

Behavioral and Brain Sciences

<http://journals.cambridge.org/BBS>

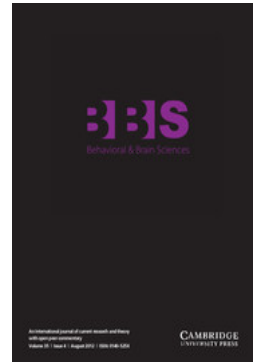
Additional services for *Behavioral and Brain Sciences*:

Email alerts: [Click here](#)

Subscriptions: [Click here](#)

Commercial reprints: [Click here](#)

Terms of use : [Click here](#)



Are developmental disorders like cases of adult brain damage? Implications from connectionist modelling

Michael Thomas and Annette Karmiloff-Smith

Behavioral and Brain Sciences / Volume 25 / Issue 06 / December 2002, pp 727 - 750

DOI: 10.1017/S0140525X02000134, Published online: 11 August 2003

Link to this article: http://journals.cambridge.org/abstract_S0140525X02000134

How to cite this article:

Michael Thomas and Annette Karmiloff-Smith (2002). Are developmental disorders like cases of adult brain damage? Implications from connectionist modelling. Behavioral and Brain Sciences, 25, pp 727-750 doi:10.1017/S0140525X02000134

Request Permissions : [Click here](#)

Are developmental disorders like cases of adult brain damage?

Implications from connectionist modelling

Michael Thomas

Neurocognitive Development Unit, Institute of Child Health and School of Psychology, Birkbeck College, University of London, London WC1E 7HX, United Kingdom

m.thomas@psychology.bbk.ac.uk

Annette Karmiloff-Smith

Neurocognitive Development Unit, Institute of Child Health, London WC1N 1EH, United Kingdom

a.karmiloff-smith@ich.ucl.ac.uk

http://www.ich.ucl.ac.uk/units/ncdu/NDU_homepage.htm

Abstract: It is often assumed that similar domain-specific behavioural impairments found in cases of adult brain damage and developmental disorders correspond to similar underlying causes, and can serve as convergent evidence for the modular structure of the normal adult cognitive system. We argue that this correspondence is contingent on an unsupported assumption that atypical development can produce selective deficits while the rest of the system develops normally (Residual Normality), and that this assumption tends to bias data collection in the field. Based on a review of connectionist models of acquired and developmental disorders in the domains of reading and past tense, as well as on new simulations, we explore the computational viability of Residual Normality and the potential role of development in producing behavioural deficits. Simulations demonstrate that damage to a developmental model can produce very different effects depending on whether it occurs prior to or following the training process. Because developmental disorders typically involve damage prior to learning, we conclude that the developmental process is a key component of the explanation of endstate impairments in such disorders. Further simulations demonstrate that in simple connectionist learning systems, the assumption of Residual Normality is undermined by processes of compensation or alteration elsewhere in the system. We outline the precise computational conditions required for Residual Normality to hold in development, and suggest that in many cases it is an unlikely hypothesis. We conclude that in developmental disorders, inferences from behavioural deficits to underlying structure crucially depend on developmental conditions, and that the process of ontogenetic development cannot be ignored in constructing models of developmental disorders.

Keywords: Acquired and developmental disorders; connectionist models; modularity; past tense; reading

1. Introduction

Behavioural impairments found in developmental disorders and in cases of acquired brain damage provide a source of information about the structure of the cognitive system. Historically, the logic of deriving implications about cognitive structure from behavioural impairments was formulated in the domain of acquired disorders in adults (see, e.g., Shallice 1988). It was argued that under some circumstances, highly selective patterns of impairment after damage could demonstrate the relative independence of different cognitive processes, predicated on an a priori assumption of modular structure within the cognitive system. Ultimately, this was thought to lead to the identification of the components of cognition.

Recently, behavioural impairments found in developmental disorders have often been interpreted within the same

cognitive neuropsychology framework (see, e.g., Baron-Cohen 1998; Leslie 1992; Temple 1997). In this case, there is an inference that selective behavioural impairments reveal discrete components of the cognitive system that have not developed properly, for example, the purported defective “theory of mind” processor in autism (Leslie 1992), or the defective phonological processor in dyslexia (Frith 1995). However, the extension of the cognitive neuropsychology framework to interpret developmental disorders has proved controversial. Indeed, some researchers (Bishop 1997a; Karmiloff-Smith 1997; 1998) have argued that the process of development itself violates key assumptions of the static cognitive neuropsychology model and thus invalidates the direct inference from impairment to cognitive structure.

Our aim in this article is to evaluate this debate from the perspective of connectionist modelling of cognitive processes. This is a useful perspective because such models

have been employed to capture both acquired deficits (when models of adult performance are damaged) and developmental deficits (when initial computational constraints are altered in models of typical development). Connectionist models therefore provide a concrete computational basis on which to anchor a debate on the potential causes of each type of deficit.

From the stance of the behavioural outcome, the impairments found in cases of acquired and developmental disorders can look very similar. For example, in relation to types of dysgraphia, dyslexia, and dyscalculia, Temple (1997, p. 324) comments that “were one to give the data from adult and child cases to a cognitive neuropsychologist and ask the question, which are the adults and which the children, there are no apparent criteria by which to distinguish them.”

One might take this similarity in behaviour as an indication that the two types of disorder are linked at a deeper level, namely, that they share a similar underlying cause. For example, where one can appeal to a static information-processing model of the adult system, one might characterise an impairment in the adult case as corresponding to selective damage to one (or more) processing components, and an impairment in the developmental case as a failure of one (or more) components to be properly acquired. Temple (1997) offers just such a characterisation for cases of developmental prosopagnosia (pp. 139 and 141), as well as two subtypes of developmental dyslexia (pp. 192 and 206), developmental disorders of spelling (pp. 238 and 244), and developmental dyscalculia (pp. 285–86).

In this target article, we argue that such a causal link between acquired and developmental disorders can only occur if, for a given domain, a very particular kind of developmental account holds true. In most cases where researchers have linked acquired and developmental disorders, the required developmental account has not been argued for, but merely assumed.

Our aim in the computational part of this article (sect. 5)

is to characterise the conditions that must hold in a developing cognitive system for acquired and developmental disorders to be linked at a causal level. We demonstrate by simulation that, in the absence of a precise developmental account of a cognitive system, behavioural data alone may be insufficient to infer underlying functional structure from a pattern of impairments. As a result, we argue that researchers working with developmental disorders must compare their data against developmental rather than static models, even though those static models may be appropriate for explaining patterns of acquired deficits in normal adults.

First, however, we introduce two concrete examples of domains in which explicit links have been drawn between acquired and developmental impairments. These are the domains of dyslexia and English past tense formation. For current purposes, these areas are important not only because they illustrate how explicit the claims have been about the relation between acquired and developmental impairments, but also because both areas have been the focus of substantial computational modelling work exploring the possible underlying causes of those impairments.

2. Comparisons of acquired and developmental deficits in two domains

2.1. Dyslexia

When adults experience difficulty in reading following brain damage, their patterns of behavioural impairments can be described according to several subtypes. Two subtypes are of particular relevance. In acquired phonological dyslexia, patients demonstrate particular difficulty in reading nonwords. In acquired surface dyslexia, they show a deficit in reading *exception* words, where the pronunciation cannot be predicted from the usual letter-to-sound correspondence. For these exception words, patients tend to display errors of regularisation, for example, *gauge* pronounced as “gorge,” *trough* as “truff,” *come* as “kome,” and *quay* as “kway” (Shallice et al. 1983).

Cognitive neuropsychologists have interpreted these two patterns as reflecting specific damage to independent sub-components of the skilled reading system. The traditional information-processing model of the skilled reading system proposes that three processing routes link print to sound (see, e.g., Patterson & Shewell 1987; Temple 1997). One route decomposes written words into their component graphemes and constructs a pronunciation via a system of grapheme-to-phoneme correspondences. This is called the *nonlexical* or *phonological* route. A second *direct* or *lexical* route recognises the whole written word form and uses this representation to recover the whole-word pronunciation. A third, *semantic*, route uses the written word form to recover the word’s meaning, and this semantic representation is then used to recover the word’s pronunciation.

Assuming this model, acquired phonological dyslexia can be interpreted as a normal adult reading system that has experienced damage to the grapheme-to-phoneme processing route; existing words can be read via the whole-word recognition routes, but reading of nonwords is impaired. Acquired surface dyslexia, however, can be interpreted as an adult system that has experienced damage to both whole-word recognition routes. Words can only be read via decomposition into component graphemes and the appli-

MICHAEL THOMAS completed his D. Phil at University of Oxford on bilingual language processing and connectionist modelling. Previously a Senior Research Fellow at the Neurocognitive Development Unit, Institute of Child Health, London, he has recently accepted a lectureship in Cognitive Neuropsychology at Birkbeck College, University of London. His research focuses on typical and atypical language development, combining empirical and computational approaches to mechanisms of learning.

ANNETTE KARMILOFF-SMITH is Professor of Neurocognitive Development at the Institute of Child Health, University College London, where she heads the Neurocognitive Development Unit. Originally a student and then collaborator of Piaget at Geneva University, she is author of 8 books, over 200 articles/chapters on infant/child cognitive development and genetic disorders, and in 1995 won the British Psychological Society’s book award for excellence in the literature of psychology. Her writings have been translated into Japanese, Chinese, Hungarian, Hebrew, Arabic, Italian, Spanish, French, German, Danish, Norwegian, Dutch and Greek. In the 1990s she was elected Fellow of the British Academy and of the Academy of Medical Sciences.

cation of grapheme-phoneme correspondences, resulting in the regularisation of exception words.

The patterns of errors defining these two subtypes of dyslexia have been reported in children, both in single case studies and in group studies (see, e.g., Castles & Coltheart 1993; Manis et al. 1996). Researchers have employed similar explanations in the developmental case – once more appealing to the structure of the adult reading system, but now replacing the notion of “specific damage” with the notion of “a specific failure to develop.” Therefore, in developmental phonological dyslexia, children may be “having difficulty with [acquiring] one or more components of the nonlexical route” (Coltheart et al. 2001) or they may exhibit an overall system with “relatively normal development of semantic, lexical, and direct reading systems but with impairment in the acquisition of the phonological reading route” (Temple 1997, p. 206). In developmental surface dyslexia, children may be having difficulty acquiring “one or more components of the lexical route” (Coltheart et al. 2001), or they may exhibit a reading system in which the “direct and semantic reading routes have failed to become established properly” (Temple 1997, p. 192).

2.2. Past tense formation

Our second example comes from the domain of inflectional morphology, and in particular the formation of the English past tense. Once again, a model has been proposed for the functional structure of the adult system in which separate subcomponents tackle different aspects of the task (Pinker 1991; 1999). One component is claimed to be responsible for forming the majority of past tenses that conform to a rule (“add -ed”) and for generating past tenses for novel verbs (wug-wugged). A second component memorises individual past tense forms, particularly those that are exceptions to the rule (e.g., go-went, sleep-slept, hit-hit, etc.).

In cases of acquired aphasia and in neurodegenerative diseases, adults can exhibit dissociations between performance on regular and exception past tense formation. Patients with nonfluent aphasia can be worse at producing and reading regular past tense forms than exception forms, whereas patients with fluent aphasia can be worse at producing and reading exception forms than regular forms (e.g., Tyler et al. 2002a; 2002b; Ullman et al. 1997; in press; though, see Bird et al. 2002). Similarly, patients with Parkinson’s disease can make more errors producing regular and novel -ed forms than exception forms, whereas patients with Alzheimer’s disease can make more errors producing exception past tense forms than regular past tense forms (Ullman et al. 1997; Ullman, in press). Assuming the dual-mechanism model of the adult system, these patterns of acquired deficit are taken to reflect selective damage to either the rule-processing component or the exception memorisation component.

Once again, parallel impairments have been reported in the developmental domain, in this case in two developmental disorders with a genetic origin, Specific Language Impairment (SLI) and Williams syndrome (WS). Ullman and Gopnik (1999) and van der Lely and Ullman (2001) reported that children with SLI perform poorly on past tense formation tasks and show a much smaller advantage of regular past tense formation over exception past tense formation – interpreted as a relative impairment in regular past formation. However, Clahsen and Almazan (1998) reported that children with WS exhibit a specific difficulty with gen-

erating exception past tense forms.¹ Pinker (1999) offered an interpretation of these respective findings in terms of the adult model: SLI represents a case where the mutation of certain genes interferes with the development of the ability to inflect new and uncommon regular verbs. WS represents a case where the rule-based computational mechanism is intact but the memory mechanism for storing exception verbs is specifically impaired. Together these disorders are argued to represent a “genetic double dissociation . . . the first group of children rarely generalise the regular pattern; the second group of children generalise it freely” (Pinker 1999, p. 262).

Both the examples of dyslexia and past tense formation illustrate the way in which developmental impairments are often interpreted by appealing to the structure of adult models. They show, too, how the central double dissociation logic of adult cognitive neuropsychology has been extended to developmental cases. Dissociable behavioural impairments are taken as evidence of independent underlying mechanisms, by virtue of the claimed independent failure of those mechanisms to develop properly. We turn now to consider why the validity of this extension is questionable, before examining specific computational implementations of deficits in these target domains.

3. Is the cognitive neuropsychology framework appropriate for the interpretation of developmental disorders?

3.1. Development in a “static” framework

When acquired damage causes selective cognitive deficits in normal adults, these deficits occur against a background of hitherto normal function. (This is also the case for acquired deficits in children, at least at the time of insult.) Such cognitive systems are hence discussed in terms of the cognitive mechanisms or processes that have become *impaired* compared to those that have remained *intact*.

When selective behavioural deficits are identified in developmental disorders, they are frequently characterised in the same way, in terms of developmental impairments against a background of normal development. We will refer to the second half of this characterisation as the *assumption of Residual Normality*. This is the assumption that, in the face of a selective developmental deficit, the rest of the system can nevertheless develop normally and independently of the deficit. It is this developmental assumption that allows researchers to relate patterns of deficits in developmental disorders to static models of the normal cognitive system. Because patterns of deficits are usually identified in older children, adolescents, or adults with the developmental disorder, static models of the normal adult system are often deemed an appropriate point of reference. In principle, however, deficits identified in the younger child could be compared against a static model of the normal system for the appropriate stage in development, were such a model to exist. In either case, the essential point here is that the assumption of Residual Normality permits developmental deficits to be compared against functional models that *themselves have no developmental component*.

The assumption of Residual Normality has been widely deployed in the study of atypical development, including in the case of disorders such as autism, WS, SLI, dyslexia, dyscalculia, Gilles de la Tourette syndrome, and develop-

mental prosopagnosia. The following quotes illustrate three explicit renditions of the claim for Residual Normality in developmental disorders:

I suggest that the study of mental retardation would profit from the application of the framework of cognitive neuropsychology (e.g., McCarthy & Warrington 1990; Shallice 1988). In cognitive neuropsychology, one key question running through the investigator's mind is "Is this process or mechanism intact or impaired in this person?" . . . In fact researchers in mental retardation have been searching for *intact* versus impaired cognitive processes for quite some time without discussing this in terms of modularity. (Baron-Cohen 1998, p. 335, emphasis added; and Footnote 1)

The analysis of the developmental dyslexias offered by the dual-route model is that, just as each of the two routes can be selectively affected by brain damage with the other remaining intact, it is possible for a child to have difficulty acquiring one of the routes, *with the other being acquired at a normal rate*. (Coltheart et al. 1993, p. 591, emphasis added)

Within modular theories, the linguistic performance of subjects with [developmental] language impairments may reflect *the architecture of the normal system* but with selective components of this system under- or over-developed. (Clahsen & Temple 2002, emphasis added)

It is interesting that Residual Normality (henceforth RN) is less frequently deployed as a developmental hypothesis in pediatric neuropsychology. For children with acquired brain damage, the clinically driven focus is usually on recovery. Researchers tend to eschew static models and explore the effect of cerebral insult on the potentially plastic process of development. Because structural damage is seen in the context of a dynamic and interactive developmental process, there is recognition of the possible influences of compensation within the cognitive system and of disruption to the acquisition of further cognitive skills, as well as family and social factors (Anderson et al. 2001b; see Thomas 2003, for discussion). If the undamaged part of the cognitive system compensates or alters across development in response to the part that has suffered a selective deficit, the undamaged part may not follow the normal path of development, in which case RN would not hold.²

However, researchers in developmental disorders of a genetic origin routinely deploy RN in their explanations, as we have seen in the cases of SLI, WS, and dyslexia. This is probably because of the fact that such disorders are used (in part) for theoretical purposes within the cognitive neuropsychology framework, as a source of evidence about the structure of the normal cognitive system (and because of the genetic origins, about the potential innateness of that structure). RN is an assumption (often implicit) about how development takes place. But, is it likely to be correct? In section 3.2, we consider two opposite claims. First, we examine the claim that *no answer to the preceding question is necessary*: From the perspective of cognitive neuropsychology, development can be ignored in the study of behavioural deficits in developmental disorders. Second, we consider the claim that, not only must development be incorporated, but that when it is, the assumptions required to use the cognitive neuropsychology framework are fatally undermined. We then propose a resolution of these two opposing positions.

3.2. Development and modularity

Jackson and Coltheart (2001) have recently defended the use of the cognitive neuropsychology framework for study-

ing developmental disorders. They have argued that the process of development is not relevant to identifying intact and impaired processes in a cognitive system, so long as modularity can be assumed for that system. In their view, the framework is equally suitable in both acquired and developmental cases for establishing what they call the *proximal cause* of the behavioural impairment. By this they mean "what is wrong with the cognitive system right now," irrespective of whether the original cause was brain damage, atypical development caused by a genetic abnormality, or even poor schooling. These latter causes are what Jackson and Coltheart term *distal*. They maintain that the distal causes of an impairment are potentially *independent* of the common proximal cause, allowing one to consider acquired and developmental deficits within the same framework. Although Jackson and Coltheart agree that a full explanation will involve both proximal and distal causes, they defend the cognitive neuropsychology framework as the appropriate way to reveal the proximal cause of any behavioural impairment, independent of distal causes.

Jackson and Coltheart discuss the case of phonological dyslexia, which, as we saw earlier, is defined by a difficulty in reading novel words. They argue that, in relation to the traditional cognitive model of reading, both acquired and developmental phonological dyslexia can be assigned the same proximal cause, namely, a problem with the processing route that maps graphemes to their respective phonemes (the "GPC" route). What differs in the acquired and developmental cases is the distal cause, respectively, brain damage and some developmental (perhaps genetic) anomaly. In short, these authors argue strongly that synchronic similarities in behavioural deficits between acquired and developmental cases *can be linked by a common underlying cause at a cognitive level of description*, and that this cognitive cause can be established by the methods of cognitive neuropsychology.

The extension of the cognitive neuropsychology framework to developmental disorders has, however, been criticised on three main grounds (Bishop 1997a; Karmiloff-Smith 1997; 1998). The first criticism is that the framework unnecessarily warps the types of data that are collected in the developmental case by focusing on the search for specific deficits, only superficially examining areas of presumed intactness. The second criticism is that the framework is unable to comment on one of the key contributory causes of the patterns of behavioural impairments found in developmental disorders, namely, the process of development itself. Where different developmental hypotheses exist for a given impairment, the cognitive neuropsychology framework cannot distinguish between them. The third criticism is that the assumption of a universal modular structure in the cognitive system on which the framework relies may not hold in the developmental case. We briefly look at each claim in turn.

First, Bishop (1997a) has argued that developmental and acquired disorders require empirical approaches with different emphases. Although researchers in adult cognitive neuropsychology look for *single cases* showing *dissociations* between cognitive abilities (as existence proofs of their dissociability), developmental disorders are likely to show patterns of *associated impairments* as a consequence of cascading effects of early deficits on subsequent development. Particular developmental disorders will be best identified by seeking consistent patterns of associated impairments in

group studies. Karmiloff-Smith (1997; 1998) argues that, particularly for disorders of a genetic origin, behavioural impairments at the end of development are likely to be the outcome of an extended atypical developmental trajectory, determined in part by the initial structural anomalies in the cognitive system and in part by the interactions of that system with its environment. She argues that, methodologically, researchers should not search exclusively for selective deficits at the end of the developmental process but should also seek differences in infancy, where the origins of the atypical trajectory may be revealed.

Second, both Bishop and Karmiloff-Smith claim that the cognitive neuropsychology framework is impoverished in the developmental domain by its exclusion of the developmental process as an explanation of patterns of behavioural impairments in the adult state. For Bishop (1997a), these processes include top-down as well as bottom-up interactions between cognitive subsystems during development, compensatory processes, and timing differences that may lead to changes in the patterns of impairment over time. As an example of the latter, she points to the hypothesis that in children with SLI, early problems in auditory discrimination that occur at a crucial stage in language development cause a lasting legacy of language impairment, even if the auditory problems subsequently resolve themselves and are undetectable. Karmiloff-Smith (1997; 1998) argues that the causes of behavioural impairments in developmental disorders are likely to be found in the low-level computational properties of the neonate brain, such as atypical neuronal firing levels or local connectivity. Such low-level differences can only lead to behavioural impairments via the developmental process, a process that may exaggerate some initial computational differences but attenuate others, depending on the nature of the domain. In the former case, the developmental process itself must be considered a key cause of the subsequent impairments.

In response to this criticism, it is worth noting Jackson and Coltheart's (2001) position that the cognitive neuropsychology framework is not designed to comment on distal causes of deficits such as development, merely on the proximal causes, that is, the functional deficits shown in the current state. As such, the cognitive neuropsychology framework simply does not have the power to address the questions of concern to Bishop and Karmiloff-Smith.

The third criticism of the extension of the cognitive neuropsychology framework to developmental disorders is potentially the most serious. The only necessary a priori assumption required to employ that framework is that in many domains the cognitive system is modular, so that selective deficits in behaviour may be traced to independent functional components. Bishop (1997a) has suggested that two of the defining properties of modules – innateness and imperviousness to top-down feedback – are clearly challenged by processes of development. Some cognitive abilities acquired by children are not innate, and top-down processing is used a great deal by children when they are performing cognitive tasks. There are problems with this criticism that concern the precise definition of what constitutes a module. Fodor (1983) identified several possible characteristics of modules (that they be domain specific, innately specified, informationally encapsulated, fast, hard-wired, autonomous, and not assembled). However, none of these was stipulated as a *necessary* property, rather, those *likely* to be associated with modular processing (Coltheart

1999). There has been significant disagreement concerning the key properties of a module, if indeed modularity is to remain a single explanatory concept (Thomas & Karmiloff-Smith 1999). For example, for Fodor (2000), the most important property is encapsulation; for Coltheart, it is domain specificity. Until there is agreement on what constitutes a functional module, it will prove difficult to demonstrate whether development violates the necessary conditions and so clearly undermines the use of the cognitive neuropsychology framework to explain developmental disorders.

However, a more grave criticism lies in wait. Even if we accept a (loose) notion of modularity, it may be that in some types of developmental disorders, individuals do not share the same set of functional modules as in the normal cognitive system. Karmiloff-Smith (1998) argues that neurobiological evidence of the development of neocortex in infants strongly suggests that genes do not code directly for high-level cognitive modules, but that processing structure is emergent and experience-dependent, the *outcome* of a developmental process. The implication is twofold. First, if modular structure is the product of development (Karmiloff-Smith 1992), in cases of atypical development, the resultant modular structure may not be the same as in the normal adult case. Second, even if early damage is limited to a specific cognitive component, if the modular structure of the cognitive system is sensitive to experience, compensation may occur elsewhere in the system, altering the function of the initially intact components.

We can illustrate this idea with reference to Jackson and Coltheart's own example, phonological dyslexia. Recall that these authors attribute the impairment in nonword reading in the developmental and acquired cases of this disorder to a common proximal cause, an impairment to the GPC route in the traditional model of adult performance. However, in the pure case of the disorder, this common proximal cause is actually shorthand for "the GPC route is impaired *and the lexical routes are functioning normally*." In cases of adult brain damage, this seems possible. But in the developmental case, Bishop and Karmiloff-Smith's position is that problems with the GPC route may also lead to differences in the way in which the lexical routes themselves develop. According to this view, the developmentally disordered system could comprise a GPC route and two lexical routes,³ *all* of which are functioning atypically. Together, however, these routes would then manifest a behavioural impairment in nonword reading.

The extent to which modular structure can vary in cases of atypical development is currently an open question (see Tager-Flusberg 2000 for discussion). Indeed, the degree of plasticity across different cognitive systems and their underlying neural substrates is an area of active investigation (Thomas 2003). For example, it remains to be seen whether limits to plasticity are different in cases of acquired damage in early childhood than in individuals with genetic developmental disorders. Nevertheless, if a common modular structure cannot be assumed, it is evident that developmental disorders cannot be straightforwardly related to static models of the normal cognitive system. It is therefore of key importance to understand how modular structure emerges and to what extent this process can be disrupted.

