Experience-dependent brain development as a key to understanding the language system

Gert Westermann

Department of Psychology, Lancaster University

August 2015

Author Note

Gert Westermann, Department of Psychology, Lancaster University, UK

Word count main text: 4244

Correspondence concerning this article should be addressed to Gert Westermann, Department of Psychology, Lancaster University, Lancaster LA1 4YF Phone: +44 (0)1524 592 942, Fax: +44 (0)1524 593 744, E-mail: g.westermann@lancaster.ac.uk

Abstract

An influential view of the nature of the language system is that of an evolved biological system in which a set of rules is combined with a lexicon that contains the words of the language together with a representation of their context. Alternative views, usually based on connectionist modeling, attempt to explain the structure of language on the basis of complex associative processes. Here I put forward a third view that stresses experience-dependent structural development of the brain circuits supporting language as a core principle of the organization of the language system. On this view, embodied in a recent neuroconstructivist neural network of past tense development and processing, initial domain-general predispositions enable the development of functionally specialized brain structures through interactions between experience-dependent brain development and statistical learning in a structured environment. Together, these processes shape a biological adult language system that appears to separate into distinct mechanism for processing rules and exceptions, whereas in reality those subsystems co-develop and interact closely. This view puts experience-dependent brain development in response to a specific language environment at the heart of understanding not only language development but adult language processing as well.

Keywords: experience-dependent brain development; neuroconstructivism; English past tense; connectionist modeling; emergent modularity, massive modularity

Experience-dependent brain development as a key to understanding the language system

Introduction

One of the most fundamental questions in the organization of the cognitive system is whether it is made up of specialized modules that have evolved to take on a specific function, or whether its organization is plastic and functional specialization arises from dynamic interactions with the environment. This debate is closely linked to the question of the origin of cortical specialization in the brain.

The modular view has seen a number of instantiations, notably Fodor's (1981) 'Modularity of Mind' hypothesis, and, more recently, 'massive modularity' (e.g., Carruthers, 2006). On this view, evolved specialized modules exist for specific cognitive functions such as theory of mind (Scholl & Leslie, 1999), number (Dehaene, 1997), face perception (Kanwisher, McDermott, & Chun, 1997), language (Pinker, 1994), and others. One argument put forward to support this view has been that from an evolutionary perspective, many highly functionally specialized modules are more efficient than fewer, more general processing systems so that the former organization would be favored in natural selection (Barrett & Kurzban, 2006). In turn, the discovery of such modules would imply that they evolved because no other process is known that can create complex functional designs in organisms.

A second argument for massive modularity has been one of computational tractability (Cosmides & Tooby, 1994; Samuels, 2005). Learning in some domains, most notably language, would not be feasible in an unconstrained learner, but if knowledge of the target domain is integrated into the learner *a priori* then the learning task becomes tractable. Therefore, evolution would again select for learning systems that are constrained with respect to solutions to a learning problem and that would therefore make learning possible in the first place.

In contrast, the ‘developmental’ view maintains that cortical processing mechanisms are domain general (O’Leary, 1989) and acquire their functional specialization in a protracted developmental process that is affected by the learner’s experience (Quartz & Sejnowski, 1997; Johnson, 2011). According to this view functional modules do not exist but development can lead to specialized areas that, however, overlap and support more than one function. This view has drawn on evidence from neuroscience showing that under altered afferent input, cortical areas assume different roles (for example, in blind people visual cortex supports processing of auditory and tactile stimuli; Burton, 2003) but also on developmental research showing progressive functional specialization and reorganization in the neural structures supporting specific functions such as face processing (de Haan, 2001).

Modular theories of cognition usually de-emphasize the role of development in favor of the evolutionary process of natural selection in shaping cognitive systems. The EvoDevo approach has aimed to redress this imbalance by drawing on recent evidence from evolutionary biology suggesting that evolution and development interact in shaping the organism through epigenetic processes (Gottlieb, 2007; Lickliter & Honeycutt, 2003). A central aspect of this approach is that an evolved organism develops its phenotype, and that development is a powerful force in shaping the adult organism. Furthermore, given recent evidence in epigenetics, there are close bi-directional interactions between development and the functional expression of genes (Moore, 2015).

