different disorders: the shared factor indexes the cause of one disorder and the elevated risk (but not direct cause) or another. Bishop (2006) recently advocated that researchers move to an explanatory framework of developmental disorders based on risk and protective factors, rather than necessary and sufficient conditions. Population modelling is a new approach that is consonant with this shift to viewing causal factors as probabilistic against a background of variability.

Lastly, population level models also permit a consideration of the effects of variations in the quality of the environment. Thomas, Ronald and Forrester (submitted) recently modelled the effects of socio-economic status (SES) on language development at the population level, evaluating the idea that one way that SES might operate on cognitive development is via a manipulation of the amount of information available to the child. This model generated the novel prediction that SES should be statistically associated with good developmental outcomes in children but not with bad developmental outcomes. Empirical data from the acquisition of inflectional morphology (Bishop, 2005) offered direct support for this novel prediction. Crucially, because the operation of the model was understood, it was possible to show that this asymmetric statistical relationship was misleading. Poor environment did indeed cause poor developmental outcomes in the model. However, because a range of other neurocomputational factors could also compromise developmental outcomes, the unique statistical predictive power of the environment was lost. By contrast, for a good developmental outcome, all factors must be good (i.e., a good learning system and a good environment). Presence or absence of good SES then becomes more uniquely predictive. In this way, the implemented mechanistic model offered a deeper understanding of causal relations than that available through simply identifying correlations between behaviour and factors in the environment.

Using models to investigate intervention

Implemented computational models of developmental deficits provide the foundation to explore possible interventions, and indeed allow for a much wider range of interventions to be considered than in human studies, where there are both practical and ethical limitations. Nevertheless, work on simulated interventions has been

relatively limited so far, the greater focus having been on building accurate models of the disorders themselves. The model of dyslexia in reading offered one example of a simulated intervention. Harm, McCandliss and Seidenberg (2003) demonstrated how improving the internal structure of poor phonological representations via training on component sounds of whole words was successful in improving the subsequent acquisition of mappings between print and sound. The model in addition demonstrated why such an intervention was more effective before the start of literacy training – once the system started to learn mappings between orthography and poorly structured phonology, these bad mappings were hard to unlearn.

A number of questions are brought to the fore in considering simulated interventions in a developmentally disordered system. First, where a cognitive system has an atypical processing property, can this be normalised by the intervention? Second, where there has been a history of development with the atypical property, can the consequences of this development be undone? In part, this relates to how the plasticity of the target cognitive system changes with age. Third, should the intervention target atypical processing properties directly, or should it operate through exposing the child to a differently structured learning environment? Fourth, if it is not possible to normalise the processing properties of the system by an intervention (so that the child cannot feasibly hope to master all aspects of a target domain), which subset of behaviours should be optimised?

These questions can be illustrated by some examples. In the model of dyslexia discussed above, it was possible to normalise the system by intervention – training to improve phonological awareness altered phonological representations sufficiently for

normal mappings to be learned between orthography and phonology. However, the consequences of a history of atypical development were harder to undo – the intervention was less successful if literacy acquisition had already commenced. In the Thomas, Knowland and Karmiloff-Smith (2011) model of autism, which simulated the disorder via over-aggressive pruning of connections in an artificial neural network, normalisation would not be possible – the connectivity was permanently lost. In this model, interventions could only aim to generate the best behavioural outcome that the altered connectivity pattern would allow. Direct interventions that target atypical processing properties might be possible in the future. Researchers working with animal models of Down syndrome have reported that a drug intervention that reduced (excess) neural inhibition in a mouse model led to improved learning on a novel object recognition task (Fernandez et al., 2007). Nevertheless, one might expect most interventions to operate via engaging the child with differently structured learning environments – perhaps those that exaggerate key dimensions of the task to be learned, or focus on prototypical behaviours. One key challenge to be addressed is how such behavioural interventions can successfully generalise beyond the items used in the intervention itself. Given that generalisation is a much-studied dimension of computational learning systems, it is an ideal challenge to be addressed by computational models of developmental disorders. In short, the modelling approach holds great promise to study and predict effective interventions and work is indeed underway in a number of labs, but the approach has yet to deliver substantial results.