3.3. A reconciliation

Jackson and Coltheart's (2001) claim that it is possible to study independently the endstate of a developmental dis-

order (proximal cause) and the developmental process by which it was reached (distal cause) must be considered carefully. Although it may be possible to *study* them independently, they are not in fact independent but are mutually constraining. It would be unwise to characterise the endstate of a cognitive system in a form that could not be reached by a feasible developmental process (Piaget 1971). This point does not just apply to the study of developmental disorders. In the same way, theories within normal cognitive psychology and normal developmental psychology must be mutually constraining, despite existing as separate fields of inquiry.

Is Residual Normality a feasible type of developmental account? Is it realistic to expect developmental patterns of specific deficits to stand against a background of normal modular function? Karmiloff-Smith and colleagues have argued that a priori, the effects of genetic abnormalities are likely to be widespread throughout the brain and unlikely to be isolated to single high-level cognitive modules (Karmiloff-Smith 1998; Karmiloff-Smith et al. 2002). When marked behavioural deficits arise in a single domain, it is likely that the cognitive processes underlying apparently intact performance in other domains are also atypical in subtle ways – which may go undetected without the sensitive testing of abilities outside of the main behavioural impairment. Such investigations are prompted only by a realistic developmental hypothesis. In support, Karmiloff-Smith cites examples such as Williams syndrome, where ostensibly intact face recognition was subsequently shown to be achieved by atypical cognitive processes (see below), and developmental dyslexia, where motor deficits have been found in children previously thought only to have a highly selective problem in reading (e.g., Bishop 1990; Fawcett et al. 1996; Hill 1998).

Despite a priori leanings, one might view the issue as one to be determined merely on empirical grounds. Are there or are there not selective cognitive deficits in developmental disorders? Unlike modularity, RN is not an assumption that is a priori required to employ the cognitive neuropsychology framework. Instead, it is a hypothesis invoked to explain a particular set of empirical data. In this sense, Jackson and Coltheart are right that the characterisation of the current deficit in a disorder could be blind to the causes of that disorder. Should RN, therefore, be seen as a neutral hypothesis, simply “calling the data”? The answer is no. The reason is that theory and data collection are clearly not independent. Disorders are typically first investigated by application of a range of standardised tests to establish which areas show behavioural deficits and which show behaviour in the normal range. If one has a predilection to believe that RN is true, the risk is that scores in the normal range will be accepted as final evidence of normal underlying processes, and data collection cut short prematurely. If, however, one is more suspicious of RN, as developmentalists usually are given the interactive nature of the developmental process, then there is a motivation to perform more fine-grained analyses to establish whether apparently normal behaviour is actually being achieved by atypical underlying processes. If so, then deficits are not specific.

An example illustrates the point. Despite deficits in visuospatial processing, face recognition in Williams syndrome was initially reported as a “spared” ability, on the basis that scores on standardised tests fell within the normal range (Bellugi et al. 1994; Udwin & Yule 1991). This prompted claims that the development of systems underly-

ing spatial reasoning is disrupted in WS, but the systems underlying face perception develop normally (Pinker 1999). If one were happy to invoke RN, one would stop at this point and perhaps use WS in combination with developmental prosopagnosia as a double dissociation implying the independence of face processing structures from general visuospatial processing.

However, suspicion of RN in genetic developmental disorders actually led to further investigation of this apparently intact ability. Closer examination of the items within the standardised tests on which individuals with WS performed well, and those on which they performed poorly, suggested that their recognition of faces proceeded atypically. Specifically, individuals with WS were better at recognising faces that could be identified by single features than those that required computation of configurations of features; control participants showed no such distinction (Karmiloff-Smith 1997).

Subsequent research with specially designed face stimuli and geometrical patterns supported the hypothesis that face processing follows an abnormal developmental course in WS (Deruelle et al. 1999; Humphreys et al. 2002). Electrophysiological brain imaging studies also indicate anomalous underlying processing in WS, including reduced sensitivity to inverted faces compared to normal faces, and an absence of the progressive developmental pattern of right hemisphere localisation found in typically developing controls (Grice et al. 2001; Mills et al. 2000). In short, when examined in detail, a superficially intact ability turned out to be associated with quite atypical cognitive and brain processes.

In reconciling the two opposing positions, a subtler picture emerges. The static cognitive neuropsychology framework is not *in principle* inappropriate for the study of developmental disorders. For example, there is nothing intrinsic to the nondevelopmental approach espoused by Jackson, Coltheart, and others that would prevent it from verifying whether RN is true in a particular child, domain, or disorder. Empirical data will eventually reveal if there are selective cognitive deficits in a given case. If there are, it will be necessary to construct and test a developmental account in which RN holds. However, a tendency simply to *assume* RN – conditioned by research in adult cognitive neuropsychology where a background of normal function can often indeed be assumed – leads to inadequate data collection. These data are then insufficient to establish beyond reasonable doubt that normal processes underlie behavioural scores that fall within the normal range. This bias has impeded progress in the study of developmental disorders, and particularly in building links to developmental cognitive neuroscience, developmental neurobiology, and, ultimately, the genetic anomalies that underlie many disorders.

This state of affairs has arisen because insufficient attention has been paid to the process of development itself in the study of developmental disorders. In particular, because RN is typically implicitly assumed, there has been no elucidation of the necessary conditions under which it would actually hold.

4. Under what developmental conditions would we expect to see similarities between developmental and acquired disorders?

Acquired deficits and developmental deficits can be related to the same model of the normal cognitive system if a com-

ponent that can be damaged in the endstate can also fail to develop in isolation of the entire developing cognitive system. This will occur under the following conditions:

- The process of development does not contribute to the pattern of behavioural deficits, so damage produces an identical effect at any stage.
- There are pre-existing modules that develop independently; so damage to one does not affect development of the others.
- If it is the case that modules emerge through a developmental process (modularisation), this process can be disrupted in such a way that some modules emerge normally and some develop atypically.

By contrast, similarities between developmental and acquired disorders will *not* occur under the following conditions:

- The process of development always contributes to the precise pattern of deficits in developmental disorders (either attenuating or exaggerating the effects of early damage) and produces patterns not found in acquired damage.
- The system has sufficient plasticity that compensation occurs; undamaged systems take on the function of damaged systems, perhaps suboptimally, perhaps also at some expense to their normal function.
- Module emergence is disrupted to produce a different functional structure that shows a behavioural pattern not found in acquired damage.

Similarities may occur between developmental and acquired disorders that *cannot* be related to the same model of the normal cognitive system under the following conditions:

- Module emergence is disrupted to produce a different functional structure that nevertheless shows a similar *behavioural* pattern to that found in acquired damage.
- The patterns of behavioural breakdown are specified not by the structure of the cognitive system but by features of the problem domain (e.g., in all suboptimal systems, performance degrades on the hard parts of the problem before it degrades on the easy parts).

What conditions *actually* hold in cognitive development? To address this issue, we need to know the answer to several further questions. First, how does the process of development interact with damage to a cognitive system to produce endstate behavioural impairments? Second, does the process of development always play a central role in producing the impairments – and if so, does development tend to produce a different pattern of endstate impairments to acquired damage in the endstate? Third, what is the origin of the specialised functional components stipulated in adult models? If components are not innately specified, how do they emerge through a process of development, as in Karmiloff-Smith's (1992) theoretical notion of emergent modularisation? Finally, how can such a process of emergent specialisation be affected or unaffected by disruption to the computational conditions existing in the early cognitive system?

These are difficult questions, and the field of developmental cognitive neuroscience is some way from having answers to all of them. In what follows, we explore potential answers to these questions by examining computational models of cognitive development. To retain a focus, we restrict our investigation to computational models applied to our two target domains, reading and past tense formation, and to research based on a single, influential class of computational learning systems, that is, connectionist networks.

5. Computer modelling

The modelling section comprises three parts. In the first, we compare the methods that researchers have used to extend connectionist models of normal development and adult function to cases of developmental and acquired impairments in reading and past tense formation. In the second, we introduce new simulations to gauge the contribution of the developmental process to producing patterns of endstate impairments, within the framework offered by those models. Specifically, we investigate the extent to which the process of development itself is a causal factor of the specific pattern of impairments shown in a developmental disorder. In the third, we use new simulations to examine the assumption of Residual Normality. In a system with emergent specialisation of function (i.e., one exhibiting modularisation), how viable is RN? Specifically, when one component of the system is prevented from developing normally, does the rest of the system nevertheless develop independently and normally? If not, what are the conditions under which learning systems *would* show RN, so that developmental impairments could be interpreted in terms of selective deficits to an adult model? From a behavioural perspective, how does assuming RN affect the inferences we can make from dissociations in behaviour to underlying structures?

6. Connectionist models of acquired and developmental deficits

Connectionist networks have been widely used in recent years to construct models of cognitive processing in adults (see, e.g., McLeod et al. 1998; Rumelhart et al. 1986). Because one of the main strengths of these networks is their ability to learn input-output functions, they have increasingly been used to model the development of cognitive processes (Elman et al. 1996; Thomas & Karmiloff-Smith 2002).⁴

When there exists a working model of a normal adult system, the validity of the model can be further tested by investigating its ability to capture patterns of acquired deficits when the model is damaged in various ways. Connectionist models have been used to capture deficits in a number of acquired disorders, including dyslexia, aphasia, alexia, prosopagnosia, epilepsy, phantom limbs, stroke, frontal lobe damage, unipolar depression, Parkinson's disease, Alzheimer's disease, and schizophrenia (see, e.g., Reggia et al. 1996b; Stein & Ludik 1998). Where knowledge is encoded in the connectionist network through a training process, acquired deficits are modelled by damaging the network *after that training process is complete*.

Where connectionist networks have been used to model phenomena within cognitive development, this has permitted the investigation of developmental disorders when development is made to follow an atypical trajectory (Oliver et al. 2000; Thomas & Karmiloff-Smith 2002). Although work in this area is relatively new, models have already been put forward attempting to capture behavioural deficits in developmental dyslexia, SLI, WS, and autism (see Thomas & Karmiloff-Smith 2002 for a review). In developmental models, knowledge is encoded in network systems via a training process, whereby the model aims to simulate both the developmental trajectory and the endstate abilities of the system. In contrast to models of acquired deficits,

changes in models of atypical development are made to the network or to the way it learns *prior to the training process*. These changes in the computational constraints of the learning system lead to atypical trajectories of development and an endstate performance that may exhibit behavioural impairments.

The contrast between connectionist models of acquired and developmental disorders is a fairly clear one. In the acquired case, damage of some sort is applied to the model at the end of a training process. In the developmental case, it is applied prior to the training process.⁵ Immediately one might ask, do modellers use the same kind of damage in each case, and does this damage cause the same behavioural impairment? If the answer to both is yes, one might conclude that the developmental process plays a limited causal role in generating the pattern of behavioural impairments.⁶ However, if the answer is no, and different impairments result from the same damage in the two cases, then the implication would be that the developmental process is an important component in determining the pattern of impairments in developmental disorders. In sections 6.1 and 6.2, we consider this question in relation to connectionist models of reading and past tense formation.

6.1. Connectionist models of reading

Connectionist models of reading assume that the computational problem in this domain is to learn to map between representational codes of the written form of a word, the spoken form of a word, and the word's meaning (Plaut et al. 1996; Seidenberg & McClelland 1989). Typically, this involves three connectionist networks, one to map from orthography to phonology, one to map from orthography to semantics, and one to map from semantics to phonology (although in many models only the first of these networks is implemented; see Harm & Seidenberg 2001 for an exception). Each network has a three-layered structure, comprising an input layer, an output layer, and a layer of hidden units in between. Some models employ recurrent connections that allow cycling activation patterns so that the model will settle into a stable output state. Sometimes a layer of "clean-up" units is connected to the output layer to aid this settling process (see, e.g., Harm & Seidenberg 1999; 2001).

Within these models, acquired dyslexia is produced by different kinds of damage to the trained model. Acquired surface dyslexia, a deficit in reading exception words, has been modelled by damaging the network that maps orthography to phonology, via the removal of hidden units or the severing of connections (Patterson 1990; Patterson et al. 1989). Such damage produces a greater impairment on reading exception words than regular words. However, failures of this approach to fit more extreme patterns of surface dyslexia subsequently led to the claim that exception word reading might be achieved via an indirect semantic route, particularly in the case of low-frequency words. Acquired surface dyslexia might represent damage to this indirect route, so naming must proceed via the orthography-to-phonology route alone, a route that has not learned to name low-frequency exception words and, as a result, regularises them (Patterson et al. 1996; Plaut et al. 1996; see Coltheart et al. 2001 for discussion).

Previously, acquired phonological dyslexia was not explicitly simulated within these learning models because, in theoretical terms, it corresponded to selective damage to

the entire orthography-to-phonology network; because most models only implemented the orthography-to-phonology network itself, such a lesion was outside their scope. Theoretically, lesioning the direct orthography-phonology route would mean that reading must be accomplished primarily or exclusively via the semantic route, so that novel words without a stored meaning would be severely impaired. However, recently Harm and Seidenberg (2001) have implemented the full connectionist reading model, including pathways between phonology, orthography, and semantics. The authors report a manipulation intended to simulate acquired phonological dyslexia, whereby noise is added to processing within the phonological component of the model (i.e., the phonological output units and associated clean-up units). Harm and Seidenberg demonstrate how this impairs the nonword reading of the model much more severely than its reading of words in the training set (Harm, personal communication, June 2001). It also accounts for several effects found in the nonword reading of acquired phonological dyslexics that were previously taken as support for the traditional model of reading.

Doubts have been raised as to the full developmental validity of several of these connectionist reading models. Nevertheless, we consider them here for their insight into systems that acquire representations appropriate for a cognitive domain through a learning process. Patterns of developmental dyslexia have been simulated by applying the relevant damage prior to this learning process. Thus surface dyslexia has been simulated in a number of models by removing units from the hidden layer in the orthography-to-phonology network prior to training (Harm & Seidenberg 1999; Plaut et al. 1996; Seidenberg & McClelland 1989). This damage produces a greater impairment in exception word reading than regular word reading or nonword reading, particularly for low-frequency exception words. Because exception word performance generally lags behind regular word performance during training, some authors have also simulated poorer performance in reading exception words simply by giving the network less training or less efficient training (Bullinaria 1997; Harm & Seidenberg 1999).

Developmental phonological dyslexia has been simulated in two main ways. The first approach reflects a prior claim that developmental phonological dyslexia may correspond to phonological representations (and perhaps orthographic representations as well) that have insufficient componentiality (Manis et al. 1996; Plaut et al. 1996). Harm and Seidenberg (1999) implemented this proposal by restricting the computational properties of the phonological component of their model (the phonological output layer, its recurrent connections, and its clean-up units). Their manipulations included the removal of the clean-up units and severing half the recurrent connections between the phonological units, or restricting the size of the weights in the recurrent connections, or making computations within the phonological component more noisy. All of these resulted in poorer nonword reading, and some impacted on exception word reading as well. Brown (1997) also demonstrated that when both orthographic representations and phonological representations are deliberately constructed with reduced componentiality, reduced nonword reading is found at the end of training.

The second approach seeks to constrain the nature of the computational function that can be learned between orthog-

raphy and phonology. For example, Zorzi et al. (1998) have argued that the reading system is better conceived of as including a network in which orthographic representations are directly connected to phonological representations. Lack of these direct connections in the initial architecture, and the presence only of a route mediated by hidden units, prevented their network from learning a simple function relating orthography and phonology, and so generalisation was reduced. Brown (1997) used another constraint on the computational function by employing several three-layer networks with progressively reduced numbers of hidden units and comparing them when their performance on regular and exception words was matched. (By necessity, this meant that the networks with fewer hidden units had experienced more training.) Networks with fewer hidden units were unable to learn a robust function linking orthography and phonology and so showed poorer nonword reading.

In summary, both surface and phonological dyslexia permit a direct comparison between simulations of the acquired and developmental forms. In the case of surface dyslexia, initial approaches used the *same* method, the removal of hidden units and/or connections, to simulate the same impairment, namely, a deficit in exception word reading. Recently, acquired accounts have appealed to the entire lesion of unimplemented routes. In the case of phonological dyslexia, several methods have been used to simulate the developmental impairment, including either altering phonological and/or orthographic representations, or constraining the computational function that the network can use to link orthographic and phonological codes. One method used by Harm and Seidenberg (1999), the addition of noise to processing within the phonological component of the model during training, was also used by Harm and Seidenberg (2001) to simulate acquired phonological dyslexia, where such noise was added to a normally trained model. Indeed, Harm and Seidenberg (2001) specifically comment that the “form of impairment is identical” in the two cases (p. 80). In short, on the basis of connectionist models of reading, one might conclude that the same form of damage before and after training creates the same behavioural deficit – as if the developmental process itself contributed nothing to the nature of that behavioural deficit.

6.2. Connectionist models of past tense formation

Connectionist models of past tense formation assume that the computational problem in this domain is to learn to map between representational codes of the phonological form of the stem of a verb and a phonological form of that verb's past tense (Plunkett & Marchman 1993; 1996; Rumelhart & McClelland 1986), sometimes in the presence of the verb's semantic representation (Hoeffner 1992; Joanisse & Seidenberg 1999), and sometimes in the presence of more restricted semantic information (MacWhinney & Leinbach 1991; Plunkett & Juola 1999).

Joanisse and Seidenberg (1999) sought to simulate two kinds of acquired deficits in their model of past tense formation, either an impairment in producing regular past tense forms (found in cases of nonfluent aphasia and Parkinson's disease), or an impairment in producing exception past tense forms (found in cases of fluent aphasia and Alzheimer's disease). Impairment to the formation of regular past tenses was achieved by randomly severing connections between the phonological output layer and that layer's bank of

clean-up units in the trained network. However, the fit to patient data was not ideal here, because the model showed a much larger decrement in extending the past tense rule to novel items (e.g., *wug-wugged*) than on the formation of existing regular past tenses (e.g., *talk-talked*). Patients, however, can show similar decrements to both (e.g., Ullman et al. 1997).⁷ It is possible that formation of existing regular past tenses was driven too much by word-specific information because the model was trained only on a single inflectional paradigm. In a larger model in which individual words can be inflected in several different ways, regularities may be pushed further into the phonological part of the network, such that regular verbs would also be amenable to selective damage. Impairment to the formation of exception past tenses, however, was achieved by randomly severing connections between the semantic representations and their clean-up units in the trained model, while adding noise to the semantic activation level. This gave a good fit to patient data (though, see Tyler et al. 2002a).

Developmental problems with regular past tense formation have been reported in SLI, although recently Ullman and collaborators have argued that the deficit is relative, in that the normal advantage for regular verbs over exception verbs is much reduced and most past tense forms are uninflected (Ullman & Gopnik 1999; van der Lely & Ullman 2001). Hoeffner and McClelland (1993) sought to simulate the developmental regular verb deficit by altering the phonological representations of their model prior to training. The phonological representations were changed in line with a hypothesis that individuals with SLI have difficulty processing fast-changing auditory signals, which particularly impairs perception of phonemes such as /t/ and /d/ (e.g., Tallal & Stark 1981; though, see Bishop et al. 1999b; Joanisse & Seidenberg 1998). Both these phonemes are involved in marking the regular past tense form in English. In the model, word final stops and fricatives were given weaker representations in the normal case to reflect their lower salience. In the impaired model, the overall strength of the phonological representations was weakened, exaggerating the disadvantage of word final stops and fricatives. When the model was trained with these altered representations, the result was poorer performance on past tense formation, such that regular past tenses showed a greater impairment than exceptions, and the predominant error pattern was a failure to inflect the verb stem. Moreover, just as in SLI, the model showed an impairment on morphemic phonemes (e.g., the final /d/ in *cared*) but not phonologically identical phonemes that were nonmorphemic (e.g., the final /d/ in *card*). The model was able to produce a differential impairment for regular verbs in its trained state, but did not successfully simulate the very low and equal performance on both regular and exception verbs (see Ullman & Gopnik 1999 for further discussion of the model).

Joanisse (2000) attempted to simulate the pattern of SLI data by applying processing noise to the phonological representations of his past tense model throughout the training process. Here, the model's level of correct performance on regular, exception, and novel verbs was closer to that shown in recent empirical data (van der Lely & Ullman 2001), with low scores on all types. However, the model did not reproduce the predominant error pattern of uninflected stems found in SLI, suggesting that a future model needs to incorporate aspects of both the Joanisse and the Hoeffner & McClelland models.

In Williams syndrome, it was initially reported that there was a selective deficit in forming the past tense of exception verbs (Clahsen & Almazan 1998). However, a larger study suggested that this apparent deficit was actually a consequence of language delay, because performance on exception verbs lags behind that on regular verbs in normal development, and language is typically delayed in WS (Thomas et al. 2001). When language delay was controlled for in this latter study, the greater deficit on exception verbs in WS disappeared. The Thomas et al. study did, however, reveal reduced generalisation of the past tense rule to novel forms in WS, a pattern which persisted even when language delay was controlled for. Using a past tense network that mapped from verb stem to past tense form in the presence of semantic information, Thomas and Karmiloff-Smith (in press) explored the manipulations to the normal model that could reproduce this pattern of developmental data. Various claims have been made that there are subtle deficits to the language system in Williams syndrome. These include the proposals that language development may be “hyper-phonological,” relying to a greater extent on phonological than lexical-semantic information (Grant et al. 1997; Vicari et al. 1996a; 1996b; Volterra et al. 2001), that the phonological representations themselves may be atypical and perhaps rely on sensitive auditory processing (Karmiloff-Smith et al. 1997; Majerus et al. 2001; Neville et al. 1994), that lexical-semantic representations may be atypical (Clahsen & Almazan 1998; Rossen et al. 1996; Temple et al. 2002), or that lexical-semantics may be poorly integrated with phonology (Frawley 2002; Karmiloff-Smith et al. 1998).

Thomas and Karmiloff-Smith found that a manipulation of the phonological representations that reduced their similarity and redundancy was sufficient to reproduce the delay for regular and exception past tense forms as well as the reduction in generalisation. However, the pattern could also be reproduced when noise was added to the information coming from the lexical-semantic system. By contrast, slowed learning failed to produce a reduction in generalisation, suggesting that delay alone was insufficient to explain the data. Although elimination or weakening of the lexical-semantic contribution produced a selective delay (but no final impairment) for exception verbs, it also failed to show the reduction in generalisation. In short, manipulations to phonology or to the integration of phonology and lexical-semantics could simulate the WS data, but a manipulation to lexical-semantics alone could not.

What if the WS data had shown a selective deficit in exceptions as initially reported in the syndrome – could the model have shown this pattern? Performance on exception verbs could be preferentially delayed under at least two conditions: by attenuating lexical-semantic input, or by restricting the computational complexity of the representations the system could learn (e.g., by employing a two-layer network or by reducing the numbers of hidden units by a certain calibrated amount). However, in both cases, the delay was not associated with an endstate impairment. The only way to achieve such a final deficit in exception verbs was to combine manipulations (e.g., attenuating lexical-semantic information and slowing down learning/terminating training at a point where regulars had reached ceiling but exceptions had not, or attenuating lexical-semantic input while restricting computational complexity).

In summary, for inflectional morphology, we have direct

comparisons of attempts to simulate acquired and developmental deficits to both regular and exception verbs. Impairments to regular verbs were simulated by damage to phonology either prior to or following training. It is worth noting that in the Hoeffner and McClelland model, a specific regular deficit in the developmental case was achieved by effectively targeting information that encoded the regular rule. However, a deficit to regular performance in the acquired case was achieved with more general damage (putting aside, for a moment, the fact that acquired damage impaired generalisation of the rule rather than its application to existing verbs). For exception verbs, an acquired impairment was simulated by damaging the input from semantics. Similar damage in a developmental model delayed the learning of exception past tense forms but, importantly, failed to produce an impairment at the end of training. Broadly, then, phonological damage targeted regular inflection/generalisation, whereas semantic damage targeted exception inflections.

6.3. Summary

What can we conclude from the detailed comparison of models of acquired and developmental deficits in these two domains? The results are somewhat contradictory. For both surface and phonological dyslexia, acquired and developmental approaches employed the same kind of damage to produce the same impairment – the intervention of the developmental process did not appear to contribute to the pattern of impairments. For impairments to regular past tense formation, however, more specific damage was required prior to training than at the end of training to generate a specific impairment in regular past tense formation – as if the developmental process risked changing the nature of the impairment. And, indeed, for the impairment of exception past tense forms, damage to semantic input only produced a delay in acquiring these forms, whereas damage at the end of training produced a marked behavioural deficit. In other words, in this case the developmental process overcame initial damage to produce a successful outcome via an altered developmental trajectory.

Our ability to gauge the contribution of the developmental process to the final impairments is compromised by the fact that in each of the preceding cases, the comparisons have involved separate models whose implementations have differed in detail. No models have afforded a direct comparison of the outcome of the same damage carried out prior to versus following the training process. For this reason, in section 7 we describe a simulation designed specifically to make such a direct comparison, something never hitherto undertaken in the literature.