Experience-dependent brain development

In this paper I focus on the important role of development for the adult cognitive system from a neuroconstructivist perspective. The recent theoretical framework of neuroconstructivism (Mareschal et al., 2007; Westermann et al., 2007; Westermann, Thomas, & Karmiloff-Smith, 2010) focuses on brain development as the basis for cognitive development and asks how different interacting constraints shape this developmental trajectory. These constraints act on different levels: experience-dependent neural development sculpts small-scale neural networks through cell-cell interactions. Different brain regions take on functional specialization through interaction and competition. The physical development of the child affords progressively more complex means to explore and manipulate the physical environment, generating new experiences that lead to cognitive, and neural, change. The changing social environment likewise shapes the child's experiences and therefore, the brain. Finally, there are epigenetic interactions between genes and experiences and these will shape brain structures either directly or through brain-body-environment loops (Westermann et al., 2007). From a neuroconstructivist perspective, cognitive development shapes the brain and the brain shapes cognitive development in a closely coupled loop. An implication of this view is that an understanding of the developmental process is essential for an understanding of the adult cognitive system which is an emergent outcome of this highly interactive developmental process.

In recent years it has become abundantly clear that neural structures can be shaped by experience, and that differences in structure have functional consequences (Casey, Giedd, & Thomas, 2000; Casey, Tottenham, Liston, & Durston, 2005; Goh & Park, 2009; Johnson, 2011; Kitayama & Uskul, 2011; Quartz & Sejnowski, 1997). In the domain of language, attention has turned to the effects of different linguistic environments on brain organization in first and second language learners (Chen et al., 2009; Kochunov et al., 2003). For example, in one study native Chinese speakers living in the United States showed increases in four small regions in frontal, temporal and parietal cortex compared with native English speakers (Kochunov et al., 2003). These differences in brain structure were interpreted as resulting from the different orthographic, phonetic and semantic characteristics of Chinese and English impacting on experience-dependent brain development. In another study, bilingual adults' grey matter differences in left inferior parietal cortex were linked to different levels of proficiency in the second language (Mechelli et al., 2004). In a longitudinal study on the effects of language experience on brain structures, adult interpreters undergoing intense language training showed structural increase in the hippocampus and several cortical areas after only three months of training (Mårtensson et al., 2012).

Given these and many other findings it is plausible to assume that brain regions arrive at functional specialization less through natural selection followed by largely pre-specified development than through interactive neuroconstructivist processes on the basis of domain general learning mechanisms. Under this view it is not necessary to assume pre-specified evolved specialization to end up with efficient and effective specialized processing systems as argued in the massive modularity view.

Tractability in developing systems

But what about the intractability argument for massive modularity? Here, it is worth going back to the origins of this argument. These origins lie in the classical view of learning as one of finding hypotheses as solution to a problem (such as learning the grammar of one's language) and evaluating these hypotheses against encountered data (Gold, 1967). When data are encountered that are incompatible with the current hypothesis the hypothesis is discarded and a new one is chosen. The intractability problem says that the space of possible hypotheses is so large that it becomes unlikely that the learner will ever find the correct hypothesis to the problem at hand. One way to overcome intractability is to restrict the hypothesis space *a priori* so that it contains only a small subset of all possible hypotheses. In this smaller subset the learning problem then becomes much easier, even under restricted experience with the environment. The universal grammar in language acquisition instantiates such an *a priori* restricted hypothesis space: from all possible grammars the subset of all possible human language grammars is pre-selected and the learner only needs to choose between one of these grammars, making language learning tractable.

One problem with restricting the set of examined hypotheses *a priori* is, of course, that the correct hypothesis might be excluded and so the solution will never be found. By invoking evolution this case can be avoided because obviously a learner who can learn the correct hypothesis will have a selective advantage over one who does not.