Conclusion

In this chapter, we outlined the advantages offered by computational modelling in advancing our understanding of the causes of developmental deficits. By implementation, models force greater specification on theoretical proposals, and test their viability. They generate novel testable predictions, and allow the model system to be evaluated in new conditions, for example to test possible interventions. We considered examples from language and reading development, and from disorders including SLI, dyslexia, autism, and WS. We considered individual cognitive systems, large-scale cognitive architectures, and interactions between reward-based learning and associative learning. We outlined the new approach of population modelling to investigate risk and protective factors modulating the relationship between disorder cause and behavioural outcome. Throughout, central to our argument has been that explanations of developmental deficits need to focus on the nature of the developmental processes itself (Karmiloff-Smith, 1998), and that computational modelling offers the means to do so.

Of course, a computational model can never demonstrate that a proffered explanation of a developmental deficit is the correct one. Models can only demonstrate that a given account is a viable one. And, as we pointed out earlier, by their nature, models will always contain simplifications, which can under some circumstances, compromise their applicability. Despite their merits, researchers should be cautious in evaluating models. For example, a number of questions might be asked of any computational model of a cognitive process: (1) How robust are the target data that are being simulated? (2) Does the model leave out any key psychological, neural, or environmental constraints? (3) Does the model include anything irrelevant or incorrect in its implementation that is instrumental in producing the target behaviour?

And (4), does the model unify a range of empirical effects and/or produce testable predictions?

Although they must be interpreted with caution, we believe, nevertheless, that computational models have great potential to complement behavioural and neuroscience methods in understanding the causes of disorders, and ultimately, in identifying the best interventions to remediate the negative consequences of these disorders on the developing child.

Acknowledgements

This research was supported by ESRC grants RES-062-23-2721 and RES-062-23-0819 and a Leverhulme Study Abroad Fellowship to Michael Thomas held at the University of Chicago. The studies of Themis Karaminis were funded by the Greek State Scholarship Foundation (IKY) and the Greek Ministry of Education. We would like to thank Marc Joanisse for his helpful comments on an earlier draft of this chapter.

References

Achard, S., Salvador, R., Whitcher, B., Suckling, J., & Bullmore, E. (2006). A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. *The Journal of Neuroscience*, *26*, 63.

Baddeley, A. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *49*, 5-28.

Bates, E., & Roe, K. (2001). Language development in children with unilateral brain injury. In C. Nelson, & M. Luciana (Eds.), *Handbook of developmental cognitive neuroscience* (pp. 281-307). Cambridge, MA: MIT Press.

Baughman, F. D., & Thomas, M. S. C. (2008). Specific impairments in cognitive development: A dynamical systems approach. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 1819-1824). Austin, TX: Cognitive Science Society.

Bishop, D. V. M. (1994). Grammatical errors in specific language impairment: Competence or performance limitations? *Applied Psycholinguistics*, *15*, 507-550.

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,* 479-490.

Bishop, D. V. M. (2006). Developmental cognitive genetics: how psychology can inform genetics and vice versa. *Quarterly Journal of Experimental Psychology*, *59*, 1153-1168.

Boersma, M., Smit, D. J. A., de Bie, H. M. A., Van Baal, G. C. M., Boomsma, D. I., de Geus, E. J. C., . . . Stam, C. J. (2011). Network analysis of resting state EEG in the developing young brain: Structure comes with maturation. *Human Brain Mapping*, *32*(3), 413-425. doi:10.1002/hbm.21030.

Bornstein, R. F. (2011). Toward a process-focused model of test score validity: Improving psychological assessment in science and practice. *Psychological Assessment*, *23*, 532.

Bressler, S. L., & Menon, V. (2010). Large-scale brain networks in cognition: Emerging methods and principles. *Trends in Cognitive Sciences*, *14*, 277-290.

Castles, A., & Coltheart, M. (1993). Varieties of developmental dyslexia. *Cognition*, *47*, 149-180.

Coltheart, M., Curtis, B, Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review*, *100*, 589-608.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A Dual Route Cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*, 204-256.

DeFries, J.C., & Alarcón, M. (1996). Genetics of specific reading disability. *Mental Retardation and Developmental Disabilities Research Reviews*, *2*, 39-47.

Dehaene, S. (2009). *Reading in the brain: The science and evolution of a human invention*. New York: Penguin Viking.

Fernandez, F., Morishita, W., Zuniga, E., Nguyen, J., Blank, M., Malenka, R. C., & Garner, C. C. (2007). Pharmacotherapy for cognitive impairment in a mouse model of Down syndrome. *Nature Neuroscience*, 10(4), 411-413.