7. Simulation one: Comparing startstate and endstate damage

7.1. Introduction

The design of the following simulations is relatively straightforward. We take a given problem domain and model architecture and train the model on the domain. This establishes its “normal performance.” We then run the model in two conditions. We either damage the model prior to its training process, or damage it following its training process. Any difference in the pattern of impairments in the two cases

must arise from the contribution of the training process, that is, from development.

Three forms of damage are reported: (1) removal of a proportion of connections from the network; (2) addition of noise to the activation levels in the network; and (3) alteration of the discriminability of the processing units, that is, the ability of a unit to produce large changes in its activation state in response to small changes in the input it receives. All manipulations have been widely used in modelling both acquired and developmental deficits. For example, lesioning of network structure has been used to model dyslexia (e.g., Hinton & Shallice 1991; Patterson et al. 1996; Plaut et al. 1996; Plaut & Shallice, 1993), alexia (Mayall & Humphreys 1996), phantom limbs (Spitzer 1996), stroke (Reggia et al. 1996a), Alzheimer's disease (Ruppin et al. 1996), prosopagnosia (Farah et al. 1993), schizophrenia (Hoffman 1996), and autism (Cohen 1998). Addition of noise to processing has been used to model dyslexia (Harm & Seidenberg 1999; 2001), SLI (Joanisse 2000), language in Williams syndrome (Thomas & Karmiloff-Smith, in press), and Alzheimer's disease (Joanisse & Seidenberg 1999). Alteration of unit discriminability has been used to model schizophrenia (Cohen & Servan-Schreiber 1992), executive dysfunction (Levine 1996) and the effects of ageing (Li & Lindenberger 1999).

For comparison with the preceding discussion, we test the contribution of the training process in a domain analogous to past tense formation.

7.2. Simulation details

Architecture: A three-layered feedforward network was used, with the architecture as shown in Figure 1a.

Training set: The training set was taken from Plunkett and Marchman (1993) and comprises an artificial language set constructed to reflect the most important structural features of English past tense formation. There were 500 monosyllabic verbs, constructed using consonant-vowel templates and the phoneme set of English. Phonemes were represented over six articulatory features, and separate banks of units were used to represent the initial, middle, and final phonemes of each monosyllable. The output layer incorporated an additional two features to represent the affix for regular verbs. This corresponds to a network with 18 input units and 20 output units. However, the current simulations involved removing connection weights, and Bullinaria and Chater (1995) have argued that when network models are lesioned, resulting patterns of impairments can be artefactual if very small networks are used. In an attempt to avoid this, the representational scheme was duplicated five times, with the addition of a small amount of noise (whereby the binary features in each duplication had a 20% chance of flipping their state). This preserved the nature of the computational problem faced by the network, but increased the network's size to 90 input units and 100 output units. Fifty hidden units were used in the hidden layer.

There were four types of verbs in the training set: (1) regular verbs that formed their past tense by adding one of the three allomorphs of the +ed rule, conditioned by the final phoneme of the verb stem (e.g., *tame-tamed*, *wrap-wrapped*, *chat-chatted*); (2) exception verbs whose past tense form was identical to the verb stem (e.g., *hit-hit*); (3) exception verbs that formed their past tenses by changing an internal vowel (e.g., *hide-hid*); (4) exception verbs

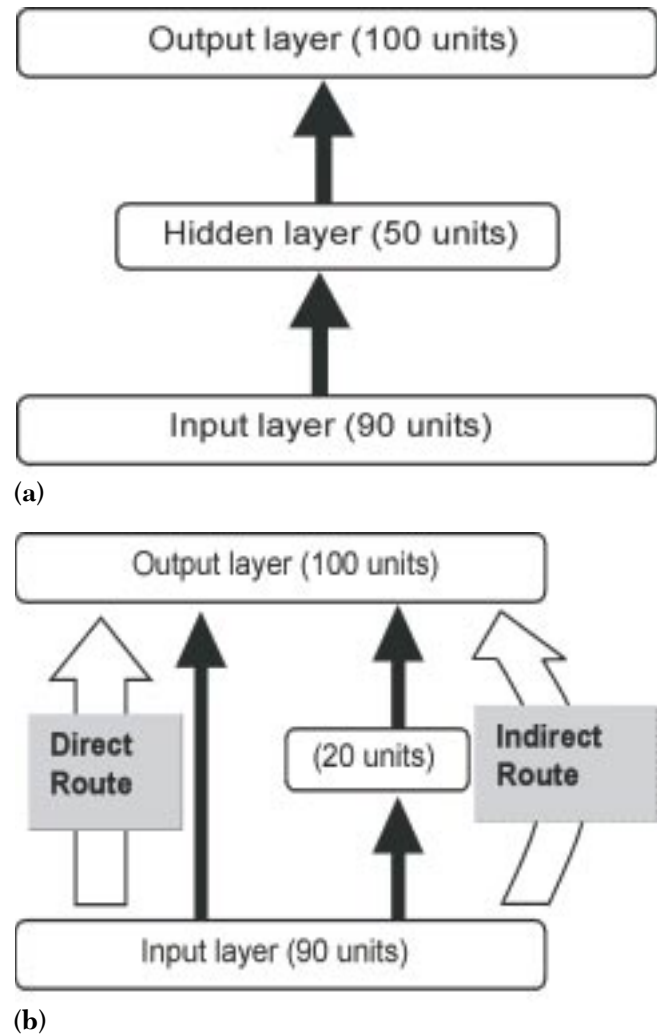


Figure 1. Architectures of the models used in Simulations One and Two. (a) Three-layer pattern associator. (b) Dual-route pattern associator.

whose past tense form bore no relation to its verb stem (e.g., *go-went*). The token frequency of this last type of exception verb had to be higher for the network to learn them successfully (see Plunkett & Marchman 1991), as is the case in real languages. As a result, this verb type experienced three times as much training as the other types. There were 410 regular verbs, and 20, 68, and 2, respectively, of each exception verb type.

A separate set of novel verbs was constructed to evaluate the generalisation performance of the network. These verbs could differ depending on their similarity to items in the training set. For simplicity, 410 novel verbs were used, each of which shared two phonemes with one of the regular verbs in the training set. Generalisation was evaluated depending on the proportion of these novel verbs, which were assigned the correct allomorph of the regular past tense rule.

Learning algorithm: The network was trained with the backpropagation learning algorithm, using cross-entropy between the output and target as the error signal (Hinton 1989). The learning rate was 0.01, and momentum was 0. The entire corpus was presented on each epoch, and pattern update was used. Networks were trained for 5,000 epochs.

Performance measure: A nearest-neighbour method was used to evaluate network performance, using a Euclidean distance metric. For each position in the output, the phoneme that the set of activation values most resembled was taken as the intended output for that position. If the resulting output string was the target output, it was marked as correct. Scores were therefore percentage correct for each verb type.

7.3. Implementation of damage

Lesioning: Weights were probabilistically set to zero throughout the network. The probability level determined the severity of the lesion. A probability of 0.3 would on average lesion 30% of the connection weights. Because of differences in sensitivity, probability levels of .01, .025, .05, .1, .2, and .3 were used for lesions applied at the end of training (with no retraining after damage), whereas probability levels of .5, .6, .7, .8, .9, and .95 were used for lesions applied at the beginning of training.

Noise: Noise was added to the activation levels of the units in the hidden layer, with a Gaussian distribution with mean zero and a standard deviation that determined the severity of the damage. Standard deviations of .025, .05, .0625, .075, .0875, .1, .2, .3, .4, .5, .6, and .7 were used. The baseline condition included no noise. Units had a maximum activation level of 1 and a minimum of 0, and noise could not take the activation state of a unit outside of these limits.

Discriminability: The activation of each processing unit in the hidden and output layers was determined by the following equation:

$$\text{Output} = \frac{1}{1 + e^{-(\text{netinput}/\text{Temperature})}}$$

where *net input* is the summed activation arriving at the unit including its bias, and where the *Temperature* parameter controls the steepness of this sigmoid function (see e.g., Hinton & Sejnowski 1986). High temperatures correspond to low discriminability, whereas low temperatures correspond to high discriminability. Values of 4 and .25 were used.

Replications. For the baseline model and for cases of damage prior to training, results were averaged over six runs of each network using different random seeds. Initial weights were randomised within the range ± 0.5 and pattern presentation during training was random without replacement. For cases of damage at the end of training, results were averaged over damage to each of the six baseline networks. For the addition of noise and the probabilistic lesioning of connection weights at the end of training, results were averaged over 10 repetitions of the damage for each of the six baseline networks. Graphs include standard error bars across the network replications as an indication of variability.

7.4. Results

The following graphs show performance on regular verbs, performance on the vowel-change exception verbs, and performance on generalisation of the regular rule. Results for the other two exception types were similar, and so are omitted. Baseline performance on the regular, exception, and rule pattern types was 100%, 100%, and 77%, respec-

tively. Figure 2a shows the effect of lesioning weights before training ("startstate" damage) and after training ("endstate" damage). For each pattern type, Figure 2 shows the relative performance of the startstate and endstate conditions for increasing levels of damage.

The results here indicate a similar pattern of impairment for both startstate and endstate damage on regular and exception verbs. Figure 2b demonstrates that in both cases, exception patterns suffer a greater impairment than regular patterns, echoing the surface dyslexia simulations. For novel items, startstate lesioning initially improves generalisation of the rule, whereas endstate lesioning is only deleterious. At higher levels of startstate lesioning, however, generalisation declines here, also. There are two major points to note from this simulation. First, very much greater damage is required in the startstate than in the endstate to produce an equivalent amount of behavioural impairment. Therefore a lesion of 2.5% of network connections in the endstate reduces performance on regular patterns to approximately 90%, whereas a lesion of 80% of the connections in the startstate is required to produce an equivalent deficit. Despite the fact that the same damage produces a similar behavioural impairment here, the training process creates a huge difference in *sensitivity* to damage between the startstate and endstate conditions. This is because the trained network is losing connections that have already stored specific knowledge, whereas the untrained network is reduced in its potential to learn and uses the remaining potential to acquire the domain as best it can.

The second finding is that the relationship among regular, exception, and rule performance in the startstate and endstate differs. For example, for a given level of performance on regular patterns in the damaged networks, the startstate network will show lower exception performance and higher generalisation performance. Despite broadly equivalent behavioural impairments, in detail the patterns of deficit are different in the acquired and developmental case.

Figure 3 illustrates the effect of adding noise to activation levels within the network, either once training is complete or throughout training. The results once more show a differential pattern of sensitivity, but now in the reverse direction to the lesioning condition. The network is much more sensitive to noise occurring in the startstate than it is to noise occurring in the endstate. In the endstate, the network has established its knowledge and, as a result of the nonlinear processing units, is able to tolerate noise in processing. As a result, performance has not yet reached floor when noise is added with a standard deviation (SD) of .7. However, when damage occurs in the startstate, the network is never provided with a reliable rendition of the knowledge it must learn. When noise is added with an SD of as little as .2, no learning is possible at all. Although the acquired and developmental phonological dyslexia models of Harm and Seidenberg (1999; 2001) are not directly comparable to each other, it is interesting to note that the acquired impairment was simulated by the addition of noise an order of magnitude greater than that used to simulate the developmental impairment, in line with the current findings. And a similar indirect comparison of Joanisse's (2000) model of SLI and Joanisse and Seidenberg's (1999) model of aphasia in the past tense domain indicates a comparable requirement for greater noise in the endstate than the startstate to produce an equivalent level of deficit.

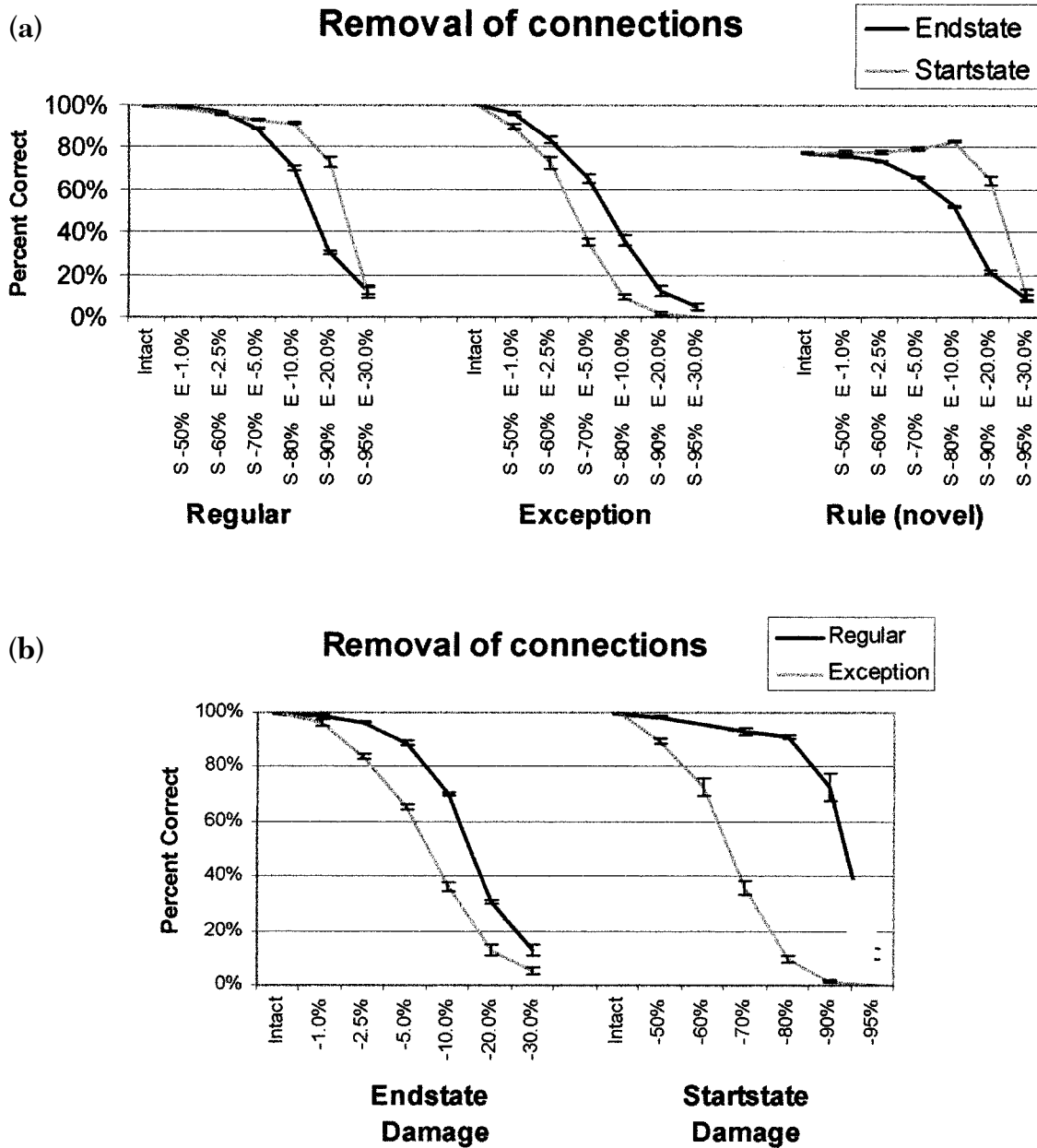


Figure 2. (a) The effect of removing connections from the network either prior to (startstate) or following (endstate) the training process for regular, exception, and novel patterns. The x-axis plots increasing levels of damage, with much greater startstate damage (S) required to produced an equivalent impairment than endstate (E). (b) Direct comparison of impairments on regular and exception patterns following startstate and endstate damage.

Noise added to the endstate led to a roughly uniform decrement across regular, exception, and novel patterns (e.g., from SD levels of .3 onwards). However, noise added to the startstate led to greater impairments to exception patterns than regular patterns. As an indication of this effect, when regular pattern performance was roughly comparable in the two conditions (startstate 90% with SD = .075, endstate 87% with SD = .25), exception patterns had fallen to 69% in the startstate against 79% in the endstate. The similarity of the mapping between regular patterns as well as their majority in the training set allows them to better overcome the addition of noise in training than the unique and minority exception patterns. In contrast, the ability of the network to deal with noise in the trained state

depends on the nonlinear functions within the hidden units – units that, broadly speaking, are shared by all patterns. Based on the relation of these different components of performance, once more one must conclude that the detailed pattern of impairments in the acquired and developmental cases was different.

In sum, the addition of noise could produce effects that were *uniform* and *global* in effect (across regulars, exceptions, and generalisation), as in the endstate, or that were *differential*, as in the startstate; but, most clearly, effects were much stronger when the damage occurred in the *startstate* than when it occurred in the endstate.

Figure 4 shows the effect of changing the discriminability of the processing units within the network. Reduced dis-

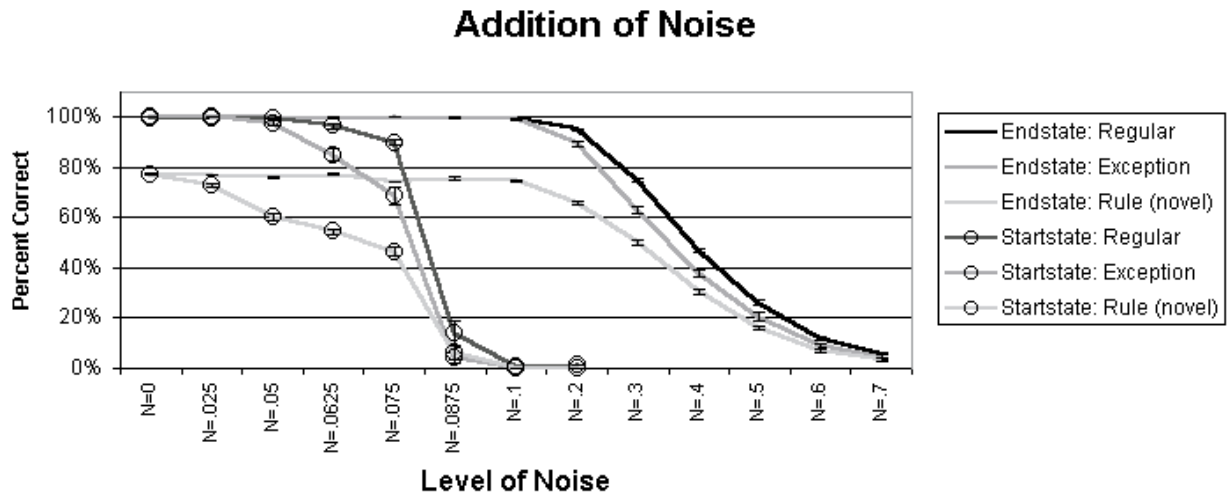


Figure 3. The effect of adding noise to activation levels within the network either prior to (startstate) or following (endstate) the training processing for regular, exception, and novel patterns.

criminability has little effect when applied either to startstate or endstate. The network can evidently compensate for it during training or tolerate this disruption at the end of training. An increase in discriminability also produces little effect when applied to the startstate – once more, the network can evidently compensate during training. However, if an increase in discriminability is applied to the endstate, the result is a marked and selective deficit in performance on exception patterns, dropping in performance from 100% to 30%. Meanwhile, regular patterns only suffer a minor dip in performance, and generalisation increases slightly. Evidently, exception patterns rely more on a certain level of discriminability in the processing units than do regular patterns. On the whole, this type of damage produces an impairment that is *selective* in the behaviour it impairs and only occurs when the damage is applied to the *endstate*.

7.5. Discussion

This simulation addressed two issues. First, does the process of development contribute to the pattern of deficits? Second, does the process of development produce patterns of deficits that are the same as those produced in acquired damage?

With regard to the first issue, direct comparison of the effects of identical damage at startstate and endstate demonstrated a complex relation between these two conditions, indicating a significant role for the process of training. Three different forms of damage produced three different possible relations. First, damage could produce a similar pattern of impairment for the startstate and endstate conditions, but the two conditions could vary in their sensitivity to the damage (removal of connections, addition of noise). Second,



Figure 4. The effect of altering unit discriminability within the network either prior to (startstate) or following (endstate) the training processing.

damage could produce impairments predominantly in the startstate (addition of noise) or predominantly in the endstate (increase in discriminability, removal of connections). Third, damage could produce impairments that were global (addition of noise, removal of connections) or selective (increase in discriminability). This complex relationship exists because the training process can sometimes play a crucial, compensatory role after damage occurs to the startstate of a learning system. However, it can play this role only to the extent that resources permit, and only to the extent that the representation of the domain remains reliable. To the extent that this is a valid model of cognitive development, the suggestion here is that the process of development will indeed contribute to patterns of deficits in a single system, but that the exact contribution will depend on the type of damage and the structure of the problem domain.⁸

With regard to the second issue, the results suggested that, at most, broad similarities were evident in the deficits caused by startstate and endstate damage, for example in the cases of lesioning and adding noise. However, in both cases, detailed examination revealed that the patterns of deficits were different – the behavioural impairments across related measures (regular, exception, and rule) did not line up. Again, to the extent that this is a valid model of cognitive development, the results do not support the idea that developmental and acquired deficits will produce precisely the same patterns in a single system.

It is instructive to see why this was the case. Take the example of lesioning the model. In the case of endstate deficits, the decline of regular and exception patterns was more closely tied because both pattern types shared a representational space that was being damaged. In the case of startstate deficits, the potential representational space was reduced, but the training process allowed the regular patterns to dominate the space that was available. The result was a system in which exception patterns were eventually squeezed out. These two impaired systems did not share a common final deficit because there is a distinction between a process of deleting parts of a representational space that is already occupied and the outcome of a process of occupying a representational space in which the initial size has been reduced.

However, the results by no means rule out the possibility that learning systems can be damaged in *different* ways prior to and following training, such that they exhibit identical endstate behavioural impairments. Generally, one must be very cautious about assuming identical causes in the case of identical outcomes. It is certainly the case that in connectionist models of developmental disorders, different forms of startstate damage can produce similar endstate behavioural impairments, as we saw in the case of phonological dyslexia and past tense formation in Williams syndrome. However, we have not yet unearthed any convincing examples in our own work or in the literature that startstate and endstate damage can separately produce identical endstate deficits in these networks. Time will tell on this point.

8. Simulation Two: Testing the assumption of Residual Normality in a simple connectionist learning system

8.1. Introduction

Many claims for RN relate to static adult models containing multiple independent, functionally specialised components.

These components are supposed to fail separately under both acquired and development damage. Claims about developmental damage, however, are quite inappropriately applied to such models, because they are not models of development (nor do they pretend to be). In this simulation, we address the issue of RN in models with specialised components that are the product of a learning process.

How do specialised processing components arise in the cognitive system? Most connectionist models of cognitive processes have focused on single domains – in effect, they have been models of components within a *modular* system (see discussion in Karmiloff-Smith 1992). Less work has examined how specialised components may actually emerge from an initially undifferentiated computational substrate. Evidence from the study of brain processes suggests that the neocortices of newborns are less structurally differentiated compared to those of adults, and that cognitive processes are less localised in this early substrate (e.g., Johnson 1999). However, the key question regarding specialised structures has been whether their emergence during development reflects the unfolding of a maturational blueprint, whether their emergence depends entirely on experience, or whether it reflects a gradual process of modularisation that lies somewhere between these two extremes (Elman et al. 1996; Karmiloff-Smith 1992).

In a recent review, Jacobs (1999) argues that the evidence points to the experience-sensitive theory of specialisation. He discusses three computational approaches that have sought to model the experience-dependent emergence of structure. In the first approach, called *mixture of experts*, the initial computational system is assumed to be computationally heterogeneous. There are components that, while not dedicated to processing any particular content, have different computational properties. These components compete to perform the computations corresponding to a new cognitive domain. The component whose computational properties best fit the demands of the domain, known as a *structure-function correspondence*, will win the competition and come to specialise in processing that domain in the future (see Jacobs 1997; Jacobs et al. 1991). In the second approach, called *neural selectionism* or *parcellation*, the initial computational system has a surplus of connections. However, during learning, many of these connections are weeded out, whereas others are stabilised depending on usage. In addition, a locality constraint favours the stabilisation of connections between nearby processing units. The result is that nearby units communicate with each other and come to perform the same functions, whereas those far apart do not communicate and come to specialise in different functions (Jacobs & Jordan 1992; Johnson & Karmiloff-Smith 1992; see Plaut 2002, for a recent application to a cognitive model of naming and gesturing). In the final approach, called the *wave of plasticity*, the initial computational system experiences differential responsiveness to learning, both spatially and temporally. Conceived of as a sheet of computational units, plasticity is reduced over time with one side of the sheet losing its plasticity earlier than the other. The result is that the later maturing units can use the functions computed by earlier maturing units as input, and derive more complex and abstract computational functions from them – in essence, the later maturing units specialise on the more abstract or high-level aspects of the problem domain (Shrager & Johnson 1996).

Let us assume that the outcome of normal development

is a set of specialised components in the endstate, which can be revealed by adult neuropsychology. For our purposes, the relevant question is, if the developmental process itself is pushed off course in a developmental disorder, could this also alter the nature of the specialised structures that are the outcome of development? Little computational work has explored this question (see Oliver et al. 2000 for some preliminary work). It seems possible that specialisation could be disrupted in any of the above computational approaches. An alteration in the initial set of computational primitives or in the competition process could disrupt specialisation in the mixture-of-experts approach. An alteration in the method or timing of pruning long connections could disrupt specialisation in the neural selectionism approach. An alteration in the timing of plasticity changes could disrupt specialisation in the wave of plasticity approach.