Yet, restricting the hypothesis space *a priori* is not the only solution to the intractability problem. It has been shown that gradual structural development of a learner provides an alternative to the initial restriction of the hypothesis space (Quartz, 1993). This argument was made with reference to artificial neural networks. The main idea is that the architecture of a neural network – specifically, the number of the network's hidden units – restricts the hypothesis space (Baum, 1989). It is well known, for example, that neural networks without any hidden units can only compute a simple class of problems that are linearly separable (Minsky & Papert, 1969). The more hidden units a network has the more complex are the problems it can learn. If a model starts with a simple architecture and progressively adds hidden units as data are encountered the hypothesis space is gradually expanded in a way that avoids intractability (White, 1990). In fact, it has been shown that such (neuro-)constructivist networks learn in polynomial time whereas training a fixed architecture network with as little as three hidden units can take exponential time or not succeed at all (Baum & Haussler, 1989).

Together, the empirical evidence for experience-dependent structural brain development and the learning theoretic arguments for constructivist learning provide a powerful argument for the importance of structural change as a core principle of psychological development (Quartz, 1999; Quartz & Sejnowski, 1997). Nevertheless, despite providing an alternative to the assumptions of pre-specification inherent in the massively modular approach, structural change has so far found only little reflection in models of psychological development (Shultz, 2003; Westermann, Ruh, & Plunkett, 2009; Westermann, Sirois, Shultz, & Mareschal, 2006).

Neuroconstructivist development in the English past tense

We have recently presented a neuroconstructivist connectionist model of learning the English past tense that takes these considerations on board (Westermann & Ruh, 2012). The English past tense has long been considered a model phenomenon for the language system in general (McClelland & Patterson, 2002; Pinker & Ullman, 2002, Seidenberg & Plaut, 2014). This is because, like language as a whole, the past tense includes rule-like processes as well as exceptions. Regular verbs are inflected by adding *–ed* to the verb stem (e.g., *look-looked*), whereas irregulars come in different varieties and need to be learned and memorized (e.g., *sing-sang* but *bring-brought*). This separation between regular and irregular verbs is one aspect of language where grammatical rules and lexical entries co-exist. The past tense has therefore been termed the ‘drosophila of psycholinguistics’ because a very detailed examination of this small part of language allows us to develop a better understanding of how language is organized in general (Pinker, 1994).

Due to its importance for understanding the architecture of the language system, the mechanisms underlying past tense inflection have been hotly debated. According to the dual-mechanism or Words-and-Rules (WR) theory (Clahsen, 1999; Pinker, 1991, 1997; Pinker & Ullman, 2002; Ullman, 2004; Ullman et al., 1997) the distinction between regular and irregular forms directly maps onto separate mechanisms in the underlying processing structure, with regular forms produced by a mental rule and irregular forms retrieved from the mental lexicon. This approach is situated within the tradition of massive modularity with the assumption of evolved, specialized qualitatively distinct modules responsible for different aspects of processing. Evidence for WR is usually derived from identifying dissociations between regular and irregular verbs that occur in many aspects of processing. For example, when learning language children often make overregularization errors with irregular forms such as *comed* and *eated* but errors for regular verbs are less frequent (Marcus et al., 1992). In adults (and children) the regular form is generalized freely to novel verbs such as *googled*. In adult aphasic patients cases have been reported with selective sparing of regular forms in some and sparing of irregular forms in others (Marslen-Wilson & Tyler, 1997; Tyler et al., 2002; Ullman et al., 1997; Ullman et al., 2005). And some imaging studies have identified brain areas that are more active for regular inflections than for irregulars and others that show the opposite activation pattern (Dhond, Marinkovic, Dale, Witzel, & Halgren, 2003; Jaeger et al., 1996).