Ferrarini, L., Veer, I. M., Baerends, E., van Tol, M. J., Renken, R. J., van der Wee, N. J. A., . . . Milles, J. (2009). Hierarchical functional modularity in the resting-state human brain. *Human Brain Mapping*, *30*(7), 2220-2231. doi:10.1002/hbm.20663.

Fransson, P., Aden, U., Blennow, M., & Lagercrantz, H. (2011). The functional architecture of the infant brain as revealed by resting-state FMRI. *Cerebral Cortex,* 21, 145-154. doi:10.1093/cercor/bhq071.

Gabrieli, J. D. E. (2009). Dyslexia: A new synergy between education and cognitive neuroscience. *Science*, *325*, 280-283.

Glushko, R. J. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance*, *5*, 674-691.

Gopnik, M. (1990). Feature blind grammar and dysphasia. *Nature*, *344*, 715.

Gopnik, M., & Crago, M. (1991). Familial aggregation of the developmental language disorders. *Cognition, 39*, 1-50.

Harm, M. W., & Seidenberg, M.S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review*, *106*, 491-528.

Harm, M. W., McCandliss, B. D., & Seidenberg, M. S. (2003). Modeling the successes and failures of interventions for disabled readers. *Scientific Studies of Reading*, *7*, 155-182.

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).

Hofbauer, J., & Sigmund, K. (1998). *Evolutionary games and population dynamics.* Cambridge University Press.

Hu, X., Guo, L., Zhang, D., Li, K., Zhang, T., Lv, J., . . . Liu, T. (2011). Assessing the dynamics on functional brain networks using spectral graph theory. In *Biomedical imaging: From nano to macro, 2011 IEEE international symposium.* April, 2011, Chicago, IL, USA.

Joanisse, M. F. (2004). Specific language impairments in children: Phonology, semantics and the English past tense. *Current Directions in Psychological Science*, *13*, 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*, *96*, 7592–7597.

Jolles, D. D., van Buchem, M. A., Crone, E. A., & Rombouts, S. A. R. B. (2011). A comprehensive study of whole-brain functional connectivity in children and young adults. *Cerebral Cortex*, *21*, 385.

Kail, R. (1994). A method for studying the generalized slowing hypothesis in children with specific language impairment. *Journal of Speech and Hearing Research*, *37*, 418-421.

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

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*. Portland, Oregon, USA.

Karmiloff-Smith, A. (1998). Development itself is the key to understanding developmental disorders. *Trends in Cognitive Sciences*, *2*, 389-398.

Kriete, T., & Noelle, D. C. (submitted). *Dopamine and the development of executive dysfunction in autism spectrum disorders*. Manuscript submitted for publication.

Kujala, T., Karma, K., Ceponiene, R., Belitz, S. Turkkila, P. Tervaniemi, M., & Naatanen, R. (2001). Plastic neural changes and reading improvement caused by audiovisual training in reading-impaired children. *Proceedings of the National Academy of Sciences*, *98)*, 10509-10514.

Leonard, L. B. (1999). *Children with specific language impairment*. Cambridge, MA: The MIT Press.

Leonard, L. B., Bortolini, U., Caselli, M. C., McGregor, K. K., & Sabbadini, L. (1992). Morphological deficits in children with specific language impairment: The status of features in the underlying grammar. *Language Acquisition*, *2*, 151-179.

Liu, Y., Liang, M., Zhou, Y., He, Y., Hao, Y., Song, M., . . . Jiang, T. (2008). Disrupted small-world networks in schizophrenia. *Brain: A Journal of Neurology*, *131*, 945.

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

Mareschal, D., Johnson, M. H., Sirois, S., Spratling, M. W., Thomas, M. S. C., & Westermann, G. (2007). *Neuroconstructivism: How the brain constructs cognition, Volume One*. Oxford: Oxford University Press.

McCandliss, B. D., Beck, I., Sandak, R., & Perfetti, C. (2003). Focusing attention on decoding for children with poor reading skills: Design and preliminary tests of the word building intervention. *Scientific Studies of Reading*, *7*, 75–104.

McClelland, J. L., & Rumelhart, D. E. (1988). *Computational models of cognition and perception: Explorations in parallel distributed processing: A handbook of models, programs, and exercises.* Cambridge, MA: The MIT Press.

Menon, V. (2010). Large-scale brain networks in cognition: Emerging principles. In: O. Sporns (Ed.), *Analysis and Function of Large-Scale Brain Networks*, (pp. 43-53). Washington, DC: Society for Neuroscience.