Decisive evidence has yet to be put forward demonstrating radical differences in specialisation in developmental disorders, although some hints have been made in this direction. For example, Karmiloff-Smith (1998) has speculated that the cognitive processes of individuals with Down syndrome may be characterised by insufficient specialisation, perhaps because of a failure to prune long connections during development. However, conceptually, it is not yet clear to what extent one could compromise the emergence of specialised cognitive structures in a disordered state and still produce a viable cognitive system.

To extend the static adult damage model to developmental disorders is to make a more precise claim, however, that damage may be highly selective and thwart the development of a single specialised module. The question now becomes, if specialisation is not predetermined, under what conditions will the rest of the system develop normally despite this early selective damage? Fortunately, some existing models of reading and past tense formation allow us to explore this question. In these models, *structure-function correspondences* have been used to generate emergent specialisation in connectionist learning systems with multiple processing routes. Such networks include two processing routes in an initially content-free network.⁹ The routes have different computational properties, and these properties line up respectively with the computational requirements of learning regular and exception patterns (see sect. 8.2). The result in both a reading model (Zorzi et al. 1998) and a past tense model (Plunkett et al. 2001) was partial specialisation of the two routes to processing regular and exception patterns (see Westermann 1998 for a related constructivist approach). We sought to evaluate the assumption of RN using this dual-route model, and in particular to answer the following question: Does disruption to one route prior to training alter the function that the initially intact route takes on at the end of training?

8.2. Simulation details

Architecture. The architecture of the dual route network is shown in Figure 1b. The feedforward network included an input layer and an output layer connected via two processing routes. The *Direct* processing route comprised a set of connections linking the input and output layer. The *Indirect* processing route connected these two layers via an intermediate layer of 20 hidden units.

Structure-function correspondences. When exception

patterns must be learned in the face of a majority of regular patterns, additional computational resources are necessary. Specifically, although a two-layer network can learn the mappings for a set of purely regular patterns, hidden units are necessary to mark out the inconsistency of the exception patterns, typically involving the use of the three-layer architecture. These effects are not all or nothing. A two-layer network can tolerate a small proportion of exception patterns in the training set; exception patterns themselves can be more or less inconsistent with the regular patterns and therefore more or less demanding of hidden units to mark out their inconsistency. (In computational terms, the role of hidden units in overcoming the inconsistency between regular and exception patterns is a question of *linear inseparability* – see Elman et al. 1996, Ch. 2 for an introductory discussion.) Furthermore, the disadvantageous effect of inconsistency can be mitigated by increasing the frequency of exception patterns in the training set. Broadly, then, in a network combining a two-layer architecture and a three-layer architecture in separate routes, the two-layer route will be best fitted to learn the regular patterns, and the three-layer route will be required to learn the exceptions, more so with the greater the inconsistency of the exception patterns with the regular patterns. Structure-function correspondences can drive specialisation in error correction networks with multiple routes, because there is competition between each route to reduce the disparity between output activations and the training target. If one route succeeds in reducing the disparity, no error signal is left to change the weight strengths in the other route(s).

Our training set includes three types of exception pattern. Those based on the No Change past tense paradigm (*hit-hit*) are the least inconsistent with regular verbs, because as with regulars, the verb stem is reproduced, whereas the affix is omitted. We term these exceptions EP1. Exception patterns based on the Vowel Change paradigm (*hide-hid*) are more inconsistent, because in addition to the omission of an affix, the central vowel of the verb stem must be transformed. These exceptions we term EP2. Finally, exception patterns based on the Arbitrary past tense paradigm (*go-went*) are the most inconsistent with the regular patterns, because the verb stem must be entirely transformed and the affix omitted. We might expect these patterns to be most dependent on the hidden units of the Indirect processing route. However, in the past tense domain, it is argued that arbitrary past tenses can only be retained in English if they are of very high token frequency. In the current training set, arbitrary mappings were presented three times as often as other forms. We term this most inconsistent exception type EP3f, to reflect the fact that high frequency may modulate patterns of specialisation.

Training set. The training set was identical to that in Simulation One, except that instead of two arbitrary patterns in the training set, there were 10 such patterns. This permitted a more sensitive evaluation of performance on this pattern type.

Learning algorithm. The learning algorithm was the same as that in Simulation One.

Residual Normality condition: In addition to the normal training scheme, for comparison the model was also trained under a Residual Normality condition. This condition assumed “guided specialisation” (see sect. 9). Here, the Direct route was trained on Regulars alone, and the Indirect route trained on Irregulars alone. Guided specialisation in a mul-

ticomponent model requires an external control system to coordinate the subsequent function of the trained components (see, e.g., the “Blocking” device in Pinker’s [1991] dual-mechanism past tense model). For simplicity, the control system was assumed in the RN condition. Routes were trained and tested independently.

Performance measure. Performance was measured using a nearest-neighbour calculation based on output activations, and scores marked as percentage correct. Specialisation of a particular pattern type to a particular route was evaluated by selectively lesioning the network at the relevant point in training. The Direct route and the Indirect route were separately given a probabilistic lesion of their weights with $p = .5$ (50% of all weights in that route). If damage to the Direct route caused more impairment on a pattern type than damage to the Indirect route, it was assumed that the function for this pattern type was specialised to the Direct route, and vice versa. Note that removal of 50% of the connections may not have equivalent effects on the Direct and Indirect routes, because the latter has two layers of weights. However, we were concerned here with differential effects between pattern types rather than routes. We did check for interactions, that is, the possibility that pattern types might show differential sensitivity to damage in each route whereby, for example, a pattern type may appear to be specialised to one route at 50% damage but the other route at 10% damage. This possibility was explored by carrying out endstate lesions with probabilities of .025, .05, .1, .2, .25, .5, and .75. Although, overall, the Indirect route showed greater sensitivity to damage than the Direct route, there was very little modulation of *relative* specialisation levels of the pattern types across damage levels. However, the *absolute* level of specialisation was affected by whether the level of damage was so great or small that it produced floor or ceiling effects in regular or exception performance. A level of 50% lesioning was used to assess specialisation because this was in the midrange of sensitivity for both regular and exception patterns.

Implementation of pretraining damage. This simulation sought to explore the implication of specific damage to either of the two routes prior to training. This was achieved by removing different proportions of the weights in each route. Probability levels of .6, .75, .9 and 1 (removal of entire route) were applied. In addition, this level of damage was performed on both routes simultaneously, as a control. After initial damage, training proceeded as normal, and levels of specialisation were then assessed in the endstate.

Replications. Results were averaged across six networks with different initial random seeds for each level of startstate damage. Probabilistic lesions were carried out 10 times and the results averaged. Graphs include standard error bars as an indication of variability across the six network replications.

8.3. Results

Figure 5c represents an index of specialisation, where positive values represent specialisation to the Direct route, and negative values represent specialisation to the Indirect route. This index corresponds to the differential impairment caused by lesioning to a single route. The first point to note is that specialisation is only partial. Using this measurement technique, in Figure 5c damage to the Direct route can cause a maximum decrement in performance only 53% greater than

damage to the Indirect route. Damage to the Indirect route can cause a maximum decrement 29% greater than damage to Direct route. Secondly, most patterns show a shift towards using the Indirect route later in training.

To understand this latter point, it is important to realise that the two routes of the network do not just differ in their computational properties, but also in their plasticity. By virtue of the learning algorithm, weights from the input layer to the hidden units change more slowly than the weights directly connecting input and output layers. In effect, this network comprises one relatively more “stupid” but fast-changing route, and one relatively more “clever” but slow-changing route. Early on in training, successful performance is largely a result of the Direct route, and this performance is best on regular patterns, generalisation of the rule, and the EP1 exceptions – those that are least inconsistent with the regular patterns. Subsequently, the slower-changing Indirect route increasingly contributes to performance, especially for the exception patterns. However, by the end of training, both regular patterns and generalisation of the rule rely more on the Direct route, whereas all exception patterns rely more on the Indirect route. As expected, the more inconsistent EP2 patterns turn out to rely more on the Indirect route than the EP1 patterns. However, the higher frequency of the EP3f patterns means that their greater inconsistency does not lead to Indirect route specialisation any more than that shown by EP2; higher frequency allows these patterns to recruit more processing from the Direct route in the face of the dominance of regular patterns.

In summary, this network shows emergent specialisation of function of different pattern types to separate processing structures. Although this specialisation is not complete, our concern here was to establish a baseline level to explore the effect of initial, route-specific damage to this process. Figure 6 shows (from left to right) the result of startstate lesions to the whole network, to the Indirect route in isolation, and to the Direct route in isolation. Figure 6a illustrates the effect of these startstate lesions on endstate performance, while Figure 6b shows their effect on the profile of specialisation in the endstate network.

Figure 6a demonstrates that when the entire Indirect route is removed, endstate performance on regular patterns and generalisation is only mildly impaired. However, performance on the exception patterns reveals a marked decrement, particularly for EP1 and EP2, which are not protected by increased frequency. When the entire Direct route is removed, regular patterns are impaired to a greater extent. Exception patterns are also impaired, but less than when the Indirect route is lost. Removal of the Direct route, however, produces a marked deficit in rule generalisation. In isolation, then, each of the two routes will attempt to acquire both regular and exception patterns, but each does so less efficiently than in the dual-route system. The routes in isolation both produce decrements in exception performance, but relatively speaking, the Indirect route is less able (but not unable) to learn the regular rule, and the Direct route is less able (but again, not unable) to learn the exception patterns.

These results mark the maximum compensation that is available to the network. Figure 6b demonstrates the specialisation when there is residual processing capacity in the damaged route.¹⁰ Here the data are unambiguous. When there is initial damage to the Indirect route, specialisation of

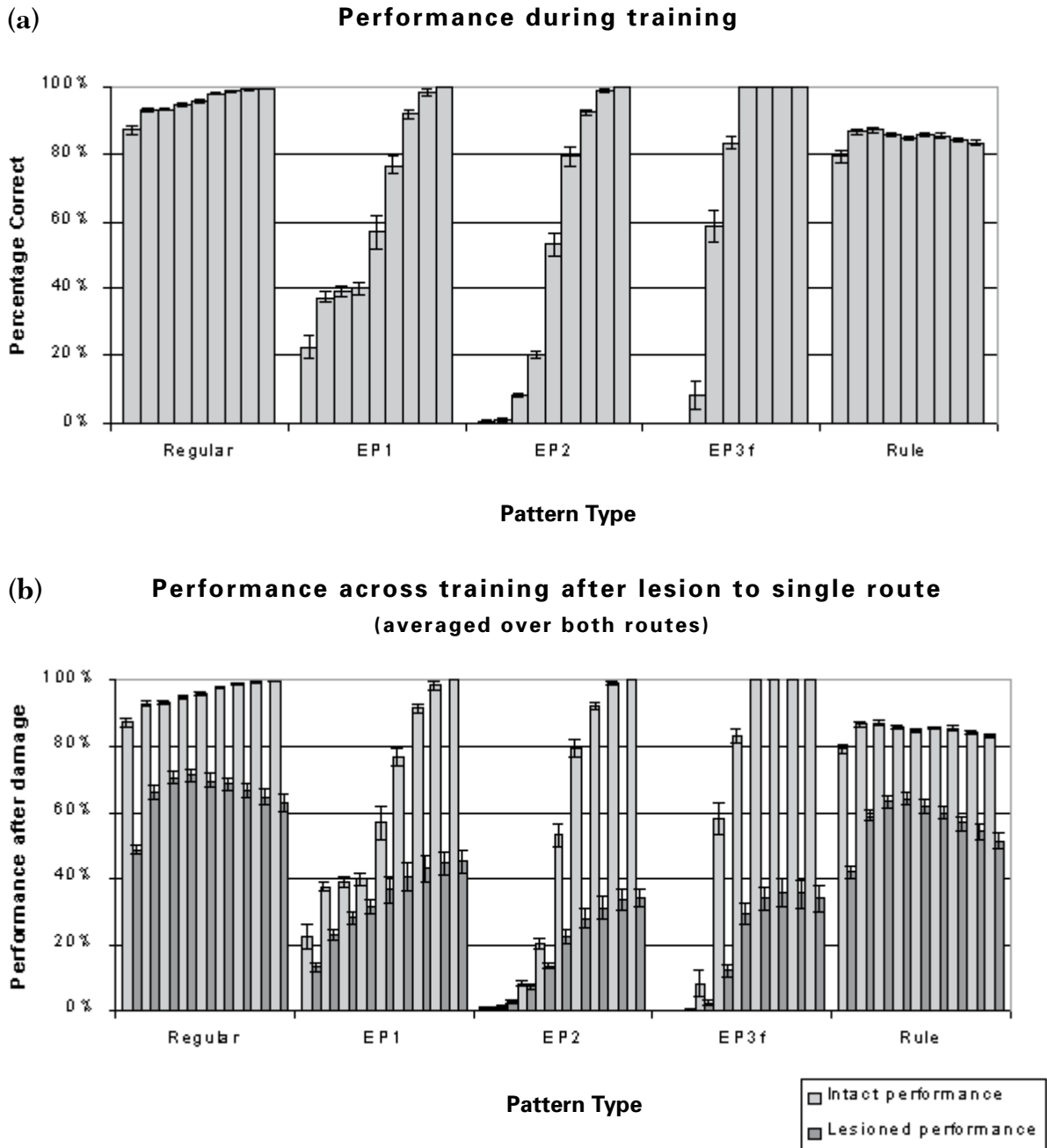


Figure 5. (a) Performance on each pattern type during training in the dual-route network, at 10, 25, 50, 100, 250, 500, 1,000, 2,000, and 5,000 epochs of training. See Simulation Details (Sec. 8.2) for description of each pattern type. (b) Performance after a 50% lesion to each route, averaged over the two routes, carried out at each point in training. The 50% lesion was used to measure specialisation of function to each route. (c) Specialisation of function for each pattern type during training, indexed by the differential impairment caused by damaging each route in isolation. Positive values indicate specialisation to the Direct route; negative values, to the Indirect route. Error bars show standard errors across network replications.

function increasingly moves over to the intact Direct route. When there is initial damage to the Direct route, specialisation of function increasingly moves over to the intact Indirect route. A crucial lesson is demonstrated by this simulation: The assumption of Residual Normality does not hold in this learning system. Damage one route and the other route will *not* develop normally. It will compensate, and take on part of the function of the damaged route, at the cost of poorer performance across all pattern types. It cannot be taken for granted that every learning system will show RN.

This simulation is interesting in two other respects. In the condition where both routes experienced initial damage, the overall outcome was reduced specialisation. A system with uniformly reduced computational resources does not have the luxury of allocating functions to different components. The effect of resources on specialisation is a point to which we will return shortly. Second, the 50% lesion used to measure specialisation produced greater deficits in systems that had experienced prodstate damage to either or both routes than those that had not. Importantly, even if

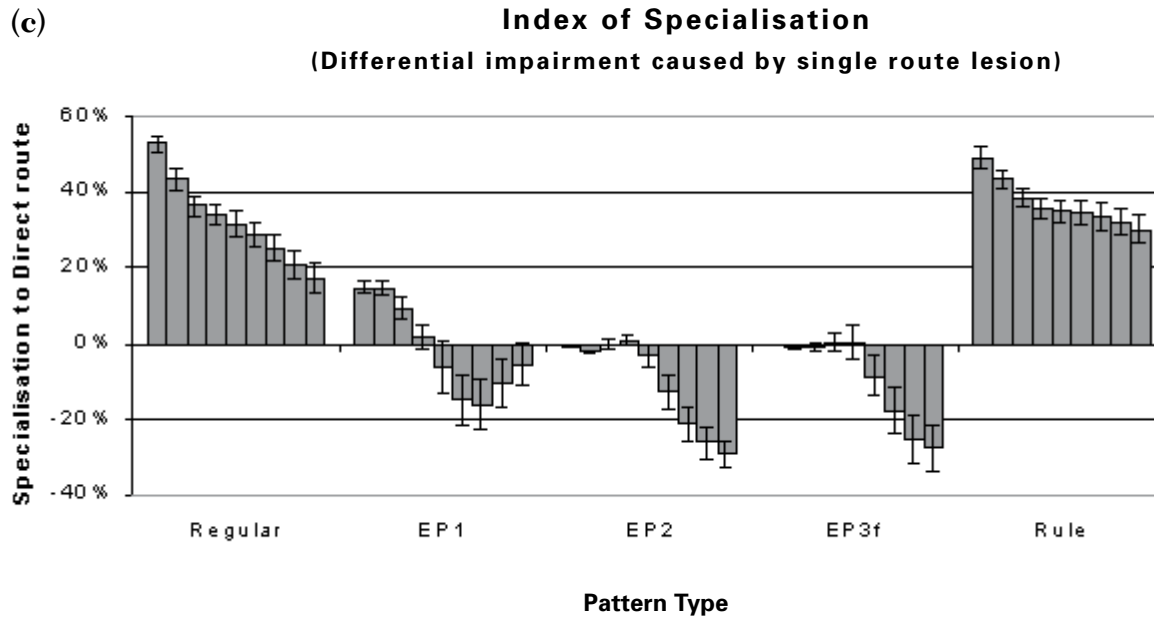


Figure 5. (continued)

systems that experience early damage achieve reasonable endstate performance, they remain more vulnerable to subsequent disruption.¹¹

8.4. Discussion

Given the compensatory characteristics of the training process evident in Simulation One, it is perhaps unsurprising to find similar compensation here. When the two routes of a dual-route network show endstate specialisation, startstate damage to either route results in compensation by the intact route and poorer performance overall. If RN had held in this model, damaging a route in the startstate would have led to the (endstate) loss of the same function for which that route was responsible in the endstate of the normal unimpaired model, with normal function in the initially undamaged route. This is precisely the claim that is made (for example) in the case of phonological dyslexia, where a developmental problem in the GPC route of the reading system is assumed to lead to the same pattern of behavioural impairments as damage to the GPC route in the adult state (Coltheart et al. 2001), with the lexical routes functioning normally in both cases.

The feedforward network presented in Simulation Two is similar to connectionist systems that have been used to successfully capture a wide range of developmental phenomena. These systems plainly do not demonstrate RN. Nevertheless, as we saw in the Introduction, RN is frequently postulated (albeit implicitly) in many explanations of developmental disorders. Perhaps, then, despite their success, current connectionist models are not the right sort of learning system to explain how the structure of the adult cognitive system comes about. Can other sorts of learning system show the emergence of specialised components while exhibiting RN after initial damage?

To date, artificial neural networks are the computational systems that have been most widely applied to the study of cognitive development. However, it is certainly possible

that other approaches will come to the fore in the future, such as decision-tree learning, Bayesian methods, production systems, reinforcement learning, instance-based learning, genetic algorithms, or indeed other types of artificial neural networks. In the meantime, from the perspective of developmental disorders, it is vital to stipulate what additional constraints any such learning systems would need to incorporate to achieve RN. We will discuss five (somewhat overlapping) notions: (1) stronger structure-function correspondences, (2) stronger competition, (3) early commitment, (4) guided specialisation, and (5) restrictions on computational resources. Note that all of these notions are based on the assumption that endstate cognitive structure is experience-sensitive. RN can of course be *stipulated*, as per accounts that propose that modular structure in the cognitive system is prespecified and that if components develop, they do so independently. (Those who stipulate innate modularity would then need to justify this claim with evidence from developmental cognitive neuroscience – evidence that we currently believe to be wanting.)

9. Ways to achieve Residual Normality in systems with emergent modularisation

9.1. Stronger structure-function correspondences

Each of the routes in the dual-route network was able to show a fair degree of compensation for the functions of the other, suggesting that the correspondences between the functions of the two routes and the regular/exception structure of the learning problem were partially overlapping. One way to assure RN would be to have much stronger structure-function correspondences, whereby the computational properties of each route were entirely inappropriate for learning the patterns on which the other route specialised.

We might illustrate this idea by stepping outside our two-example domains for a moment and considering a connectionist model of the development of object-oriented behav-

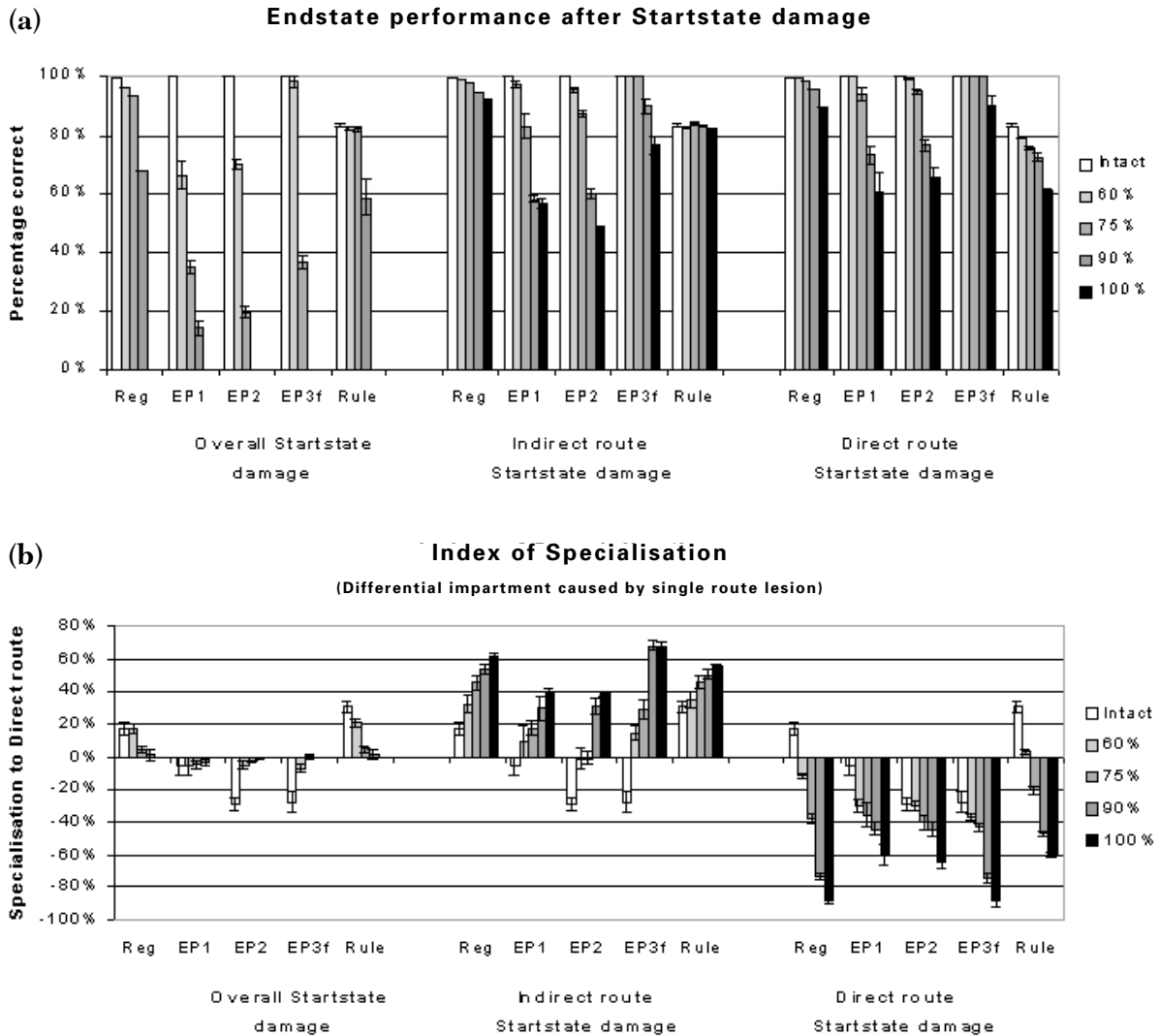


Figure 6. (a) Performance at the end of training following initial damage to both routes, to the Indirect route only, or to the Direct route only. (b) Patterns of endstate specialisation for each pattern type after initial damage. Error bars show standard errors across network replications.

our proposed by Mareschal et al. (1999). This model used an input retina to represent the trajectories of various sorts of moving objects. The system had to learn to reach for some of the objects it saw but not for others. To achieve this task, it was given two routes, one that processed spatiotemporal information about the position of each object (the “where” channel), the other that processed featural information about the identity of each object, such as its colour and shape (the “what” channel). A final layer of the system combined the two processing routes to achieve reaching behaviour. The relevance of this model is that the computational properties of each processing route were very different. The “what” channel utilised competitive learning to achieve translation-invariant feature recognition across the entire retina. The “where” channel employed recurrent circuits to encode time-varying information about position. If either of these channels were to be damaged prior to the training process, the remaining route simply would not have the appropriate computational primitives to compensate for the function of the damaged route: Translation-invariant featural information contains no clues to location, and spatial

trajectory information contains no clues about object identity.¹² The result would be a system with RN, where the remaining initially intact route would develop the only skill it had the capacity to learn and nothing else.¹³

9.2. Stronger competition

In the mixture-of-experts approach to specialisation, separate components compete to represent a new domain. The component best able to represent the domain is given “sole rights” to it, and the other components are inhibited. In the dual-route network, however, both routes worked in harness to learn the appropriate mappings – neither was prevented from adjusting its weights to improve performance on any pattern. The result was to encourage cooperation between the routes. Much stronger competitive processes might permit a component to claim sole rights only to patterns that best suited its computational properties, and might inhibit it from acquiring patterns outside of that set. If the same component were to win the competition in both normal and atypical development (a big “if”), then the component would exhibit RN and no compensation.