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

Plunkett, K., & Juola, P. (1999). A connectionist model of English past tense and plural morphology. *Cognitive Science*, *23*, 463–490.

Plunkett, K., & Marchman, M. (1991). U-shaped learning and frequency effects in a multi-layered perceptron: Implications for child language acquisition. *Cognition*, *38*, 43-102.

Rezazadeh, S. M., Wilding, J., & Cornish, K. (2011). The relationship between measures of cognitive attention and behavioral ratings of attention in typically developing children. *Child Neuropsychology*, *17*, 197-208.

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

Richardson, F., & Thomas, M. S. C. (2006). The benefits of computational modelling for the study of developmental disorders: extending the Triesch et al. model to ADHD. *Developmental Science*, *9*, 151–155.

Rougier, N. P., Noelle, D. C., Braver, T. S., Cohen, J. D., & O'Reilly, R. C. (2005). Prefrontal cortex and flexible cognitive control: Rules without symbols. *Proceedings of the National Academy of Sciences*, *102*, 7338-7343.

Seidenberg, M. S. & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*, 523-568.

Shallice, T. (1988). *From neuropsychology to mental structure.* Cambridge University Press.

Spencer, J. P., Perone, S., & Johnson, J. S. (2009). The dynamic field theory and embodied cognitive dynamics. In: Spencer, Thomas, & McClelland (Eds.), *Toward a new unified theory of development: Connectionism and dynamical systems theory re-considered*, (p. 146-202). Oxford: Oxford University Press.

Spencer, J., Thomas, M. S. C., & McClelland, J. L. (2009). *Toward a new unified theory of development: Connectionism and dynamical systems theory re-considered*. Oxford: Oxford University Press.

Stam, C. J., de Haan, W., Daffertshofer, A., Jones, B. F., Manshanden, I., van Cappellen van Walsum, A. M., . . . Scheltens, P. (2009). Graph theoretical analysis of magnetoencephalographic functional connectivity in Alzheimer's disease. *Brain: A Journal of Neurology*, *132*(Pt 1), 213-24. doi:10.1093/brain/awn262

Stavrakaki, S., Koutsandreas, K., & Clahsen, H. (in press). The perfective past tense in Greek children with specific language impairment. *Morphology*.

Stephany, U. (1997). The acquisition of Greek. In D. I. Slobin (Ed.), *The cross-linguistic study of language acquisition (Vol. 4)*. Hillsdale, NJ: Erlbaum.

Stiles, J. (2001). Spatial cognitive development. In C. A. Nelson & M. Luciana (Eds.), *Handbook of developmental cognitive neuroscience* (pp. 399–414). Cambridge, MA: MIT Press.

Tallal, P., & Piercy, M. (1973a). Defects of non-verbal auditory perception in children with developmental aphasia. *Nature*, *241*, 468-469.

Tallal, P., & Piercy, M. (1973b). Developmental aphasia: Impaired rate of non- verbal processing as a function of sensory modality. *Neuropsychologia*, *11*, 389-398.

Tallal, P., Miller, S., & Fitch, R. (1993). Neurobiological basis of speech: A case for the pre-eminence of temporal processing. *Annals of the New York Academy of Sciences*, *682*, 27-47.

Temple, C. M. (1997). Cognitive neuropsychology and its application to children. *Journal of Child Psychology and Psychiatry*, *38*, 27-52.

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action.* Cambridge, MA: MIT Press.

Thelen, E., Schöner, G., Scheier, C., & Smith, L. B. (2001). The dynamics of embodiment: A field theory of infant perseverative reaching. *Behavioral and Brain Sciences*, *24*, 1-34.

Thomas, M. S. C. (2003). Limits on plasticity. *Journal of Cognition and Development*, *4*, 95-121.

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

Thomas, M. S. C. (2006). Williams syndrome: Fractionations all the way down? *Cortex*, *42*, 1053-1057.

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

Thomas, M. S. C., & Karmiloff-Smith, A. (2002). Are developmental disorders like cases of adult brain damage? Implications from connectionist modelling. *Behavioural and Brain Sciences*, *25*, 772-780.

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

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., & Redington, M. (2004). Modelling atypical syntax processing. In W. G. Sakas (Ed.), *COLING 2004: Psycho-Computational Models of Human Language Acquisition* (p. 87-94). Geneva, Switzerland: COLING.

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*, 143-176.

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.

Torgesen, J. K. (2004). Preventing early reading failure. *American Educator*, *28*, 6-9.