9.3. Early commitment

If plasticity reduces rapidly over time, then early commitment may contribute to RN. In this scenario, early commitment of separate components to functions must occur before the developmental disruption takes place. This can be either because the damage has not yet occurred, or because the underlying damage does not make itself apparent because it is not relevant for the cognitive processes appropriate for the current stage of development (see discussion in Thomas 2003). Early and irreversible (yet experience-dependent) specialisation could contribute to a modular system where RN holds (see, e.g., Miller & Erwin 2001), provided an account exists of the required delay in the emergence of the developmental disruption.

9.4. Guided specialisation

In the dual-route network, prior analysis of the computational properties of two- and three-layer networks suggested that the Direct route would be better suited for learning regular patterns and the Indirect route would be better suited to learning the exception patterns. In principle, we could have determined to label each pattern as Regular or Exception in advance, and then only allowed the Direct route to alter its weights in response to Regular patterns and the Indirect route to alter its weights in response to Exception patterns. Unsurprisingly, the result of this form of guided training would be independent specialisation. Clearly if one route were damaged prior to training without any change to the advanced labelling system, this route would fail to learn the patterns assigned to it. The other route would be unaffected and would hence show RN.

In such a case, we are of course left with the burning question of where the advanced labelling information comes from. Some would argue it is innate. Alternatively, the labelling information could be the product of an earlier phase of learning, in which an analysis of the target domain identified the presence of regular and exception patterns *before* the dual-route system was engaged. It should be evident here that if one chooses to appeal to guided specialisation to support RN, then claims about the availability of a priori knowledge need serious substantiation.

9.5. Restrictions on computational resources

We saw in the results of Simulation Two that specialisation was eliminated under severe resource limitation. Resource limitations may also have implications for limits on compensation and, indirectly, for RN. This idea can be illustrated with an example from the past tense domain and claims made for the (computationally unimplemented) traditional model. The traditional model comprises two mechanisms, one employing rule-based representations and nominally responsible for learning regular inflections, and the other employing an associative memory and nominally responsible for learning exception inflections. If these two mechanisms are to specialise appropriately on their respective inflections (in the absence of external guidance; see Pinker 1999 for discussion of a possible mechanism for guided specialisation in this theoretical model), it is important that neither mechanism be too powerful. Given sufficient “rule” resources, for example, all past tenses could be learned in terms of a large set of rules (see, e.g., Ling & Marinov 1993; Taatgen & Anderson 2002). In contrast,

given sufficient “associative memory” resources, all past tenses could be learned as specific instances. Generalisation to novel exemplars could be achieved in either case by an analogy-based strategy, such as similarity-to-the-nearest-known exemplar. Given two overly powerful mechanisms, the result would be duplicated processing systems, each able to perform the whole task. To achieve specialisation, therefore, one needs to restrict the resources of each mechanism. For example, the English past tense has one rule and about 150 exceptions. If the relevant mechanisms were restricted to these limits, then they would show little compensation in the event of damage to the other mechanism.

The case of SLI illustrates the point. Children with SLI show low levels of inflection on both regular and exception verbs, and poor generalisation of the regular rule to novel strings. Appealing to the traditional two-mechanism model of past tense formation, Ullman and Gopnik (1999), Pinker (1999), and van der Lely and Ullman (2001) have all argued that there is a startstate deficit to the mechanism intended to learn the regular rule. All that remains is the exception mechanism, which learns the past tenses of some exception and some regular verbs, and (presumably by analogy) can struggle to offer a few correct generalisations to novel verbs. The implicit claim here is that there is a *particular limit in computational resources* in the exception mechanism that prevents it from learning more than a handful of regular past tenses by way of compensation. Specifically, the exception mechanism must be able to learn the couple of hundred exceptions to explain its performance in normal development. But, to explain the lack of compensation in SLI, it must be unable to learn a couple of thousand high-frequency regular past tenses that would be sufficient to get by in everyday language use and therefore mask the regular-mechanism deficit. The explanation of the endstate impairment crucially relies on this precise memory stipulation which constrains compensation. (It is interesting that no empirical support is offered for such a memory limitation in the preceding accounts of SLI.)

In short, resource limitations may be a necessary component of RN but not a sufficient one.

10. Residual Normality and the inference from behaviour to structure

In section 4, we identified two developmental conditions where behavioural similarities between acquired and developmental disorders should not lead to the inference that they reflect selective impairments to the same components of a static adult model. One condition was when features of the problem domain determine the pattern of breakdown rather than features of the processing structure. In the simulations, this situation is reflected in the greater vulnerability of exception patterns, irrespective of damage type. However, closer inspection revealed different patterns of deficit to regulars, exceptions, and generalisation for different damage types, implying an effect of residual processing structures on performance.

The second condition was when similar patterns of behaviour in atypical development are generated by a different underlying structure of specialised components. In section 9, we have argued that the emergent pattern of functional specialisation depends on computational constraints operating during cognitive development. However,

Jackson and Coltheart (2001) maintain that developmental disorders in their endstate are potentially no different from acquired disorders and can be used with reference to the normal adult model quite independently of the nature of the developmental process that produced them. One reading of their position is that, at a given moment in time, inferences can be made from behavioural impairments directly to underlying structure, irrespective of the developmental processes that produced the system.

Computational modelling allows us to explore this claim, because we can simultaneously generate patterns of behavioural deficits while knowing the underlying cause and the background developmental account in each case. Figure 7 illustrates a behavioural impairment generated from the dual-route model following selective damage to one of its routes. Two versions of the model were damaged, the one we have already studied in which RN does not hold, and a second version in which RN *does* hold (in this case, RN was achieved by guided specialisation of regulars and exceptions to the two routes. See simulation details in section 8.2). Figure 7 shows that both versions of the model can generate similar behavioural impairments after startstate and after endstate damage, (although in the absence of RN, regular performance is not quite at ceiling). However, as summarised in Table 1, the inferences that one can make from intact behaviour to intact underlying process, or from impaired behaviour to impaired underlying process, crucially depend on the developmental constraints of the system. With regard to *developmental deficits*, where RN holds, intact behaviour implies intact underlying process and a dissociation of independent structures. Where RN does not hold, intact behaviour implies atypical underlying process, in a system that has experienced compensation during training. One has to know the developmental story with regard to *acquired deficits*, however, the developmental constraints are not key in inferring underlying structure, because specialised processing components have formed prior to damage.

This simulation illustrates that the second condition identified in section 4 is a quite viable one. A pattern of intact and impaired behaviour can be produced by atypical underlying processes in the developmental case. It supports the idea that in developmental disorders, underlying structure cannot be inferred from behavioural deficits without explicit reference to the developmental constraints under which processing structures have emerged. In contrast, the logic of acquired deficits in the endstate is agnostic as to whether specialisation was prespecified or emergent.

11. Conclusion

In developmental cognitive neuropsychology, the hypothesis of Residual Normality is widely deployed, explicitly or implicitly, in explanations of developmental deficits, but with little supporting evidence. The motivation is clear – it allows developmental disorders to be interpreted according to static adult models that themselves have no developmental component. We have argued that this tendency to focus on specific deficits and take RN for granted has had serious consequences for the type of data collected in characterising developmental disorders. Researchers have been too prone to assume that superficially normal behaviour (e.g., in standardised tests) corresponds to normal underlying processes, irrespective of the developmental plausibility of RN in each cognitive domain. We believe this tendency has also impeded the discipline in building links to developmental cognitive neuroscience and developmental neurobiology, because the use of RN with static models stipulates rather than explores the developmental origins of deficits.

There may be no universal causal link between developmental and acquired disorders. Rather, RN may hold for some cognitive domains and not for others. Compensation and plasticity mitigate against it on a fine-scale decomposition of cognitive domains, but RN may obtain across some broader functional distinctions, such as ventral versus dorsal

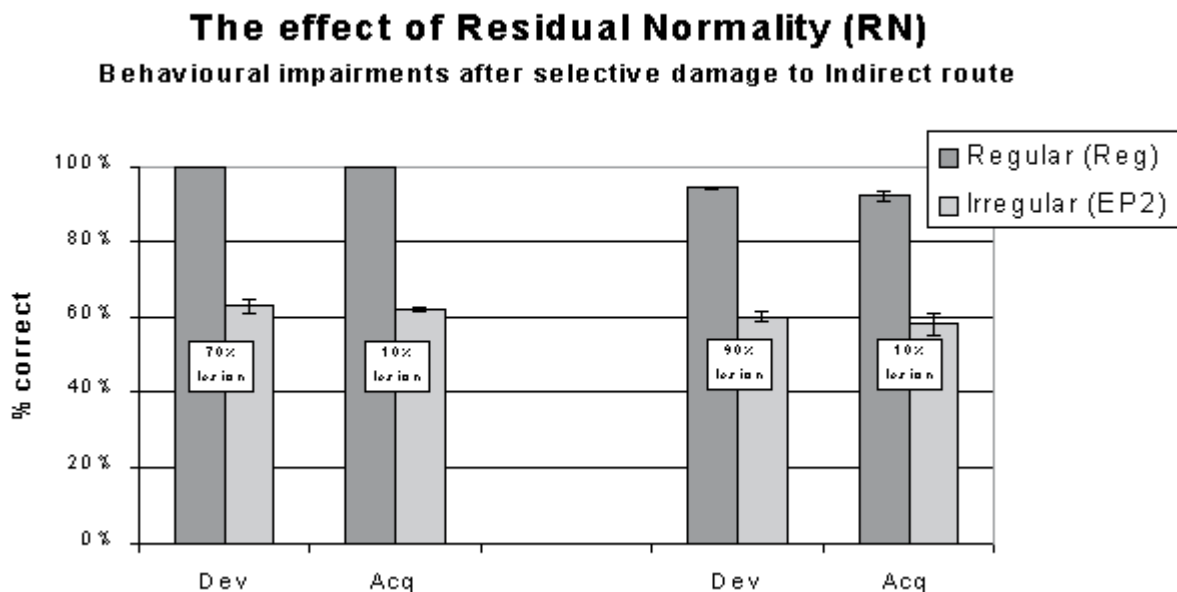


Figure 7. Selective behavioural impairment caused by startstate (Developmental) and endstate (Acquired) lesion to the Indirect route of the Dual-Route network, in the cases when RN operates and when it does not (see Simulation Details, sect. 8.2). From left to right, the lesions to the Indirect routes were 70%, 10%, 90%, and 10% of connections, respectively.

Table 1. *Inferences from behaviour to structure. In two different models, selective damage to the Indirect route of the dual route system causes a similar selective behavioural impairment to Irregular patterns (EP2), with largely intact performance on Regular patterns (Reg), in both developmental and acquired cases. In one model, Residual Normality (RN) holds; in the other it does not. The correct inference from behaviour to underlying process depends on the developmental assumption (of whether RN holds)*

	What is producing normal (intact) behaviour? [Reg]		Does the dissociation imply independent underlying processes?	
	RN operates	RN does not operate	RN operates	RN does not operate
Acquired deficit	Normal process in Direct route	Mix of Normal process in Direct route and residual process in damaged Indirect route	Yes	Yes *
Developmental deficit	Normal process in Direct route	Atypical processes in both Direct and residual Indirect route after development	Yes	No

	What is producing impaired behaviour? [EP2]	
	RN operates	RN does not operate
Acquired deficit	Impaired process in residual damaged Indirect route	Predominantly Impaired process in residual damaged Indirect route
Developmental deficit	Impaired process in residual damaged Indirect route	Atypical processes in both Direct and residual Indirect route after development

*Partial independence (see sect. 8.3)

systems, anterior versus posterior systems, or cortical versus subcortical systems. The plausibility of RN across various systems and domains is a topic in need of urgent investigation.

We conclude that RN cannot be taken for granted, but must be argued for on the basis of developmental evidence. It is essential that developmental *disorders* be compared against developmental *models*, and that the process of ontogenetic development itself be taken seriously.

ACKNOWLEDGMENTS

Many thanks to Max Coltheart for helpful discussions on the arguments within this article. We thank Denis Mareschal and Birkbeck College, and Kim Plunkett and the McDonnell-Pew Centre for Cognitive Neuroscience for the use of time on their computers. Thanks also to Michael Rutter, Denis Mareschal, Marc Joanisse, Lorraine Tyler, and Mike Harm for helpful comments. This work was funded by MRC Project Grant No. G9809880 and MRC Programme Grant No. G9715642 to Annette Karmiloff-Smith.

NOTES

1. With a much larger sample size, Thomas et al. (2001) subsequently found that this deficit appeared to be an artefact of delayed language development in Williams syndrome. Their results suggested that the main difference between the Williams syndrome group and controls was a reluctance to inflect novel forms. However, for current purposes, Clahsen and Almazan's (1998) claim serves to illustrate the explanatory framework within which those authors were operating.

2. Where we use the term *compensation* in this article, we do not mean that alterations in response to damage elsewhere in the

system must necessarily produce a beneficial outcome. Undamaged (i.e., previously normally developing) components may have their development disrupted through attempts to compensate for the impairment, such that they function more poorly on the task that they normally acquire. For example, in the simulations presented in section 8.3, a component that normally specialises in learning regular patterns achieves this function less efficiently when, as a consequence of damage elsewhere in the system, it also has to learn exception patterns. Similarly, a component that normally specialises in exception patterns achieves this function less efficiently when, as a consequence of damage elsewhere in the system, it also has to learn regular patterns.

3. Or indeed more routes, less routes, or other routes.

4. Because of space restrictions, our discussion focuses on fixed architecture models. We suspect similar arguments will hold for constructivist models that change their architecture as a function of learning. However, the general interaction of early damage and constructivist processes of development requires further investigation (see Thomas & Karmiloff-Smith, in press, for discussion of constructivist models and developmental disorders in the specific case of past tense formation).

5. In principle, models of atypical development could employ damage that occurs *during* training. For example, in typical development, a computational constraint might change as a function of developmental time. Such constraints might include the rate at which connections are pruned, or the rate at which connection strengths are altered in response to error signals. Although constructivist neural networks add hidden units as part of the learning algorithm, thus far, limited modelling work has explored how typical development may *rely* on continuously changing computational constraints, let alone how atypical development may arise from dysfunctions in such changes.

6. The developmental process is equated here, as in most connectionist models of development, to the training process – see Thomas and Karmiloff-Smith (2003) for discussion of this assumption.

7. A debate continues as to (1) whether patients can indeed show decrements in past tense performance whereby regular inflections for existing words are much poorer than exception inflections (see, e.g., Bird et al. 2002; Ullman et al. 1997); and (2), whether deficits in regular inflection for existing words are necessarily associated with phonological deficits, as predicted by the Joanisse and Seidenberg (1999) model (see e.g., Bird et al. 2002; Tyler et al. 2002b).

8. The results reported here were fairly robust to changes in hidden unit numbers and input-output coding schemes. The coding scheme determines the complexity of the mapping problem, and the architecture in some sense determines the computational capacity of the system. Networks closer to their capacity limit (for example, when a more complex encoding of the mappings is used or there are fewer hidden units) are more vulnerable to damage, but breakdown patterns appeared to be robust.

9. Although the network is content-free in terms of an absence of domain-specific knowledge, it nevertheless has knowledge of a sort, in terms of the computational mechanisms that constrain the content that it can learn (see Elman et al. 1996).

10. In Figure 6b, a value of specialisation is reported when one of the routes has experienced a 100% startstate lesion. On the face of it, this may seem confusing. This value is in fact derived by calculating the difference between the decrement in performance caused by a 50% lesion to the intact route and that caused by a 50% lesion to the (100%) damaged route. However, because no further damage is possible beyond 100%, the second term in this comparison is zero. The specialisation value therefore just indexes the decrement in performance caused by a 50% lesion to the single intact route.

11. With regard to the robustness of the findings for the dual-route network, decreasing the hidden units decreased functional specialisation, whereas increasing hidden units had little effect. Recalling that the coding scheme determined the complexity of the mapping problem and the architecture the computational capacity of the system, increasing the mapping complexity reduced specialisation for a fixed capacity. The network showed no critical period for specialisation when damage occurred at different points in training, provided training time was equated (Thomas, in preparation). Changes in learning algorithm, learning rate, discriminability, and noise were not examined in these simulations.

12. As implemented by Mareschal et al., this isn't strictly true. If the quality of featural information in this model were to vary to any extent across the retina, this variation could in principle be used to track position, so that the *what* channel would implicitly code information about location. Moreover, spatiotemporal codes must retain information about shape to predict the movement of all the parts of an object. If objects differ in shape, the *where* channel in the model would implicitly encode information about identity (Mareschal, personal communication, May 2001). However, it seems unlikely that these residual sources of information would be sufficient to generate robust object-reaching behaviour in an impaired model.

13. For illustration, as applied to Coltheart et al.'s (2001) non-developmental model of reading, structure-function correspondences would ensure RN under the following conditions: (1) The GPC route is constrained to learn context-sensitive rules that encode *just enough* context to enable this mechanism to relate graphemes to phonemes, but *not so much* context that it is able to learn whole exception words. For example, a priori, the system must be able to learn variously that *t* → /t/, that *h* → /h/, that *t* followed by *h* in *th* → /T/, that *i* → /i/, that *i* followed by *e* in *_i_e* → /I/, that *i* followed by *g* followed by *h* in *_igh* → /I/, but *not* that *i* preceded by *p* and followed by *n* and *t* in *pint* → /pInt/. The difficulty of calibrating initial computational constraints so precisely is illustrated by the fact that when Coltheart et al. (1993) attempted

to deliberately handcraft a rule-learning algorithm to acquire solely the GPC regularities (that is, the regular pronunciations alone), they could not prevent the GPC route from learning to pronounce 5 of the 47 exception words in the training set. (2) The lexical route is constrained to learn representations of written word forms that preserve zero similarity between the compositional structure of different input forms. Any remaining similarity would permit analogy-based generalisation of novel pronunciations, based on pronunciations of orthographically related lexical forms. For example, if *nog* produced partial activation of the lexical representations for *nod* and *dog*, the pronunciations /nod/ and /dog/ could be combined to generate a pronunciation for the novel word. This computational constraint entails that there is a strict winner-takes-all competition between lexical representations prior to the activation of phonemic output, where “no winner” implies no remaining activation of lexical representations bearing any similarity to the input. The psychological plausibility of these two computational conditions within the developing reading system remains to be seen.

Text removed due to third party copyright

Text removed due to third party copyright

References

Letters “a” and “r” appearing before authors’ initials refer to target article and response, respectively.

- Ahmed, A. & Dutton, G. N. (1996) Cognitive visual dysfunction in a child with cerebral damage. *Developmental Medicine and Child Neurology* 38:736–43. [CT]
- Akshoomoff, N., Pierce, K. & Courchesne, E. (2002) The neurobiological basis of autism from a developmental perspective. *Development and Psychopathology* 14:613–34. [R-AM]
- Anderson, D. I., Campos, J. J., Anderson, D. E., Thomas, T. D., Witherington, D. C., Uchiyama, I. & Barbu-Roth, M. A. (2001a) The flip side of perception-action coupling: Locomotor experience and the ontogeny of visual-postural coupling. *Human Movement Science* 20(4–5):461–87. [rMT]
- Anderson, V., Northam, E., Hendy, J. & Wrennall, J. (2001b) *Developmental neuropsychology: A clinical approach*. Psychology Press. [arMT]
- Ansari, D., Donlan, C., Thomas, M. S. C., Ewing, S., Peen, T. & Karmiloff-Smith, A. (2003) How children with Williams syndrome understand why counting counts. *Journal of Experimental Child Psychology* 85:50–62 [rMT]
- Ansari, D. & Karmiloff-Smith, A. (2002) A typical trajectories of number development. *Trends in the Cognitive Sciences* 6(12):511–16. [rMT]
- Antshel, K. M. & Waisbren, S. E. (in press) Timing is everything: Executive functions in children exposed to elevated levels of phenylalanine. *Neuropsychology*. [rMT]
- Bach, R. P. (1990) Brain plasticity as a basis for recovery of function in humans. *Neuropsychologia* 28:547–54. [YL]
- Backman, J., Bruck, M., Hebert, M. & Seidenberg, M. (1984) Acquisition and use of spelling sound correspondences in reading. *Journal of Experimental Child Psychology* 38:114–33. [CW]
- Baddeley, A. D. (2000) The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences* 11:417–23. [JB]
- Baltes, P. B., Cornelius, S. W., Spiro, A., Nesselroade, J. R. & Willis, S. L. (1980) Integration versus differentiation of fluid/crystallized intelligence in old age. *Developmental Psychology* 6:625–35. [S-CL]
- Baltes, P. B., Lindenberger, U. & Staudinger, U. M. (1998) Life-span theory in developmental psychology. In: *Handbook of child psychology, vol. 1: Theoretical models of human development*, ed. R. M. Lerner. [Editor-in-chief: W. Damon.] Wiley. [S-CL]
- Baltes, P. B. & Singer, T. (2001) Plasticity and the aging mind: An exemplar of the biocultural orchestration of brain and behaviour. *European Review* 9:59–76. [S-CL]
- Banich, M. T., Levine, S. C., Kim, H. & Huttenlocher, P. (1990) The effects of developmental factors on IQ in hemiplegic children. *Neuropsychologia* 28:35–47. [SCL]
- Barbeau, D., Liang, J. J., Robitaille, Y., Quirion, R. & Srivastava, L. K. (1995) Decreased expression of the embryonic form of the neural cell adhesion molecule in schizophrenic brains. *Proceedings of the National Academy of Sciences USA* 92:2785–89. [RC]
- Baron-Cohen, S. (1998) Modularity in developmental cognitive neuropsychology: Evidence from autism and Gilles de la Tourette syndrome. In: *Handbook of mental retardation and development*, ed. J. A. Burack, R. M. Hodapp & E. Zigler. Cambridge University Press. [arMT]
- Baron-Cohen, S., Ring, H. A., Wheelwright, S., Bullmore, E. T., Brammer, M. J., Simmons, A. & Williams, S. C. (1999) Social intelligence in the normal and autistic brain: An fMRI study. *European Journal of Neuroscience* 11:1891–98. [R-AM]
- Basser, L. S. (1962) Hemiplegia of early onset and the faculty of speech with special reference to the effects of hemispherectomy. *Brain* 85:427–60. [SCL]
- Bates, E., Friederici, A. & Wulfeck, B. (1987) Comprehension in aphasia: A cross-linguistic study. *Brain and Language* 32:19–37. [PJ]
- Bates, E., Vicari, S. & Trauner, D. (1999) Neural mediation of language development: Perspectives from lesion studies of infants and children. In: *Neurodevelopmental disorders. Developmental cognitive neuroscience*, ed. H. Tager-Flusberg. MIT Press. [SCL]
- Bellugi, U., Lichtenberger, L., Jones, W., Lai, Z. & St. George, M. (2000) The neurocognitive profile of Williams syndrome: A complex pattern of strengths and weaknesses. *Journal of Cognitive Neuroscience* 12(1):7–30. [CP]
- Bellugi, U., Wang, P. & Jernigan, T. L. (1994) Williams syndrome: An unusual neuropsychological profile. In: *Atypical cognitive deficits in developmental disorders: Implications for brain function*, ed. S. Broman & J. Grafman. Erlbaum. [arMT]
- Benasich, A. A. & Tallal, P. (1996) Auditory temporal processing thresholds, habituation, and recognition memory over the first year. *Infant Behavior and Development* 19:339–57. [PT]
- (2002) Infant discrimination of rapid auditory cues predicts later language impairment. *Behavioural Brain Research* 136:31–49. [PT]
- Bertenthal, B. I., Campos, J. J. & Kermoian, R. (1994) An epigenetic perspective on the development of self-produced locomotion and its consequences. *Current Directions in Psychological Science* 3(5):140–45. [rMT]
- Betancur, C., Corbex, M., Spielesoy, C., Philippe, A., Laplanche, J. L., Launay, J. M., Gillberg, C., Mouren-Simeoni, M. C., Hamon, M., Giros, B., Nosten-Bertrand, M. & Leboyer, M. (2002) Serotonin transporter gene polymorphisms and hyperserotonemia in autistic disorder. *Molecular Psychiatry* 7:67–71. [R-AM]
- Bird, H., Lambon Ralph, M. A., Seidenberg, M. S., McClelland, J. L. & Patterson, K. (in press) Deficits in phonology and past-tense morphology: What’s the connection? *Journal of Memory and Language*. [aMT, JM]
- Bishop, D. V. M. (1990) Handedness, clumsiness and developmental language disorders. *Neuropsychologia* 28:681–90. [aMT]
- (1997a) Cognitive neuropsychology and developmental disorders: Uncomfortable bedfellows. *Quarterly Journal of Experimental Psychology* 50A:899–923. [aMT, YL]
- (1997b) *Uncommon understanding: Development and disorders of language comprehension in children*. Psychology Press. [JB]
- Bishop, D. V. M., Bishop, S. J., Bright, P., James, C., Delaney, T. & Tallal, P. (1999a) Different origin of auditory and phonological processing problems in children with language impairment: Evidence from a twin study. *Journal of Speech, Language, and Hearing Research* 42:155–68. [FR]
- Bishop, D. V. M., Carlyon, R. P., Deeks, J. M. & Bishop, S. J. (1999b) Auditory temporal processing impairment: Neither necessary nor sufficient for causing language impairment in children. *Journal of Speech, Language and Hearing Research* 42:1295–310. [aMT]
- Blackwell, A. & Bates, E. (1995) Inducing agrammatic profiles in normals: Evidence for the selective vulnerability of morphology under cognitive resource limitation. *Journal of Cognitive Neuroscience* 7:228–57. [PJ]
- Bradshaw, J. L. (2001) *Developmental disorders of the frontostriatal system: Neuropsychological, neuropsychiatric, and evolutionary perspectives*. Psychology Press. [rMT]
- Bray, N. W., Reilly, K. D. & Grupe, L. A. (1997) Neural network models and mechanisms of strategy development. *Developmental Review* 17:525–66. [MHC]
- Bretherton, I. (1996) Internal working models of attachment relationships as related to resilient coping. In: *Development and vulnerability in close relationships*, ed. G. G. Noam & K. W. Fischer. Erlbaum. [MF]
- Briscoe, J., Bishop, D. V. M. & Norbury, C. (2001a) Phonological processing, language and literacy: A comparison of children with mild to moderate hearing loss and those with specific language impairment. *Journal of Child Psychology and Psychiatry and Allied Disciplines* 42. [JB]
- Briscoe, J., Gathercole, S. E. & Marlow, N. (2001b) Everyday memory and