Triesch, J., Teuscher, C., Deak, G., & Carlson, E. (2006). Gaze following: why (not) learn it? *Developmental Science*, *9*, 125-147.

Ullman, M., & Pierpont, E. (2005). Specific Language Impairment is not specific to language: The Procedural Deficit hypothesis. *Cortex*, *41*, 399-433.

van den Heuvel, M. P., & Pol, H. E. H. (2010). Exploring the brain network: A review on resting-state fmri functional connectivity. *EUR NEUROPSYCHOPHARM*, *20*(8), 519-534. doi:10.1016/j.euroneuro.2010.03.008.

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.

van der Maas, H. L., & Molenaar, P. C. (1992). Stage-wise cognitive development: An application of catastrophe theory. *Psychological Review*, *99*, 395-417.

van Geert, P. (1991). A dynamic systems model of cognitive and language growth. *Psychological Review*, *98*, 3-53.

Vernes, S. C., Newbury, D. F., Abrahams, B. S., Winchester, L, Nicod, J., Groszer, M., . . . Fisher, S.E. (2008). A functional genetic link between distinct developmental language disorders. *New England Journal of Medicine*, *359*, 2337-45.

Wang, L., Zhu, C., He, Y., Zang, Y., Cao, Q. J., Zhang, H., . . . Wang, Y. (2009). Altered small-world brain functional networks in children with attention-deficit/hyperactivity disorder. *Human Brain Mapping*, *30*, 638-649.

Williams, J. O. H., & Dayan, P. (2005). Dopamine, learning, and impulsivity: a biological account of ADHD. *Journal of Child and Adolescent Psychopharmacology*, *15*, 160–179.

Figure captions

Figure 1. Architecture of the Multiple Inflection Generator (MIG), which integrates multiple cues to output the phonological form of content words appropriate to the sentential grammatical context (Karaminis & Thomas, 2010). The model has been applied to simulating the acquisition of inflectional morphology in English and in Greek, for both typically developing children and children with SLI (Karaminis, 2011).

Figure 2. (a) The non-developmental Dual-Route Cascaded model of reading (Coltheart et al., 2001). (b) The 'triangle' model of reading development (Seidenberg & McClelland, 1989).

Figure 3: Candidate architectures for multi-component cognitive systems

Figure 4: The mutualism model, applied to the study of disorders (Baughman & Thomas, 2008)

Figure 5: Developmental trajectories for each architecture, distinguishing between key and peripheral processes

Figure 6: A simulated typical developmental trajectory with upper and lower bounds

Figure 7: The connectivity of ‘small world’ architectures: (a) connections are limited to processes that are near each other; (b) and (c): small-world networks with additional random connections that shorten pathways.

Figures

Figure 1.

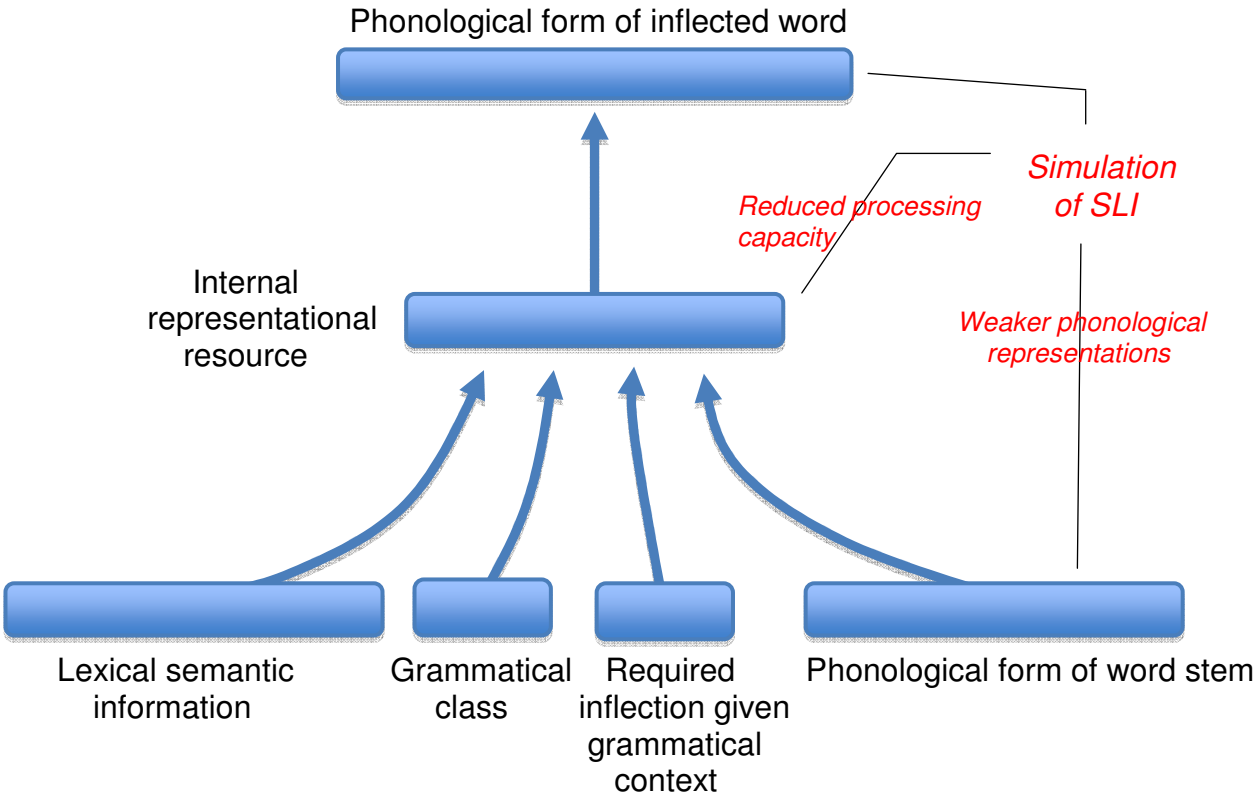
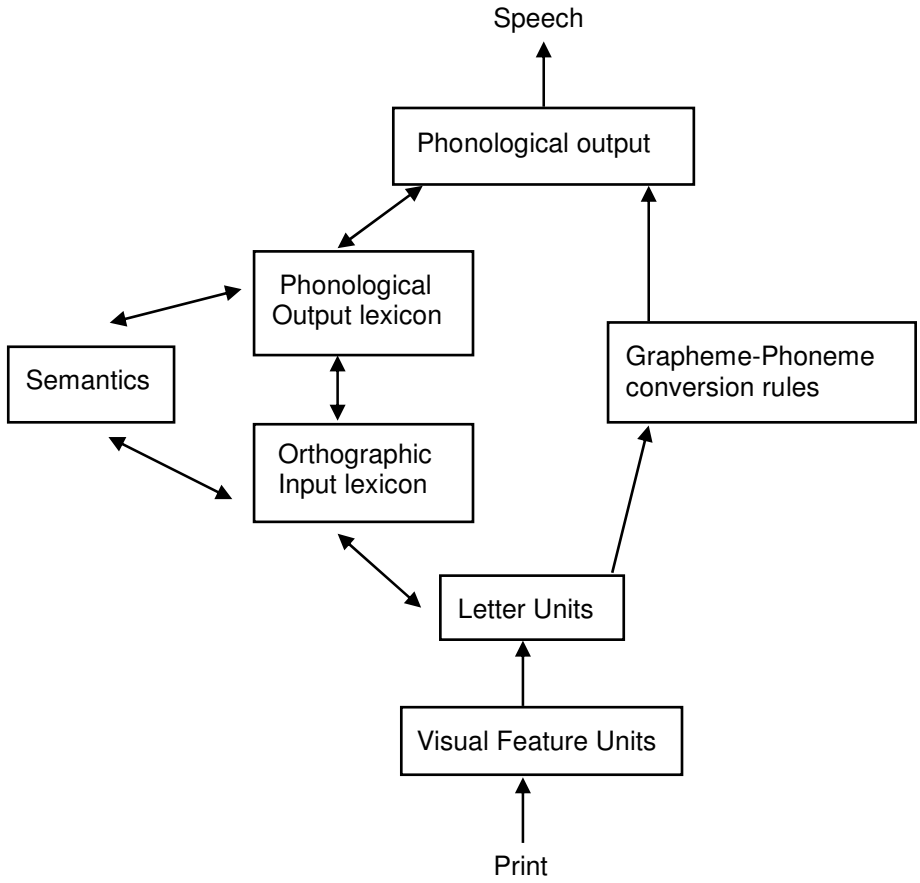

Phonological form of inflected word
Simulation of SLI
Reduced processing capacity
Internal representational resource
Weaker phonological representations
Lexical semantic information
Grammatical class
Required inflection given grammatical context
Phonological form of word stem

Figure 2

(a)