- cognitive ability in children born very prematurely. *Journal of Child Psychology and Psychiatry and Allied Disciplines* 42(6):749–54. [JB]
- Broca, P. (1861) Remarks on the seat of the faculty of articulated language, following an observation of aphemia (loss of speech). *Bulletin de la Société Anatomique* 6:330–57. [ET]
- Brown, G. D. A. (1997) Connectionism, phonology, reading, and regularity in developmental dyslexia. *Brain and Language* 59:207–35. [MHC, aMT]
- Brown, G. D. A. & Loosmore, R. (1995) A computational approach to dyslexic reading and spelling. In: *Developmental and acquired dyslexia*, ed. C. K. Leong & R. M. Joshi. Kluwer Academic. [MHC]
- Brown, J. H., Johnson, M. H., Paterson, S. J., Gilmore, R., Longhi, E. & Karmiloff-Smith, A. (2003) Spatial representation and attention in toddlers with Williams syndrome and Down syndrome. *Neuropsychologia* 41(8):1037–46. [rMT]
- Bruner, J. (2001) Self-making and world-making. In: *Narrative and identity: Studies in autobiography, self, and culture*, ed. J. Brockmeier & D. Carbaugh. John Benjamins. [MF]
- Brunsdon, R. K., Hannan, T. J., Nickels, L. & Coltheart, M. (2002) Successful treatment of sublexical reading deficits in a child with dyslexia of the mixed type. *Neuropsychological Rehabilitation* 12:199–229. [CT]
- Bryne, B. (1993) Learning to read in the absences of phonemic awareness? A comment on Cossu, Rossini and Marshall (1993). *Cognition* 48:285–88. [YL]
- Buitelaar, J. K. & Willemsen-Swinkels, S. H. (2000) Medication treatment in subjects with autistic spectrum disorders. *European Child and Adolescent Psychiatry* 9(1):185–197. [R-AM]
- Bullinaria, J. A. (1997) Modelling reading, spelling, and past tense learning with artificial neural networks. *Brain and Language* 59:236–66. [aMT]
- Bullinaria, J. A. & Chater, N. (1995) Connectionist modelling: Implications for cognitive neuropsychology. *Language and Cognitive Processes* 10:227–64. [aMT]
- Bunge, S. A., Dudukovic, N. M., Thomason, M. E., Vaidya, C. J. & Gabrieli, J. D. (2002) Immature frontal lobe contributions to cognitive control in children: Evidence from fMRI. *Neuron* 33:301–11. [ET]
- Cabeza, R. (2002) Hemispheric asymmetry reduction in older adults. *Psychology and Aging* 17:85–100. [S-CL]
- Calder, A. J., Lawrence, A. D., Keane, J., Scott, S. K., Owen, A. M., Christoffels, I. & Young, A. W. (2002) Reading the mind from eye gaze. *Neuropsychologia* 40:1129–38. [R-AM]
- Campos, J. J., Bertenthal, B. I. & Kermonian, R. (1992) Early experience and emotional development: The emergence of wariness of heights. *Psychological Science* 3(1):61–64. [rMT]
- Capirci, O., Sabbadini, L. & Volterra, V. (1996) Language development in Williams syndrome: A case study. *Cognitive Neuropsychology* 13(7):1017–39. [YL, rMT]
- Caramazza, A. (1984) The logic of neuropsychological research and the problem of patient classification in aphasia. *Brain and Language* 21:9–20. [CT]
- Cardell, E. A. & Chenery, H. J. (1999) A cognitive neuropsychological approach to the assessment and remediation of acquired dysgraphia. *Language Testing* 16:353–88. [CT]
- Carey, S. & Diamond, R. (1977) From piecemeal to configurational representation of faces. *Science* 195:312–14. [CT]
- Castelli, F., Frith, C., Happé, F. & Frith, U. (2002) Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain* 125:1839–49. [R-AM]
- Castles, A. & Coltheart, M. (1993) Varieties of developmental dyslexia. *Cognition* 47:149–80. [JB, FR, aMT]
- (1996) Cognitive correlates of developmental surface dyslexia: A single case study. *Cognitive Neuropsychology* 13:25–50. [CT]
- Castles, A., Datta, H., Gayan, J. & Olson, R. (1999) Varieties of developmental reading disorder: Genetic and environmental influences. *Journal of Experimental Child Psychology* 72:73–94. [NEJ]
- Cattell, R. B. (1971) *Abilities: Their structure, growth and action*. Houghton Mifflin. [S-CL]
- Chapman, L. J. & Chapman, J. P. (2001) Commentary on two articles concerning generalized and specific cognitive deficits. *Journal of Abnormal Psychology* 110:31–39. [RC]
- Chomsky, N. (1959) A review of *Verbal behavior*, by B. F. Skinner. *Language* 35:26–58. [GFM]
- (2002) *On Nature and language*. Cambridge University Press. [R-AM]
- Christiansen, M. H. & Chater, N. (2001) Connectionist psycholinguistics: Capturing the empirical data. *Trends in Cognitive Sciences* 5:82–88. [MHC]
- Chugani, D. C. (2002) Role of altered brain serotonin mechanisms in autism. *Molecular Psychiatry* 7(2):S16–S17. [R-AM]
- Churchill, J. D., Grossman, A. W., Irvin, S. A., Galvez, R., Klintsova, A. Y., Weiler, I. J. & Greenough, W. T. (2002) A converging methods approach to FragileX Syndrome. *Developmental Science* 40:323–38. [rMT]
- Clahsen, H. (1999) Lexical entries and rules of language: A multidisciplinary study of German inflection. *Behavioral and Brain Sciences* 22:991–1060. [CP]
- Clahsen, H. & Almazan, M. (1998) Syntax and morphology in children with Williams syndrome. *Cognition* 68:167–98. [CP, CT, arMT]
- (2001) Compounding and inflection in language impairment: Evidence from Williams syndrome (and SLI). *Lingua* 111:729–57. [YL, CT]
- Clahsen, H., Ring, M. & Temple, C. (2003) Lexical and morphological skills in English children with Williams syndrome. In: *Language competence and development in Williams syndrome*, ed. S. Bartke & J. Siegmüller. Benjamins. [CT]
- Clahsen, H. & Temple, C. (2002) Words and rules in Williams Syndrome. In: *Towards a definition of specific language impairment in children*, ed. Y. Levy & J. Schaeffer. Erlbaum. [aMT]
- (2003) Words and rules in children with Williams Syndrome. In: *Language competence across populations*, ed. Y. Levy & J. Schaeffer. Erlbaum. [CT, rMT]
- Cohen, I. L. (1998) Neural network analysis of learning in autism. In: *Neural networks and psychopathology*, ed. D. J. Stein & J. Ludik. Cambridge University Press. [MHC, aMT]
- Cohen, J. D. & Servan-Schreiber, D. (1992) Context, cortex, and dopamine: A connectionist approach to behaviour and biology in schizophrenia. *Psychological Review* 99:45–77. [RC, aMT]
- Cole, M. (1999) Culture in development. In: *Developmental psychology: An advanced textbook, 4th edition*, ed. M. H. Bornstein & M. E. Lamb. Erlbaum. [S-CL]
- Coltheart, M. (1999) Modularity and cognition. *Trends in Cognitive Sciences* 3:115–20. [aMT]
- 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. [YM]
- Coltheart, M., Rastle, K., Perry, C., Langdon, R. & Ziegler, J. C. (2001) DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review* 108:204–56. [NEJ, CT, arMT]
- Condray, R., Steinhauer, S. R., van Kammen, D. P. & Kasperek, A. (in press) The language system in schizophrenia: Effects of capacity and linguistic structure. *Schizophrenia Bulletin*. [RC]
- Conway, C. M., Ellefson, M. R., Dale, R., Christiansen, M. H. & Schaper, K. (in preparation) *Connectionist models of developmental disorders: A critical appraisal*. [MHC]
- Cornish, K. M. (in press) *Deconstructing the attention deficit in Fragile X syndrome: A developmental neuropsychological approach* (submitted). *Journal of Child Psychology and Psychiatry* [rMT]
- Cornish, K. M., Munir, F. & Cross, G. (1998) The nature of the spatial deficit in young females with Fragile X syndrome: A neuropsychological and molecular perspective. *Neuropsychologia* 36(11):1239–46. [rMT]
- Cossu, G. & Marshall, J. C. (1990) Are cognitive skills prerequisite for learning to read and write? *Cognitive Neuropsychology* 7:21–40. [YL]
- Cossu, G., Rossini, F. & Marshall, J. C. (1993) When reading is acquired but phonemic awareness is not. *Cognition* 46:129–38. [YL]
- Crick, F. (1989) The recent excitement about neural networks. *Nature* 337:129–32. [RPC]
- Crombie, A. C. (1994) *Styles of scientific thinking in the European tradition: The history of argument and explanation especially in the mathematical and biomedical sciences, vols. 1, 2, and 3*. Duckworth. [MF]
- Davis, C. J., Gayan, J., Knopik, V. S., Smith, S. D., Cardon, L. R., Pennington, B. F., Olson, R. K. & DeFries, J. C. (2001) Etiology of reading difficulties and rapid naming: The Colorado Twin Study of Reading Disability. *Behav Genet* 31(6): 625–35. [FR]
- Dawson, G., Webb, S., Schellenberg, G. D., Dager, S., Friedman, S., Aylward, E. & Richards, T. (2002) Defining the broader phenotype of autism: Genetic, brain, and behavioral perspectives. *Development and Psychopathology* 14:581–611. [R-AM]
- Deary, I. J. (2001) Human intelligence differences. A recent history. *Trends in Cognitive Science* 5:127–30. [S-CL]
- De Haan, M., Pascalis, O. & Johnson, M. H. (2002) Specialization of neural mechanisms underlying face recognition in human infants. *Journal of Cognitive Neuroscience* 14:199–209. [S-CL]
- Denenberg, V. H. (1999) A critique of Mody, Studdert-Kennedy and Brady's "Speech perception deficits in poor readers: Auditory processing of phonological coding?" (1997) *Journal of Learning Disabilities* 32:379–83. [PT]
- De Ribaupierre, A. (1993). Structural invariants and individual differences: On the difficulty of dissociating developmental and differential processes. In: *The new structuralism in cognitive development: Theory and research on individual pathways*, ed. R. Case & W. Edelman. Karger. [S-CL]
- Deruelle, C., Mancini, J., Livet, M. O., Cassé-Perrot, C. & de Schonen, S. (1999) Configural and local processing of faces in children with Williams syndrome. *Brain and Cognition* 41:276–98. [YL, CT, aMT]
- Diamond, A. (2001) A model system for studying the role of dopamine in prefrontal cortex during early development in humans. In: *Handbook of*

- developmental cognitive neuroscience*, ed. C. Nelson & M. Luciana. MIT Press. [rMT]
- Diamond, A., Prevor, M., Callender, G. & Druin, D. P. (1997) Prefrontal cortex cognitive deficits in children treated early and continuously for PKU. *Monographs of the Society for Research in Child Development (Monograph No. 252)* 62(4):1–207. [rMT]
- Donnai, D. & Karmiloff-Smith, A. (2000) Williams syndrome: From genotype through to cognitive phenotype. *American Journal of Medical Genetics: Seminars in Medical Genetics* 97:164–71. [PG]
- Dunn, J. & Kirsner, K. (2003) What can we infer from double dissociations? *Cortex* 39(1):1–7. [P]
- Dykens, E. M., Heap, R. M., Ort, S., Finucane, B., Shapiro, L. & Leckman, J. F. (1989) The trajectory of cognitive development in males with FragileX syndrome. *Journal of American Academy of Child and Adolescent Psychiatry* 28(3):422–26. [rMT]
- Ellis, A. W. (1993) *Reading, writing, and dyslexia: A cognitive analysis, 2nd edition*. Erlbaum. [NF]
- Ellis, A. W. & Young, A. W. (1996) *Human cognitive neuropsychology: A textbook with readings*. Psychology Press. [NF]
- (1988) *Human cognitive neuropsychology*. Erlbaum. [CT]
- Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D. & Plunkett, K. (1996) *Rethinking innateness: A connectionist perspective on development*. MIT Press. [PG, S-CL, arMT]
- Engle, R. W. (2002) Working memory capacity is executive attention. *Current Directions in Psychological Science* 11:19–23. [S-CL]
- Farah, M. J., Hammond, K. M., Mehta, Z. & Ratcliff, G. (1989) Category-specificity and modality-specificity in semantic memory. *Neuropsychologia* 27:193–200. [ET]
- Farah, M. J. & McClelland, J. L. (1991) A computational model of semantic memory impairment: Modality specificity and emergent category-specificity. *Journal of Experimental Psychology: General* 120:339–57. [JLM]
- Farah, M. J., O'Reilly, R. C. & Vecera, S. P. (1993) Dissociated overt and covert recognition as an emergent property of a lesioned neural network. *Psychological Review* 100:571–88. [aMT]
- Fawcett, A. J., Nicolson, R. I. & Dean, P. (1996) Impaired performance of children with dyslexia on a range of cerebellar tasks. *Annals of Dyslexia* 46:259–83. [FR, aMT]
- Fisch, G. S., Carpenter, N., Holden, J. J., Howard-Peebles, P. N., Maddalena, A., Borghgraef, M., Steyaert, J. & Fryns, J. P. (1999) Longitudinal changes in cognitive and adaptive behaviour in Fragile X females: A prospective multicenter analysis. *American Journal of Medical Genetics* 83(4):308–12. [rMT]
- Fitch, R. H., Brown, C. P., Tallal, P. & Rosen, G. D. (1997) Effects of sex and MK-801 on auditory-processing deficits associated with developmental microgyric lesions in rats. *Behavioral Neuroscience* 111:404–12. [FR]
- Flax, J. F., Realpe-Bonilla, T., Hirsch, L. S., Brzustowicz, L., Bartlett, C. & Tallal, P. (in press) Specific language impairments in families: Evidence for co-occurrence with reading impairments. *Journal of Speech, Language, and Hearing Research*. [PT]
- Fodor, J. A. (1983) *The modularity of mind*. MIT Press. [YL, aMT]
- (2000) *The mind doesn't work that way: The scope and limits of computational psychology*. MIT Press. [RC, aMT]
- Ford, D. H. (1987) *Human as self-constructing living systems. A developmental perspective on behavior and personality*. Erlbaum. [S-CL]
- Frangiskakis, J. M., Ewart, A. K., Morris, A. C., Mervis, C. B., Bertrand, J., Robinson, B. F., Klein, B. P., Ensing, G. J., Everett, L. A., Green, D., Proschel, C., Cutowski, N. J., Noble, M., Atkinson, D. L., Odelberg, S. J. & Keating, M. T. (1996) LIM-kinase1 hemizygosity implicated in impaired visuospatial constructive cognition. *Cell* 86:59–69. [rMT]
- Frawley, W. (2002) Control and cross-domain mental computation: Evidence from language breakdown. *Computational Intelligence* 18:1–28. [aMT]
- Friedmann, N. & Gvion, A. (2001) Letter position dyslexia. *Cognitive Neuropsychology* 18:673–96. [NF]
- (2002) Letter position dyslexia following left occipito-parietal lesion. Paper presented at the Science of Aphasia: Functional Neuroimaging Studies of Language and its Impairment Conference, Maratea, Italy, June 2002. [NF]
- Friedmann, N. & Nachman-Katz, I. (in press) Neglect dyslexia in a Hebrew-reading child. *Cortex*. [NF]
- Friedmann, N. & Novogrodsky, R. (2003) Syntactic movement in Hebrew-speaking children with G-SLI. Paper presented at EUCLIDES meeting, Wales, UK, May 2003. [NF]
- Friedmann, N. & Rahamim, E. (2002) Developmental letter position dyslexia: The effect of diacritics and comparison of treatment methods. Paper presented at the SCRIPT conference, Tel Aviv University, Tel Aviv, Israel, December 2002. [NF]
- Frith, U. (1985) Beneath the surface of developmental dyslexia. In: *Surface dyslexia*, ed. K. E. Patterson, J. C. Marshall & M. Coltheart. Erlbaum. [CT]
- (1986) A developmental framework for developmental dyslexia. *Annals of Dyslexia* 36: 69–81. [FR]
- (1995) Dyslexia: Can we have a shared theoretical framework? *Educational and Child Psychology* 12:6–17. [aMT]
- (in press) Nativism and neuroconstructivism in the explanation of Williams syndrome. *Biology and Philosophy*. [PG]
- Galaburda, A. (2001) Brain and sounds: Lessons from “dyslexic” rodents. In: *Language, brain and cognitive development: Essays in honor of Jacques Mehler*, ed. E. Dupoux. MIT Press. [FR]
- Galaburda, A. M., Menard, M. T. & Rosen, G. D. (1994) Evidence for aberrant auditory anatomy in developmental dyslexia. *Proceedings of the National Academy of Science USA* 91:8010–13. [FR]
- Galaburda, A. M., Sherman, G. F., Rosen, G. D., Aboitiz, F. & Geschwind, N. (1985) Developmental dyslexia: Four consecutive patients with cortical anomalies. *Annals of Neurology* 18:222–33. [FR]
- Gerrans, P. (2002) Theory of mind in evolutionary psychology. *Biology and Philosophy* 17:305–21. [PG]
- (2003) Nativism and neuroconstructivism in the explanation of Williams syndrome. *Biology and Philosophy* 18:1–11.
- Gerrans, P. & McGeer, V. (2003) Theory of mind in autism and schizophrenia. In: *Individual differences in theory of mind: Influences on typical and atypical development*, ed. B. Repacholi & V. Slaughter. Psychology Press. [PG]
- Goldman-Rakic, P., Isseroff, A., Schwartz, M. & Bugbee, N. (1983) The neurobiology of cognitive development. In: *Handbook of child psychology, vol. 2: Infancy and developmental psychobiology*, ed. M. Haith & J. Campos. Wiley. [SCL]
- Gooding, D. C. & Iacono, W. G. (1995) Schizophrenia through the lens of a developmental psychopathology perspective. In: *Developmental psychopathology: Vol. II. Risk, disorder, and adaptation*, ed. D. Cicchetti & D. J. Cohen. Wiley. [RC]
- Gottlieb, G. & Halpern, C. T. (2002) A relational view of causality in normal and abnormal development. *Development and Psychopathology* 14:421–35. [R-AM]
- Grant, J., Karmiloff-Smith, A., Gathercole, S. A., Paterson, S., Howlin, P., Davies, M. & Udwin, O. (1997) Phonological short-term memory and its relationship to language in Williams syndrome. *Cognitive Neuropsychiatry* 2:81–99. [aMT]
- Grice, S., Spratling, M. W., Karmiloff-Smith, A., Halit, H., Csibra, G., de Haan, M. & Johnson, M. H. (2001) Disordered visual processing and oscillatory brain activity in autism and Williams Syndrome. *NeuroReport* 12:2697–2700. [arMT]
- Hacking, I. (1998) *Mad travelers: Reflections on the reality of transient mental illnesses*. University of Virginia Press. [MF]
- (2002) *Historical ontology*. Harvard University Press. [MF]
- Halliday, L. & Briscoe, J. (2002) Reading strategies in children with mild to moderate hearing loss. Unpublished raw data. [JB]
- Happé, F., Ehlers, S., Fletcher, P. C., Frith, U., Johansson, M., Gillberg, C., Dolan, R. J., Frackowiak, R. S. J. & Frith, C. D. (1996) “Theory of mind” in the brain: Evidence from a PET scan study of Asperger syndrome. *NeuroReport* 8:197–201. [R-AM]
- Hari, R. & Renvall, H. (2001) Impaired processing of rapid stimulus sequences in dyslexia. *Trends in Cognitive Science* 5:525–32. [FR]
- Harm, M. W. & Seidenberg, M. S. (1999) Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review* 106:491–528. [MHC, NEJ, aMT]
- Harm, M. W. & Seidenberg, M. S. (2001) Are there orthographic impairments in phonological dyslexia? *Cognitive Neuropsychology* 18(1):71–92.
- Hauser, M. D., Chomsky, N. & Fitch, W. T. (2002) The faculty of language: What it is, who has it, and how did it evolve? *Science* 298:1569–79. [GFM]
- Hebb, D. O. (1949) *The organization of behavior: A neuropsychological theory*. Wiley. [SCL, PT]
- Herman, A. E., Galaburda, A. M., Fitch, R. H., Carter, A. R. & Rosen, G. D. (1997) Cerebral microgyria, thalamic cell size and auditory temporal processing in male and female rats. *Cerebral Cortex* 7:453–64. [FR]
- Hill, E. (1998) A dyspraxic deficit in specific language impairment and developmental co-ordination disorder? Evidence from hand and arm movements. *Developmental Medicine and Child Neurology* 40:388–95. [aMT]
- Hinton, G. (1989) Connectionist learning procedures. *Artificial Intelligence* 40:185–234. [aMT]
- Hinton, G. E. & Sejnowski, T. J. (1986) Learning and relearning in Boltzmann machines. In: *Parallel distributed processing: Explorations in the microstructure of cognition, vol. 1: Foundations*, ed. D. E. Rumelhart, J. L. McClelland & the PDP Research Group. MIT Press. [aMT]
- Hinton, G. E. & Shallice, T. (1991) Lesioning an attractor network: Investigations of acquired dyslexia. *Psychological Review* 98:74–95. [aMT]
- Hodges, J. R., Patterson, K., Oxbury, S. & Funnell, E. (1992) Semantic dementia: Progressive fluent aphasia with temporal lobe atrophy. *Brain* 115:1783–1806. [CT]

- Hoeffner, J. (1992) Are rules a thing of the past? The acquisition of verbal morphology by an attractor network. *Proceedings of the 14th Annual Meeting of the Cognitive Science Society*. Erlbaum. [aMT]
- Hoeffner, J. H. & McClelland, J. L. (1993) Can a perceptual processing deficit explain the impairment of inflectional morphology in developmental dysphasia? A computational investigation. *Proceedings of the 25th Child Language Research Forum*, ed. E. V. Clark. Stanford University Press. [MHC, JLM, aMT]
- Hoffman, R. E. (1996) Neural networks, cortical connectivity and schizophrenic psychosis. In: *Neural modeling of brain and cognitive disorders*, ed. J. A. Reggia, E. Ruppín & R. Berndt. World Scientific. [aMT]
- Hoffman, R. E. & McClashan, T. H. (1997) Synaptic elimination, neurodevelopment, and the mechanism of hallucinated "voices" in schizophrenia. *American Journal of Psychiatry* 154:1683–89. [MHC]
- Horn, J. L. (1968) Organization of abilities and the development of intelligence. *Psychological Review* 75:242–59. [S-CL]
- Horrobin, D. (1998) The membrane phospholipid hypothesis as a biochemical basis for the neurodevelopmental concept of schizophrenia. *Schizophrenia Research* 30:193–208. [RC]
- Humphreys, K., Ewing, S. & Karmiloff-Smith, A. (2002) *Face processing in Williams syndrome: Infant precursors in developmental disorders*. Presentation to the International Conference on Infant Studies, Toronto, April 2002. [aMT]
- Humphreys, P., Kaufmann, W. E. & Galaburda, A. M. (1990) Developmental dyslexia in women: Neuropathological findings in three patients. *Annals of Neurology* 28:727–38. [FR]
- Huttenlocher, P. R. (2002) *Neural plasticity: The effects of the environment on the development of the cerebral cortex*. Harvard University Press. [rMT]
- Isaacs, E. B., Lucas, A., Chong, W. K., Wood, S. J., Johnson, C. L., Marshall, C., Vargha-Khadem, F. & Gadian, D. G. (2000) Hippocampal volume and everyday memory in children of very low birthweight. *Pediatric Research* 47:73–720. [JB]
- Jackson, N. E. & Coltheart, M. (2001) *Routes to reading success and failure: Toward an integrated cognitive psychology of atypical reading*. Psychology Press. [NEJ, aMT]
- Jacobs, R. A. (1997) Nature, nurture, and the development of functional specializations: A computational approach. *Psychonomic Bulletin Review* 4:229–309. [aMT]
- (1999) Computational studies of the development of functionally specialized neural modules. *Trends in Cognitive Sciences* 3:31–38. [aMT]
- Jacobs, R. A. & Jordan, M. I. (1992) Hierarchical mixtures of experts and the EM algorithm. *Neural Computation* 6:181–214. [aMT]
- Jacobs, R. A., Jordan, M. I., Nowlan, S. J. & Hinton, G. E. (1991) Adaptive mixtures of local experts. *Neural Computation* 3:79–87. [aMT]
- Jarrold, C., Hartley, S. J., Phillips, C. & Baddeley, A. D. (2000) Word fluency in Williams syndrome: Evidence for unusual semantic organisation? *Cognitive Neuropsychiatry* 5:293–319. [CP, rMT]
- Joanisse, M. F. (2000) Connectionist phonology. Ph.D. dissertation, University of Southern California. [aMT]
- Joanisse, M. F. & Seidenberg, M. S. (1998) Specific language impairment in children: A deficit in grammar or processing? *Trends in Cognitive Sciences* 2:240–47. [aMT]
- (1999) Impairments in verb morphology following brain injury: A connectionist model. *Proceedings of the National Academy of Sciences USA* 96:7592–97. [JLM, aMT]
- Johnson, M. H. (1997) *Developmental cognitive neuroscience*. Blackwell. [rMT, GW]
- (1999) Cortical plasticity in normal and abnormal cognitive development: Evidence and working hypothesis. *Development and Psychopathology* 11:419–37. [aMT]
- (2001) Functional brain development in humans. *Nature Review Neuroscience* 2:475–83. [S-CL]
- Johnson, M. H. & Karmiloff-Smith, A. (1992) Can neural selectionism be applied to cognitive development and its disorders? *New Ideas in Psychology* 10(1):35–46. [aMT]
- Johnson, S. C. & Carey, S. (1998) Knowledge enrichment and conceptual change in folkbiology: Evidence from Williams syndrome. *Cognitive Psychology* 37:156–200. [CP]
- Juola, P. & Plunkett, K. (1998) Why double dissociations don't mean much. In: *Proceedings of the Twentieth Annual Meeting of the Cognitive Science Society, Orlando, Florida, March 1998*, pp. 561–66, ed. H. MacKenzie. Erlbaum. [PJ]
- Karmiloff-Smith, A. (1992) *Beyond modularity: A developmental perspective on cognitive science*. MIT Press/Bradford Books. [arMT]
- (1997) Crucial differences between developmental cognitive neuroscience and adult neuropsychology. *Developmental Neuropsychology* 13:513–24. [YL, aMT, CT]
- (1998) Development itself is the key to understanding developmental disorders. *Trends in Cognitive Sciences* 2:389–98. [PG, YL, arMT]
- Karmiloff-Smith, A., Grant, J., Berthoud, I., Davies, M., Howlin, P. & Udwin, O. (1997) Language and Williams syndrome: How intact is "intact"? *Child Development* 68:246–62. [YL, aMT]
- Karmiloff-Smith, A., Grant, J., Ewing, S., Cayette, M. J., Metcalfe, K., Donnai, D., Read, A. P. & Tassabehji, M. (2003a) Using case study comparisons to explore genotype/phenotype correlations in Williams Syndrome. *Journal of Medical Genetics* 40:136–40. [rMT]
- Karmiloff-Smith, A., Scerif, G. & Ansari, D. (2003b) Double dissociations in developmental disorders: Theoretically misconceived, empirically dubious. *Cortex* 39:161–63. [rMT]
- Karmiloff-Smith, A., Scerif, G. & Thomas, M. S. C. (2002) Different approaches to relating genotype to phenotype in developmental disorders. *Developmental Psychobiology* 40:311–22. [arMT]
- Karmiloff-Smith, A. & Thomas, M. S. C. (2003) Developmental disorders. In: *The handbook of brain theory and neural networks, 2nd edition*, ed. M. A. Arbib. MIT Press. [rMT]
- (in press) What can developmental disorders tell us about the neurocomputational constraints that shape development? The case of Williams syndrome. *Development and Psychopathology*. [rMT]
- Karmiloff-Smith, A., Tyler, L. K., Voice, K., Sims, K., Udwin, O., Howlin, P. & Davies, M. (1998) Linguistic dissociations in Williams syndrome: Evaluating receptive syntax in on-line and off-line tasks. *Neuropsychologia* 36: 343–51. [aMT]
- Keating, D. P. & Miller, F. K. (1999) Individual pathways in competence and coping: From regulatory systems to habits of mind. In: *Developmental health and the wealth of nations: Social, biological, and educational dynamics*, ed. D. P. Keating & C. Hertzman. Guilford Press. [MF]
- Knight, R. A. & Silverstein, S. M. (2001) A process-oriented approach for averting confounds resulting from general performance deficiencies in schizophrenia. *Journal of Abnormal Psychology* 110:15–30. [RC]
- Koch, R., Moseley, K., Ning, J., Romstad, A., Guldberg, P. & Guttler, F. (1999) Long-term beneficial effects of the phenylalanine-restricted diet in late-diagnosed individuals with phenylketonuria. *Molecular Genetics and Metabolism* 67:148–55. [rMT]
- Kolb, B. (1995) *Brain plasticity and behavior*. Erlbaum. [SCL]
- Kónya, A. & Pintér, G. (1986) Kategória norma a verbális emlékezet vizsgálatához. (Category norms for testing verbal memory). *Magyar Pszichológiai Szemle* 20:93–111. [CP]
- Kronbichler, M., Hutzler, F. & Wimmer, H. (2002) Dyslexia: Verbal impairments in the absence of magnocellular impairments. *NeuroReport* 13:617–20. [FR]
- Ladd, G. T. (1892) Psychology as a so-called "natural science." *Philosophical Review* 1:24–53. [MF]
- Lai, S. L., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F. & Monaco, A. (2001) A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* 413:519–23. [R-AM, rMT]
- Laing, E., Butterworth, G., Ansari, D., Gsödl, M., Longhi, E., Panagiotaki, G., Paterson, S. & Karmiloff-Smith, A. (2002) Atypical development of language and social communication in toddlers with Williams syndrome. *Developmental Science* 5:233–46. [PG, rMT]
- Laing, E., Hulme, C., Grant, J. & Karmiloff-Smith, A. (2001) Learning to read in Williams syndrome: Looking beneath the surface of atypical reading development. *Journal of Child Psychology and Psychiatry* 42(6):729–39. [rMT]
- Lautrey, J. (in press) A pluralistic approach to cognitive differentiation and development. In: *Models of intelligence for the new millennium*, ed. R. J. Sternberg, J. Lautrey & T. I. Lubart. American Psychological Association. [S-CL]
- Leonard, L. B. (1998) *Children with specific language impairment*. MIT Press. [JLM, PT]
- Leonard, L. B., McGregor, K. K. & Allen, G. D. (1992) Grammatical morphology and speech perception in children with specific language impairment. *Journal of Speech and Hearing Research* 35:1076–85. [JLM]
- Leslie, A. M. (1992) Pretense, autism, and the theory-of-mind module. *Current Directions in Psychological Science* 1:18–21. [R-AM, aMT]
- Levine, D. S. (1996) Modeling dysfunction of the prefrontal executive system. In: *Neural modeling of brain and cognitive disorders*, ed. J. A. Reggia, E. Ruppín & R. Berndt. World Scientific. [aMT]
- Levine, S. C., Huttenlocher, P. R., Banich, M. T. & Duda, E. (1987) Factors affecting cognitive functioning in hemiplegic children. *Developmental Medical Child Neurology* 29:27–35. [SCL]
- Levy, Y., Tennebaum, A. & Ormoy, A. (2000) Spontaneous language of children with specific neurological syndromes. *Journal of Speech, Language and Hearing Research* 43:351–65. [YL]
- Li, S.-C. (2003) Biocultural orchestration of developmental plasticity across levels: The interplay of biology and culture in shaping the mind and behaviour across the lifespan. *Psychological Bulletin* 129:171–94. [S-CL]
- Li, S.-C. & Lindenberger, U. (1999) Cross-level unification: A computational exploration of the link between deterioration of neurotransmitter systems and

- the dedifferentiation of cognitive abilities in old age. In: *Cognitive neuroscience of memory*, ed. L.-G. Nilsson & H. Markowitsch. Hogrefe and Huber. [aMT]
- Li, S.-C., Lindenberger, U., Hommel, B., Aschersleben, G., Prinz, W. & Baltes, P. B. (in press) Lifespan developmental transformations in the couplings of mental abilities and underlying cognitive processes. *Psychological Science*. [S-CL]
- Li, S.-C., Lindenberger, U. & Sikström, S. (2001) Aging cognition: From neuromodulation to representation. *Trends in Cognitive Sciences* 5:479–86. [S-CL]
- Ling, C. X. & Marinov, M. (1993) Answering the connectionist challenge: A symbolic model of learning the past tenses of English verbs. *Cognition* 49:235–90. [aMT]
- Livingstone, M., Rosen, G., Drislane, F. & Galaburda, A. (1991) Physiological and anatomical evidence for a magnocellular defect in developmental dyslexia. *Proceedings of the National Academy of Science USA* 88:7943–47. [FR]
- Logan, J. M., Sanders, A. L., Snyder, A. Z., Morris, J. C. & Buckner, R. L. (2002) Under-recruitment and non-selective recruitment: Dissociable neural mechanisms associated with aging. *Neuron* 33:827–40. [S-CL]
- Lovegrove, W. J., Bowling, A., Badcock, B. & Blackwood, M. (1980) Specific reading disability: Differences in contrast sensitivity as a function of spatial frequency. *Science* 210:439–40. [FR]
- Lukács, Á. & Pléh, C. (1999) Hungarian cross-modal priming and treatment of nonsense words supports the dual process hypothesis. Commentary on Harald Clahsen's "Lexical entries and rules of language." *Behavioral and Brain Sciences* 22:6. [CP]
- Lukács, Á., Pléh, C. & Racsmány, M. (in press) Language in Hungarian Williams syndrome children. [CP]
- Lukács, Á., Racsmány, M. & Pléh, C. (2001) Vocabulary and morphological patterns in Hungarian children with Williams syndrome: A preliminary report. *Acta Linguistica* 48:243–69. [CP]
- Luna, B., Thulborn, K. R., Munoz, D. P., Merriam, E. P., Garver, K. E., Minshew, N. J., Keshavan, M. S., Genovese, C. R., Eddy, W. F. & Sweeney, J. A. (2001) Maturation of widely distributed brain function subserves cognitive development. *Neuroimage* 13:786–93. [ET]
- Lutz, A. (2002) Toward a neurophenomenology as an account of generative passage: A first empirical case study. *Phenomenology and the Cognitive Sciences* 1:133–67. [MF]
- Lynch, E. D., Lee, M. K., Morrow, J. E., Welch, P. L., Leon, P. E. & King, M. C. (1997) Nonsyndromic deafness DFNA1 associated with mutation of a human homolog of the *Drosophila* gene *diaphanous*. *Science* 278:1315–18. [rMT]
- MacWhinney, B. & Leinbach, J. (1991) Implementations are not conceptualizations: Revising the verb learning model. *Cognition* 40:121–57. [aMT]
- Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S. J. & Frith, C. D. (2000) Navigation-related structural change in the hippocampi of taxi drivers. *Proceedings of the National Academy of Sciences USA* 97:4398–403. [S-CL]
- Majerus, S., Palmisano, I., van der Linden, M., Barisnikov, K. & Poncelet, M. (2001) An investigation of phonological processing in Williams syndrome. *Journal of the International Society* 7(2):153. [aMT]
- Makeig, S., Westerfield, M., Jung, T.-P., Enghoff, S., Townsend, J., Courchesne, E. & Sejnowski, T. J. (2002) Dynamic brain sources of visual evoked responses. *Science* 295:690–94. [S-CL]
- Manis, F., Seidenberg, M. S., Doi, L., McBride-Chang, C. & Petersen, A. (1996) On the basis of two subtypes of developmental dyslexia. *Cognition* 58:157–95. [aMT]
- Marchman, V. A. (1993) Constraints on plasticity in a connectionist model of the English past tense. *Journal of Cognitive Neuroscience* 5:215–34. [SCL]
- Marcus, G. F. (2001) *The algebraic mind: Integrating connectionism and cognitive science*. MIT Press. [GFM]
- (2003) *The birth of the mind: The genetic origins of our mental life*. Basic Books. [GFM]
- Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., Rosen, T. J. & Xu, F. (1992) Overregularization in language acquisition. *Monographs of the Society for Research in Child Development* 4 (entire volume). [GW]
- Marengo, S. & Weinberger, D. R. (2000) The neurodevelopmental hypothesis of schizophrenia: Following a trail of evidence from cradle to grave. *Development and Psychopathology* 12:501–27. [RC]
- Mareschal, D., Plunkett, K. & Harris, P. (1999) A computational and neuropsychological account of object-oriented behaviours in infancy. *Developmental Science* 2:306–17. [YM, aMT]
- Mareschal, D. & Shultz, T. R. (1996) Generative connectionist networks and constructivist cognitive development. *Cognitive Development* 11:571–603. [GW]
- Margulies, E. H., Kardia, S. L. & Innis, J. W. (2001) A comparative molecular analysis of developing mouse forelimbs and hindlimbs using serial analysis of gene expression (SAGE). *Genome Research* 11:1686–98. [GFM]
- Marshall, J. C. & Newcombe, F. (1973) Patterns of paralexia: A psycholinguistic approach. *Journal of Psycholinguistic Research* 2:179–99. [NF]
- Matthews, A., Ellis, A. E. & Nelson, C. A. (1996) Development of preterm and full-term infant ability on AB, recall memory, transparent barrier detour, and means-end tasks. *Child Development* 67(6):2658–76. [rMT]
- Mayall, K. & Humphreys, G. W. (1996) Covert recognition in a connectionist model of pure alexia. In: *Neural modeling of brain and cognitive disorders*, ed. J. A. Reggia, E. Ruppin & R. Berndt. World Scientific. [aMT]
- Mayr, E. (1998) *Evolutionary synthesis: Perspectives on the unification of biology*. Harvard University Press. [S-CL]
- McCarthy, R. & Warrington, E. (1990) *Cognitive neuropsychology*. Academic Press. [aMT]
- McClelland, J. L. (1979) On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review* 86:287–330. [GFM]
- McCloskey, M., Rapp, B., Yantis, S. et al. (1995) A developmental deficit in localising objects from vision. *Psychological Science* 6:112–17. [CT]
- McLeod, P., Plunkett, K. & Rolls, E. T. (1998) *Introduction to connectionist modelling of cognitive processes*. Oxford University Press. [aMT]
- Merzenich, M. M. & Jenkins, W. M. (1995) Cortical plasticity, learning and learning dysfunction. In: *Maturational windows and adult cortical plasticity*, ed. B. Julesz & I. Kovacs. Addison-Wesley. [PT]
- Merzenich, M., Jenkins, W., Johnston, P., Schreiner, C., Miller, S. & Tallal, P. (1996) Temporal processing deficits of language-learning children ameliorated by training. *Science* 271:77–81. [JB, PT]
- Miller, K. D. & Erwin, E. (2001) Effects of monocular deprivation and reverse suture on orientation maps can be explained by activity-instructed development of geniculocortical connections. *Visual Neuroscience* 18:821–34. [aMT]
- Mills, D. L., Alvarez, D., St. George, M., Appelbaum, G., Bellugi, U. & Neville, H. (2000) Electrophysiological studies of face processing in Williams syndrome. *Journal of Cognitive Neuroscience* 12(S):47–64. [aMT]
- Mody, M., Studdert-Kennedy, M. & Brady, S. (1997) Speech perception deficits in poor readers: Auditory processing or phonological coding? *Journal of Experimental Child Psychology* 64:199–231. [PT]
- Morris, D. B. (1998) *Illness and culture in the postmodern age*. University of California Press. [MF]
- Moscovic, M. & Umiltà, C. (1990) Modularity and neuropsychology: Modules and central processes in attention and memory. In: *Modular deficits in Alzheimer-type dementia*, ed. M. Schwarz. MIT Press. [YL]
- Müller, R.-A., Rothermel, R. D., Behen, M. E., Muzik, O., Chakraborty, P. K. & Chugani, H. T. (1999) Language organization in patients with early and late left hemisphere lesion: A PET study. *Neuropsychologia* 37:545–57. [R-AM]
- Nagarajan, S., Mahncke, H., Salz, T., Tallal, P., Roberts, T. & Merzenich, M. (1999) Cortical auditory signal processing in poor readers. *Proceedings of the National Academy of Sciences USA* 96:6483–648. [PG]
- Nazzi, T. & Karmiloff-Smith, A. (2002) Categorization abilities in infants and toddlers with Williams syndrome. *NeuroReport* 13:1259–62. [rMT]
- Nazzi, T., Paterson, S. & Karmiloff-Smith, A. (2003) Early word segmentation by infants and toddlers with Williams syndrome. *Infancy* 4:257–71. [rMT]
- Nelson, C. A. & Luciana, M. (2001) *Handbook of developmental cognitive neuroscience*. MIT Press. [rMT]
- Nelson, K. (1993) The psychological and social origins of autobiographical memory. *Psychological Science* 4:7–14. [MF]
- (1996) *Language in cognitive development*. Cambridge University Press. [S-CL]
- Nesselroade, J. R. (1991) The warp and the woof of developmental fabric. In: *Visions of aesthetics, the environment, and development: The legacy of Joachim F. Wohlwill*, ed. R. M. Downs, L. S. Liben & D. S. Palermo. Erlbaum. [S-CL]
- Neville, J. J., Mills, D. L. & Bellugi, U. (1994) Effects of altered auditory sensitivity and age of language acquisition on the development of language-related neural systems: Preliminary studies of Williams syndrome. In: *Atypical cognitive deficits in developmental disorders: Implications for brain function*, ed. S. Broman & J. Grafman. Erlbaum. [aMT]
- Newcombe, F. & Marshall, J. C. (1981) On psycholinguistic classifications of the acquired dyslexias. *Bulletin of the Orton Society* 31:29–46. [NF]
- Nichols, G. & Prigogine, I. (1989) *Exploring complexity: An introduction*. W. H. Freeman. [PJ]
- Nicolson, R. I. & Fawcett, A. J. (2000) Long-term learning in dyslexic children. *European Journal of Cognitive Psychology* 12:357–93. [FR]
- Novogrodsky, R. & Friedmann, N. (2002) Relative clause comprehension in Hebrew-speaking school-age children with G-SLI. Paper presented at the Euroesco Syntax of Normal and Impaired Language Conference, Corinth, Greece, June 2002. [NF]
- O'Laughlin, C. & Thagard, P. (2000) Autism and coherence: A computational model. *Mind and Language* 15:375–92. [MHC]
- Oliver, A., Johnson, M. H., Karmiloff-Smith, A. & Pennington, B. (2000)