Speech
Phonological output
Phonological
Output lexicon
Semantics
Grapheme-Phoneme
conversion rules
Orthographic
Input lexicon
Letter Units
Visual Feature Units
Print

(b)

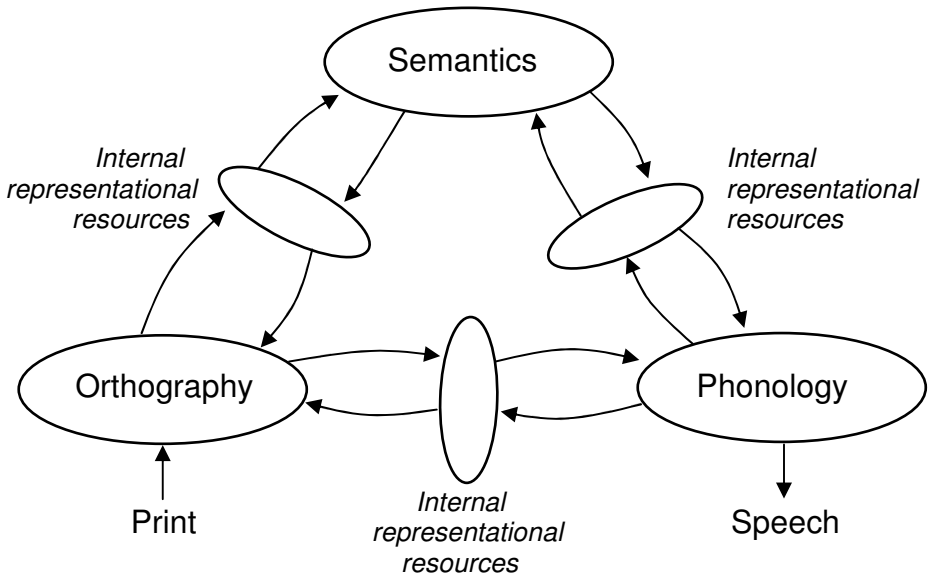
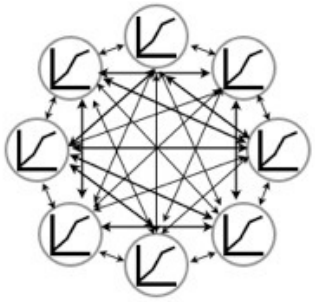
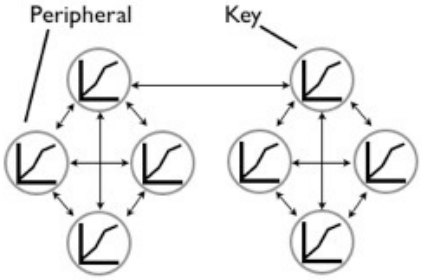

Semantics
Internal
representational
resources
Internal
representational
resources
Orthography
Phonology
Print
Internal
representational
resources
Speech