- Deviations in the emergence of representations: A neuroconstructivist framework for analysing developmental disorders. *Developmental Science* 3:1–23. [arMT]
- Olson, R. & Datta, H. (2002). Visual-temporal processing in reading-disabled and normal twins. *Reading and Writing* 15(1–2):127–149. [FR]
- O'Reilly, R. C. (1998) Six principles for biologically based computational models of cortical cognition. *Trends in Cognitive Sciences* 2(11):455–62. [rMT]
- O'Reilly, R. C. & Munakata, Y. (2000) *Computational explorations in cognitive neuroscience: Understanding the mind by simulating the brain*. MIT Press. [YM]
- Patterson, K. E. (1981) Neuropsychological approaches to the study of reading. *British Journal of Psychology* 72:151–74. [NF]
- (1990) Alexia and neural nets. *Japanese Journal of Neuropsychology* 6:90–99. [aMT]
- Patterson, K., Plaut, D. C., McClelland, J. L., Seidenberg, M. S., Behrmann, M. & Hodges, J. R. (1996) Connections and disconnections: A connectionist account of surface dyslexia. In: *Neural modeling of brain and cognitive disorders*, ed. J. A. Reggia, E. Ruppini & R. Berndt. World Scientific. [aMT]
- Patterson, K., Seidenberg, M. S. & McClelland, J. L. (1989) Connections and disconnections: Acquired dyslexia in a computational model of reading processes. In: *Parallel distributed processing: Implications for psychology and neuroscience*, ed. R. G. M. Morris. Oxford University Press. [aMT]
- Patterson, K. & Shewell, C. (1987) Speak and spell: Dissociations and word-class effects. In: *The cognitive neuropsychology of language*, ed. M. Coltheart, G. Sartori & R. Job. Erlbaum. [aMT]
- Paulesu, E., McCrory, E., Fazio, F., Menoncello, L., Brunswick, N., Cappa, S. F., Cotelli, M., Cossu, G., Corte, F., Lorusso, M., Pesenti, S., Gallagher, A., Perani, D., Price, C., Frith, C. D. & Frith, U. (2000) A cultural effect on brain function. *Nature Neuroscience* 3:91–96. [S-CL]
- Penke, M., Janssen, U. & Krause, M. (1999) The representation of inflectional morphology: Evidence from Broca's aphasia. *Brain and Language* 68:225–32. [GW]
- Piaget, J. (1971) *Biology and knowledge*. University of Chicago Press. [aMT]
- Pinker, S. (1991) Rules of language. *Science* 253:530–35. [JLM, CP, aMT]
- (1999) *Words and rules*. Basic Books/Weidenfeld and Nicolson. [PG, CP, R-AM, arMT]
- (2001) Talk of genetics and vice versa. *Nature* 413:465–66. [rMT]
- Pinker, S. & Ullman, M. T. (2002) The past and future of the past tense. *Trends in Cognitive Sciences* 6:456–63. [JLM]
- Plaut, D. C. (1995) Double dissociation without modularity: Evidence from connectionist neuropsychology. *Journal of Clinical and Experimental Neuropsychology* 17:291–321. [JLM]
- (2002) Graded modality-specific specialisation in semantics: A computational account of optic aphasia. *Cognitive Neuropsychology* 19(7):603–39. [aMT]
- 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. [aMT, CT]
- Plaut, D. C. & Shallice, T. (1993) Deep dyslexia: A case study of connectionist neuropsychology. *Cognitive Neuropsychology* 10:377–500. [aMT]
- Plunkett, K., Bandelow, S. & Juola, P. (2001) *Stochastic approaches to the breakdown of inflectional morphology in a connectionist network*. (in preparation). [aMT]
- Plunkett, K. & Juola, P. (1999) A connectionist model of English past tense and plural morphology. *Cognitive Science* 23:463–90. [aMT]
- Plunkett, K. & Marchman, V. (1991) U-shaped learning and frequency effects in a multi-layered perceptron: Implications for child language acquisition. *Cognition* 38:1–60. [MHC, aMT]
- (1993) From rote learning to system building: Acquiring verb morphology in children and connectionist nets. *Cognition* 48:21–69. [JLM, aMT]
- (1996) Learning from a connectionist model of the English past tense. *Cognition* 61:299–308. [aMT]
- Posner, M. I., Walker, J. A., Friedrich, F. J. & Rafal, R. D. (1984) Effects of parietal injury on covert orienting of attention. *Journal of Neuroscience* 4:1863–74. [ET]
- Precl, K. & Friedmann, N. (2002) The relationship between attention deficits in ADHD and in specific developmental dyslexia. Paper presented at the Israeli Association for Literacy 17th SCRIPT conference, Bar-Ilan University, Ramat Gan, Israel, June 2002. [NF]
- Quartz, S. (1993) Neural networks, nativism, and the plausibility of constructivism. *Cognition* 48:223–42. [GW]
- (in press) Innateness and the brain. *Biology and Philosophy*. [PG]
- Quartz, S. R. & Sejnowski, T. J. (1997) The neural basis of cognitive development: A constructivist manifesto. *Behavioral and Brain Sciences* 20:537–96. [PG, GW]
- Ramus, F. (2001) Dyslexia – talk of two theories. *Nature* 412:393–95. [FR]
- (2003) Developmental dyslexia: Specific phonological deficit or general sensorimotor dysfunction? *Current Opinion in Neurobiology* 13. [FR]
- Ramus, F., Rosen, S., Dakin, S. C., Day, B. L., Castellote, J. M., White, S. & Frith, U. (2003) Theories of developmental dyslexia: Insights from a multiple case study of dyslexic adults. *Brain* 126:841–65. [GFM, FR]
- Rastle, K. & Coltheart, M. (1999) Serial and strategic effects in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance* 25:482–503. [NEJ]
- Raz, N. (2000) Aging of the brain and its impact on cognitive performance: Integration of structural and functional findings. In: *Handbook of aging and cognition*, ed. F. I. M. Craik & T. A. Salthouse. Erlbaum. [S-CL]
- Reggia, J., Goodall, S., Chen, Y., Ruppini, E. & Whitney, C. (1996a) Modeling post-stroke cortical map reorganization. In: *Neural modeling of brain and cognitive disorders*, ed. J. A. Reggia, E. Ruppini & R. Berndt. World Scientific. [aMT]
- Reggia, J. A., Ruppini, E. & Berndt, R. (1996b) *Neural modeling of brain and cognitive disorders*. World Scientific. [aMT]
- Reiss, A. L., Eliez, S., Schmitt, J. E., Straus, E., Lai, Z., Jones, W. & Bellugi, U. (2000) Neuroanatomy of Williams syndrome: A high-resolution MRI study. *Journal of Cognitive Neuroscience* 12(1):65–73. [GW]
- Reuchlin, M. (1978) Processus vicariants et différences individuelles. *Journal de Psychologie* 2:133–45. [S-CL]
- Reuter-Lorenz, P. A. (2002) New visions of the aging mind and brain. *Trends in Cognitive Sciences* 6:394–400. [S-CL]
- Rice, M. (1999) Specific grammatical limitations in children with Specific Language Impairment. In: *Neurodevelopmental disorders*, ed. H. Tager-Flusberg. MIT Press. [CT]
- Rice, M. L., Wexler, K. & Hershberger, S. (1998) Tense over time: The longitudinal course of tense acquisition in children with specific language impairment. *Journal of Speech, Language and Hearing Research* 41:1412–31. [YL]
- Riva, D. & Cazzaniga, L. (1986) Effects of brain lesion sustained before and after age one. *Neuropsychologia* 24:423–28. [SCL]
- Rosen, G. D., Herman, A. E. & Galaburda, A. M. (1999) Sex differences in the effects of early neocortical injury on neuronal size distribution of the medial geniculate nucleus in the rat are mediated by perinatal gonadal steroids. *Cerebral Cortex* 9:27–34. [FR]
- Rosen, H. J., Petersen, S. E., Linenweber, M. R., Snyder, A. Z., White, D. A., Chapman, L., Dromerick, A. W., Fiez, J. A. & Corbetta, M. D. (2000) Neural correlates of recovery from aphasia after damage to left inferior frontal cortex. *Neurology* 55:1883–94. [R-AM]
- Rossen, M., Klima, E. S., Bellugi, U., Bihrl, A. & Jones, W. (1996) Interaction between language and cognition: Evidence from Williams syndrome. In: *Language learning and behaviour*, ed. J. H. Beitchman, N. J. Cohen, M. M. Konstantareas & R. Tannock. Cambridge University Press. [aMT]
- Rothweiler, M. & Clahsen, H. (1994) Dissociations in SLI children's inflectional systems: A study of participle inflection and subject-verb agreement. *Journal of Logopedics and Phoniatrics* 18:169–79. [CT]
- Rumelhart, D. E. & McClelland, J. L. (1986) On learning the past tense of English verbs. In: *Parallel distributed processing: Explorations in the microstructure of cognition. Vol. 2: Psychological and biological models*, ed. J. L. McClelland, D. E. Rumelhart & the PDP Research Group. MIT Press. [JLM, aMT]
- Rumelhart, D. E., McClelland, J. L. & the PDP Research Group (1986) *Parallel distributed processing: Explorations in the microstructure of cognition, vols. 1 and 2*. MIT Press. [JLM, aMT]
- Rumsey, J. M. & Ernst, M. (2000) Functional neuroimaging of autistic disorders. *Mental Retardation and Developmental Disabilities Research Reviews* 6:171–79. [R-AM]
- Ruppini, E., Horn, D., Levy, N. & Reggia, J. A. (1996) Computational studies of synaptic alterations in Alzheimer's disease. In: *Neural modeling of brain and cognitive disorders*, ed. J. A. Reggia, E. Ruppini & R. Berndt. World Scientific. [aMT]
- Rutherford, M. D., Baron-Cohen, S. & Wheelwright, S. (2002) Reading the mind in the voice: A study with normal adults and adults with Asperger syndrome and high functioning autism. *Journal of Autism and Developmental Disorders* 32:189–94. [R-AM]
- Saffran, E. M. (1982) Neuropsychological approaches to the study of language. *British Journal of Psychology* 73:317–37. [CT]
- Scerif, G., Cornish, K., Wilding, J., Driver, J. & Karmiloff-Smith, A. (in press) Visual selective attention in typically developing toddlers and toddlers with Fragile X and Williams syndrome. *Developmental Science*. [rMT]
- Scholl, B. J. & Leslie, A. M. (2001) Minds, modules, and meta-analysis. *Child Development* 72:696–701. [R-AM]
- Scott, P., Mervis, C. B., Bertrand, J., Klein, B. P., Armstrong, S. C. & Ford, A. L. (1995) Semantic organization and word fluency in 9- and 10-year-old children with Williams syndrome. [Special Issue.] *Genetic Counseling* 6:172–73. [rMT]
- Scoville, W. B. & Milner, B. (2000) Loss of recent memory after bilateral hippocampal lesions (1957). (Reprint of 1957 article). *Journal of Neuropsychiatry and Clinical Neurosciences* 12:103–13. [ET]
- Seidenberg, M. S. & McClelland, J. L. (1989) A distributed, developmental model of word recognition and naming. *Psychological Review* 96:523–68. [aMT]

- Shallice, T. (1988) *From neuropsychology to mental structure*. Cambridge University Press. [JLM, aMT]
- Shallice, T., Warrington, E. K. & McCarthy, R. (1983) Reading without semantics. *Quarterly Journal of Experimental Psychology* 35A:111–38. [aMT]
- Shaywitz, B. A., Shaywitz, S. E., Pugh, K. R., Mencl, P. O., Fulbright, R. K., Skudlarski, P., Constable, R. T., Marchione, K. E., Fletcher, J. M., Lyon, G. R. & Gore, J. C. (2002) Disruption of posterior brain systems for reading in children with developmental dyslexia. *Biological Psychiatry* 52:101–10. [ET]
- Sherman, G. F., Morrison, L., Rosen, G. D., Behan, P. O. & Galaburda, A. M. (1990) Brain abnormalities in immune defective mice. *Brain Research* 532:25–33. [FR]
- Shrager, J. & Johnson, M. H. (1996) Dynamic plasticity influences the emergence of function in a simple cortical array. *Neural Networks* 9:1119–29. [aMT]
- Shultz, T. R., Schmidt, W. C., Buckingham, D. & Mareschal, D. (1995) Modeling cognitive development with a generative connectionist algorithm. In: *Developing cognitive competence: New approaches to process modeling*, ed. T. Simon & G. Halford. Erlbaum. [GW]
- Siegler, R. S. (1994) Cognitive variability: A key to understanding cognitive development. *Current Directions in Psychological Science* 3:1–5. [S-CL]
- Simos, P. G., Fletcher, J. M., Bergman, E., Breier, J. I., Foorman, B. R., Castillo, E. M., Davis, R. N., Fitzgerald, M. & Papanicolaou, A. C. (2002) Dyslexia-specific brain activation profile becomes normal following successful remedial training. *Neurology* 58:1203–13. [ET]
- Singer, W. (1995) Development and plasticity of cortical processing architectures. *Science* 270:758–75. [S-CL]
- Singer Harris, N. G., Bellugi, U., Bates, E., Jones, W. & Rossen, M. (1997) Contrasting profiles of language development in children with Williams and Down syndromes. *Developmental Neuropsychology* 13:345–70. [rMT]
- Snow, C. E., Burns, S. M. & Griffin, P. (1998) *Preventing reading difficulties in young children*. National Academy Press. [PT]
- Snowling, M. J. (2000) *Dyslexia, 2nd edition*. Blackwell. [FR]
- Sokol, S. M., Macaruso, P. & Gollan, T. H. (1994) Developmental dyscalculia and cognitive neuropsychology. *Developmental Neuropsychology* 10:413–41. [CT]
- Spearman, C. (1904) General intelligence, objectively determined and measured. *American Journal of Psychology* 15:201–93. [S-CL]
- Spitzer, M. (1996) Phantom limbs, self-organizing feature maps, and noise-driven neuroplasticity. In: *Neural modeling of brain and cognitive disorders*, ed. J. A. Reggia, E. Ruppini & R. Berndt. World Scientific. [aMT]
- Sprien, O., Rissler, A. H. & Edgell, D. (1995) *Developmental neuropsychology*. Oxford University Press. [SCL]
- Stark, R. & Tallal, P. (1981) Speech acoustic-cue discrimination abilities of normally developing and language-impaired children. *Journal of the Acoustical Society of America* 69:568–74. [PT]
- Stein, D. J. & Ludik, J. (1998) *Neural networks and psychopathology*. Cambridge University Press. [aMT]
- Stein, J. & Walsh, V. (1997) To see but not to read; The magnocellular theory of dyslexia. *Trends in Neurosciences* 20:147–52. [FR]
- Stone, V. E., Baron-Cohen, S. & Knight, R. T. (1998) Frontal lobe contributions to theory of mind. *Journal of Cognitive Neuroscience* 10:640–56. [R-AM]
- Stoodley, C. J., Talcott, J. B., Carter, E. L., Witton, C. & Stein, J. F. (2000) Selective deficits of vibrotactile sensitivity in dyslexic readers. *Neuroscience Letters* 295:13–16. [FR]
- Strauss, E., Satz, P. & Wada, J. A. (1990) An examination of the crowding hypothesis in epileptic patients who have undergone carotid amytal test. *Neuropsychologia* 28:1221–27. [SCL]
- Strauss, E. & Verty, C. (1983) Effects of hemispherectomy in infantile hemiplegia. *Brain and Language* 20:1–11. [SCL]
- Strauss, M. E. (2001) Demonstrating specific cognitive deficits: A psychometric perspective. *Journal of Abnormal Psychology* 110:6–14. [RC]
- Stuart, M. & Howard, D. (1995) KJ: A developmental deep dyslexic. *Cognitive Neuropsychology* 12:793–824. [NF]
- Taatgen, N. A. & Anderson, J. R. (2002) Why do children learn to say “broke”? A model of learning the past tense without feedback. *Cognition* 86(2):123–55. [aMT]
- Tager-Flusberg, H. (2000) Differences between neurodevelopmental disorders and acquired lesions. *Developmental Science* 3:33–34. [aMT]
- (2003) Language impairment in children with complex neurodevelopmental disorders: The case of autism. In: *Language competence across populations: Toward a definition of specific language impairment*, ed. J. Levy, & J. Schaeffer. Erlbaum. [YL]
- Takeuchi, J. K., Koshida-Takeuchi, K., Matsumoto, K., Vogel-Hopker, A., Naitoh-Matsuo, M., Ogura, K., Takahashi, N., Yasuda, K. & Ogura, T. (1999) Tbx5 and Tbx4 genes determine the wing/leg identity of limb buds. *Nature* 398:810–14. [GFM]
- Tallal, P. (1980) Auditory temporal perception, phonics, and reading disabilities in children. *Brain and Language* 9:182–98. [FR, PT]
- (1995) Development and disorders of speech and language: Implications for neural and behavioral plasticity. In: *Maturation windows and adult cortical plasticity*, ed. B. Julesz & I. Kovacs. Addison-Wesley. [JLM]
- Tallal, P., Miller, S. L., Bedi, G., Byrna, G., Wang, X., Nagarajan, S. S., Schreiner, C., Jenkins, W. M. & Merzenich, M. M. (1996) Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science* 271:81–84. [PT]
- Tallal, P. & Piercy, M. (1973a) Defects of non-verbal auditory perception in children with developmental aphasia. *Nature* 241:468–69. [PT]
- (1973b) Developmental aphasia: Impaired rate of non-verbal processing as a function of sensory modality. *Neuropsychologia* 11:389–98. [PT]
- (1974) Developmental aphasia: Rate of auditory processing and selective impairment of consonant perception. *Neuropsychologia* 12:83–93. [PT]
- (1975) Developmental aphasia: The perception of brief vowels and extended stop consonants. *Neuropsychologia* 13:69–74. [PT]
- Tallal, P. & Stark, R. (1981) Speech acoustic cue discrimination abilities of normally developing and language impaired children. *Journal of the Acoustical Society of America* 69:568–74. [aMT]
- Tassabehji, M., Metcalfe, K., Karmiloff-Smith, A., Carette, M. J., Grant, J., Dennis, N., Reardon, W., Splitt, M., Read, A. P. & Donnai, D. (1999) Williams syndrome: Use of chromosomal microdeletions as a tool to dissect cognitive and physical phenotypes. *American Journal of Human Genetics* 64:118–25. [rMT]
- Temple, C. M. (1984) New approaches to the developmental dyslexias. In: *Advances in neurology, 42: Progress in aphasiology*, ed. F. C. Rose. Raven Press. [NF]
- (1991) Procedural dyscalculia and number fact dyscalculia: Double dissociation in developmental dyscalculia. *Cognitive Neuropsychology* 8:155–76. [CT]
- (1992) Developmental memory impairment: Faces and patterns. In: *Mental lives: Case studies in cognition*, ed. R. Campbell. Blackwell. [CT]
- (1994) The cognitive neuropsychology of the developmental dyscalculias. *Cahiers de Psychologie Cognitive/ Current Psychology of Cognition* 133:351–70. [CT]
- (1997) *Developmental cognitive neuropsychology*. Psychology Press. [NF, CT, arMT]
- (2001) Magnetic resonance imaging of reading disorders: Neural correlates of reading difficulty in adults and children. Unpublished doctoral dissertation, Stanford University Medical School. [ET]
- (2002) Brain mechanisms in normal and dyslexic readers. *Current Opinion in Neurobiology* 12:178–83. [ET]
- Temple, C. M., Almazan, M. & Sherwood, S. (2002). Lexical skills in Williams syndrome: A cognitive neuropsychological analysis. *Journal of Neurolinguistics* 15(6):463–95. [CP, aMT]
- Temple, C. M. & Marshall, J. C. (1983) A case study of developmental phonological dyslexia. *British Journal of Psychology* 74:517–33. [NF]
- Temple, E., Deutsch, G. K., Poldrack, R. A., Miller, S. L., Tallal, P., Merzenich, M. M. & Gabrieli, J. D. E. (2003) Neural deficits in children with dyslexia ameliorated by behavioral remediation: Evidence from functional MRI. *Proceedings of the National Academy of Science USA* 100:2860–65. [ET]
- Temple, E., Poldrack, R. A., Protopoulos, A., Nagarajan, S., Salz, T., Tallal, P., Merzenich, M. M. & Gabrieli, J. D. (2000) Disruption of the neural response to rapid acoustic stimuli in dyslexia: Evidence from functional MRI. *Proceedings of the National Academy of Science USA* 97:13907–12. [ET]
- Temple, E., Poldrack, R. A., Salidis, J., Deutsch, G. K., Tallal, P., Merzenich, M. M. & Gabrieli, J. D. (2001) Disrupted neural responses to phonological and orthographic processing in dyslexic children: An fMRI study. *NeuroReport* 12:299–307. [ET]
- Thal, D., Marchman, V. A., Stiles, J., Aram, D., Trauner, D., Nass, R. & Bates, E. (1991) Early lexical development in children with focal brain injury. *Brain and Language* 40:491–527. [SCL]
- Thomas, K. M., Drevets, W. C., Whalen, P. J., Eccard, C. H., Dahl, R. E., Ryan, N. D. & Casey, B. J. (2001) Amygdala response to facial expressions in children and adults. *Biological Psychiatry* 49:309–16. [ET]
- Thomas, M. S. C. (2000) Neuroconstructivism's promise. *Developmental Science* 3(1):35–37. [YM, rMT]
- (2003) Limits on plasticity. *Journal of Cognition and Development* 4(1):95–121. [arMT]
- (in press) Multiple causality in developmental disorders: Methodological implications from computational modelling. *Developmental Science*. [rMT]
- Thomas, M. S. C., Grant, J., Barham, Z., Gsödl, 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* 16:143–76. [CP, CT]
- Thomas, M. S. C. & Karmiloff-Smith, A. (1999) Quo vadis modularity in the 1990s? *Learning and Individual Differences* 10:245–50. [aMT]
- (2002) Modelling typical and atypical cognitive development. *Handbook of childhood development*, ed. U. Goswami. Blackwell. [CP, arMT]
- (2003) Connectionist models of development, developmental disorders and

- individual differences. In: *Models of intelligence: International perspectives*, ed. R. J. Sternberg, J. Lautrey & T. Lubart. American Psychological Association. [arMT]
- (in press) Modeling language acquisition in atypical phenotypes. *Psychological Review*. [MHC, arMT]
- (in preparation) Critical periods and plasticity in feedforward connectionist networks: Implications for cognitive development and language acquisition. [aMT]
- Thomas, M. S. C., Dockrell, J. E., Messer, D., Parmigiani, C., Ansari, D. & Karmiloff-Smith, A. (2003) *Naming in Williams syndrome*. (submitted). [rMT]
- Thomas, M. S. C., Grant, J., Barham, Z., Gsödl, 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* 16:143–76. [arMT]
- Thompson, G. B. & Johnston, R. (2000). Are nonword and other phonological deficits indicative of a failed reading process? *Reading and Writing* 12:63–97. [NEJ]
- Thulborn, K. R., Carpenter, P. A. & Just, M. A. (1999) Plasticity of language-related brain function during recovery from stroke. *Stroke* 30:749–54. [R-AM]
- Toga, A. W. (2002) Neuroimage databases: The good, the bad, and the ugly. *Nature Reviews Neuroscience* 3:302–309. [S-CL]
- Tyler, L. K., de Mornay Davies, P., Anokhina, R., Longworth, C., Randall, B. & Marslen-Wilson, W. D. (2002a) Dissociations in processing past tense morphology: Neuropathology and behavioural studies. *Journal of Cognitive Neuroscience* 14:79–94. [aMT]
- Tyler, L. K., Randall, B. & Marslen-Wilson, W. D. (2002b) Phonology and neuropsychology of the English past tense. *Neuropsychologia* 40:1154–66. [aMT]
- Udwin, O. & Yule, W. (1991) A cognitive and behavioural phenotype in Williams syndrome. *Journal of Clinical and Experimental Neuropsychology* 13:232–44. [aMT]
- Ullman, M. T. (in press) Evidence that lexical memory is part of temporal lobe declarative memory, and that grammatical rules are processed by the frontal/basal-ganglia procedural system. *Brain and Language*. [aMT]
- Ullman, M. T., Corkin, S., Coppola, M., Hickok, G., Growdon, J. H., Koroshetz, W. J. & Pinker, S. (1997) A neural dissociation within language: Evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system. *Journal of Cognitive Neuroscience* 9:266–76. [aMT]
- Ullman, M. T. & Gopnik, M. (1999) Inflectional morphology in a family with inherited specific language impairment. *Applied Psycholinguistics* 20:51–117. [CP, aMT]
- Ullman, M. T., Izvorksi, R., Love, T., Yee, E., Swinney, D. & Hickok, G. (in press). Neural correlates of lexicon and grammar: Evidence from the production, reading, and judgement of inflection in aphasia. *Brain and Language*. [aMT]
- Van der Lely, H. K., Rosen, S. & McClelland, A. (1998) Evidence for a grammar-specific deficit in children. *Current Biology* 8:1253–58. [GFM]
- Van der Lely, H. K. J. & Ullman, M. T. (2001) Past tense morphology in specifically language impaired and normally developing children. *Language and Cognitive Processes* 16:177–218. [arMT]
- Vargha-Khadem, F., Gadian, D. G., Watkins, K. R., Connelly, A., Van-Paesschen, W. & Mishkin, M. (1997) Differential effects of early hippocampal pathology on episodic and semantic memory. *Science* 277:376–80. [JB]
- Vargha-Khadem, F. & Mishkin, M. (1997) Speech and language outcome after hemispherectomy in childhood. In: *Paediatric epilepsy syndromes and their surgical treatment*, ed. I. Tuxhorn, H. Holthausen & H. E. Boenigk. John Libbey. [R-AM]
- Vellutino, F. R. (1979) *Dyslexia: Research and theory*. MIT Press. [FR]
- Vicari, S., Brizzolara, D., Carlesimo, G., Pezzini, G. & Volterra, V. (1996a) Memory abilities in children with Williams syndrome. *Cortex* 32:503–14. [aMT]
- Vicari, S., Carlesimo, G., Brizzolara, D. & Pezzini, G. (1996b) Short-term memory in children with Williams syndrome: A reduced contribution of lexical-semantic knowledge to word span. *Neuropsychologia* 34:919–25. [aMT]
- Volterra, V., Capirci, O. & Caselli, M. C. (2001) What atypical populations can reveal about language development: The contrast between deafness and Williams syndrome. *Language and Cognitive Processes* 16:219–39. [aMT]
- Vuletic, L. & Ferrari, M. (in press) *Perspectives on unusualness: A transfer boy*. Jessica Kingsley Press. [MF]
- Warrington, E. K. (1981) Concrete word dyslexia. *British Journal of Psychology* 72:175–96. [JLM]
- Warrington, E. K. & McCarthy, R. (1987) Categories of knowledge: Further fractionation and an attempted integration. *Brain* 110:1273–96. [JLM]
- Watkins, K. E., Dronkers, N. F. & Vargha-Khadem, F. (2002a) Behavioural analysis of an inherited speech and language disorder: Comparison with acquired aphasia. *Brain* 125(3):452–64. [R-AM]
- Watkins, K. E., Vargha-Khadem, F., Ashburner, J., Passingham, R. E., Connelly, A., Friston, K. J., Frackowiak, R. S., Mishkin, M. & Gadian, D. G. (2002b) MRI analysis of an inherited speech and language disorder: Structural brain abnormalities. *Brain* 125(3):465–78. [R-AM]
- Weinberger, D. R. (1996) On the plausibility of “the neurodevelopmental hypothesis” of schizophrenia. *Neuropsychopharmacology* 14:11S–11S. [RC]
- Westermann, G. (1998) Emergent modularity and U-shaped learning in a constructivist neural network learning the English past tense. In: *Proceedings of the 20th Annual Meeting of the Cognitive Science Society*, ed. M. A. Gensbacher & S. J. Derry. Erlbaum. <http://www.cogsci.ed.ac.uk/~gert/publications/cogsci98.pdf> [aMT, GW]
- (1999) Single mechanism but not single route: Learning verb inflections in constructivist neural networks. *Behavioral and Brain Sciences* 22:1042–43. <http://www.cogsci.ed.ac.uk/~gert/publications/bbs-clahsen.pdf> [GW]
- (2000) Constructivist neural network models of cognitive development. Ph.D. thesis, Division of Informatics, University of Edinburgh. <http://www.cogsci.ed.ac.uk/~gert/publications/thesis.pdf> [GW]
- Whitaker-Azmitia, P. M. (2001) Serotonin and brain development: Role in human developmental diseases. *Brain Research Bulletin* 56:479–85. [R-AM]
- White, S., Milne, E., Rosen, S., Hansen, P. C., Swettenham, J., Frith, U. & Ramus, F. (submitted) The role of sensorimotor processing in dyslexia: A multiple case study of dyslexic children. [FR]
- Woods, B. & Teuber, H-L. (1978) Early onset of complementary specialization of the cerebral hemispheres in man. *Transactions of the American Neurological Association* 98:113–17. [SCL]
- Wright, B. A., Buonomano, D. V., Mahncke, H. W. & Merzenich, M. M. (1997a) Learning and generalization of auditory temporal-interval discrimination in humans. *Journal of Neuroscience* 17:3956–63. [PG]
- Wright, B. A., Lombardino, L. J., King, W. M., Puranik, C. S., Leonard, C. M. & Merzenich, M. M. (1997b) Deficits in auditory temporal and spectral resolution in language-impaired children. *Nature* 387:176–78. [PG]
- Zabell, C. & Everatt, J. (2002) Surface and phonological subtypes of adult developmental dyslexia. *Dyslexia* 8:160–77. [FR]
- Zorzi, M., Houghton, G. & Butterworth, B. (1998) Two routes or one in reading aloud? A connectionist dual-process model. *Journal of Experimental Psychology: Human Perception and Performance* 24:1131–61. [aMT]
- Zukowski, A. (2001) Uncovering grammatical competence in children with Williams syndrome. Unpublished doctoral dissertation, Boston University. [rMT]