Figure 3

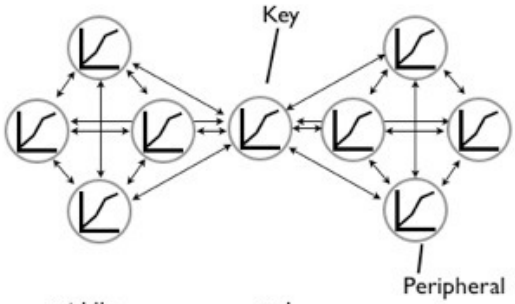
(a) Fully distributed



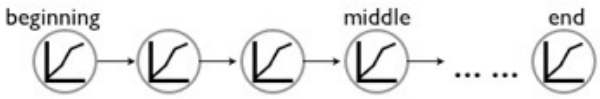
(b) Hemispheric



(c) Central processor



(d) Hierarchical



(e) Modular

Figure 4

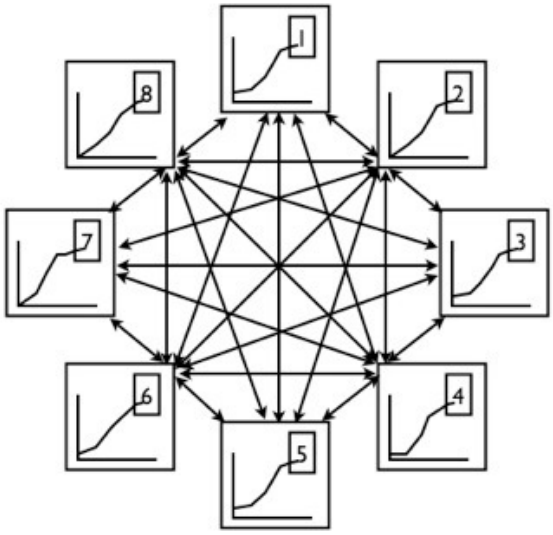


Figure 5

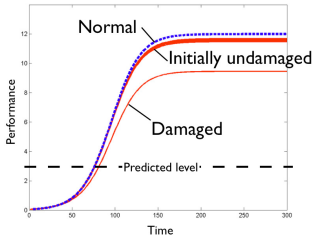
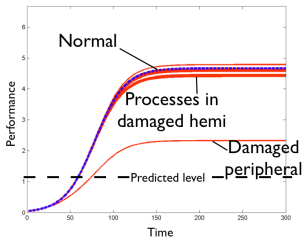
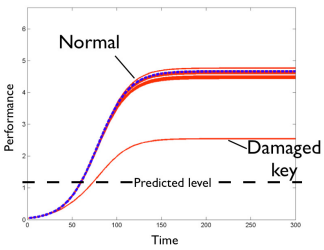
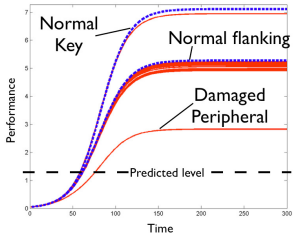
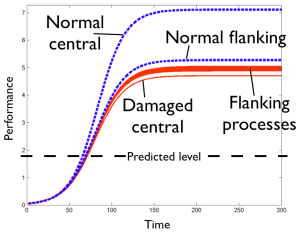
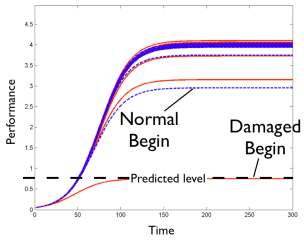
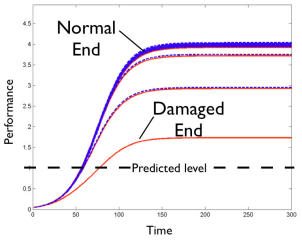
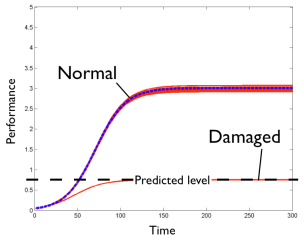
Architecture	Analysis of peripheral deficit	Analysis of key deficit
Distributed		All processes equal
Hemispheric		
Central processor		
Hierarchical		
Modular		All processes equal

Figure 6

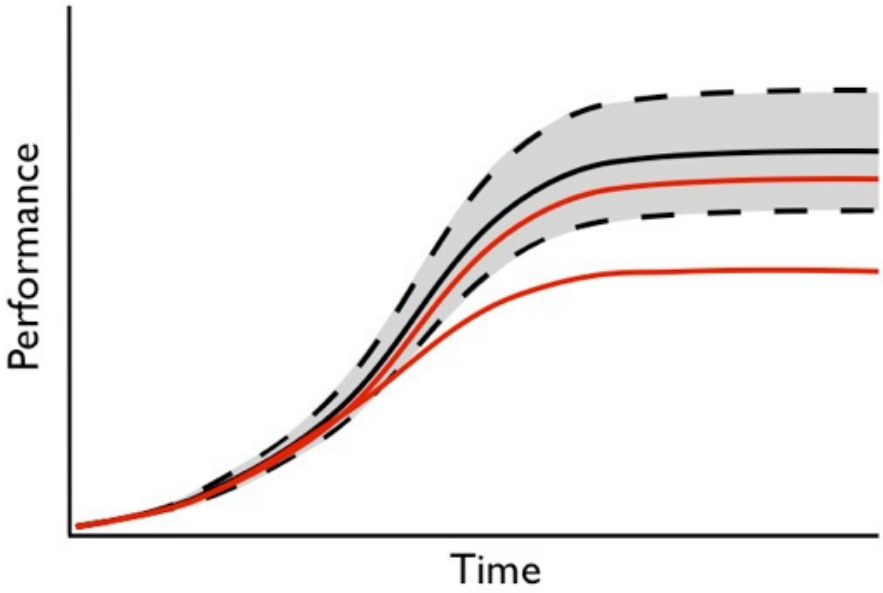


Figure 7

(a) (b) (c)