Although WR is intuitively elegant in that it maps a grammatical property (regularity) onto the underlying processing structure, this account nevertheless has been undermined by empirical data showing that dissociations between regulars and irregulars are much less clear cut than predicted by this theory. For example, it has been shown that errors in acquisition are predicted by statistical factors such as frequency, the phonological complexity of the verb stem, and the number of phonological friends (verbs with a similar stem and similar past tense form such as *sing* and *ring*) and enemies (verbs with a similar stem and different past tense forms such as *sing* and *bring*) (Marchman, 1997). Most imaging studies have not found clear regular/irregular dissociations, but some found enhanced activity in some areas for irregulars but no specific areas for regulars (Joanisse & Seidenberg, 2005), and others have argued that observed differences are best explained by phonological, not grammatical, differences between verbs (Desai, Conant, Waldron, & Binder, 2006; Joanisse & Seidenberg, 2005). Likewise, dissociation profiles in aphasic patients are affected by phonological factors (Bird, Lambon-Ralph, Seidenberg, McClelland, & Patterson, 2003), and there are virtually no reported cases of fully preserved irregular inflection with impaired regular inflection (Faroqi-Shah, 2007), a pattern that would be expected if the rule-system were selectively damaged. While proponents of WR have tried to integrate some of this empirical evidence into a modified theory in which regular forms can also be stored in the lexicon and therefore show aspects of associative memory such as frequency effects (Pinker & Ullman, 2002), this modification has made the theory overly descriptive and hard to falsify since it is not clear exactly which regulars should be stored and which should not. As WR has never been implemented in a working model it remains difficult to evaluate its assumptions and predictions in detail.

The empirical evidence showing graded dissociations and effects of statistical factors is more in line with an alternative theory to WR, the single mechanism or connectionist view (Joanisse & Seidenberg, 1999; McClelland & Patterson, 2002; Plunkett & Juola, 1999). On this view, all forms, regulars and irregulars alike, are produced by the same associative mechanism, and processing differences between verbs arise from statistical factors such as frequency, phonological complexity, and similarities between different verbs. Some instantiations of this view claim that whereas regular inflection relies on phonological representations of the verb, irregular forms are produced on the basis of semantic information (Joanisse & Seidenberg, 1999). The single mechanism view is closely associated with connectionist models and there have been several neural network implementations of past tense models (e.g., Joanisse & Seidenberg, 1999; Plunkett & Juola, 1999; Rumelhart & McClelland, 1986).

While the empirical evidence appears to be incompatible with a modular, dual mechanism system of verb inflection, single mechanism approaches as well have come under criticism. Single mechanism arguments are generally put forward in the form of implemented computational models, and by necessity they can be analyzed in more detail than the verbal WR theory and therefore offer intrinsically more scope for criticism. Nevertheless, two valid points can be raised: first, different models have each focused on a small subset of phenomena (Pinker & Ullman, 2003). There are models for acquisition, others for adult generalization, and yet others for breakdown after brain damage. However, given that it is the same mind/brain that passes through acquisition, adult processing and breakdown in principle a single model should be able to account for this range of behaviors. Second, most of the single mechanism models have accounted for data more in a proof-of-concept manner rather than in ways that stand up to close comparison with empirical data. For example, models of children's errors in learning the past tense have mostly not closely compared error patterns with reported children's errors (a task that admittedly is not facilitated by the relative scarcity of such empirical data). Similarly, the best-known model of breakdown after brain damage (Joanisse & Seidenberg, 1999) was unable to produce cases in which irregular performance is higher than that for regulars, although this pattern is frequently found in aphasic patients (Faroqi-Shah, 2007).

Recently Westermann and Ruh (2012) put forward an alternative theory to the established positions. This neuroconstructivist view of past tense processing argued that the core to understanding the inflection system is to appreciate that it emerges from experience-dependent brain development in a structured environment. Like in other connectionist approaches there is a single domain general learning mechanism, but experience-dependent structural change leads to a differentiated architecture in the adult system. Specifically, because new structure is added when the existing structure is insufficient to learn new forms, the system develops areas of functional specialization for forms that are easy to process and those that are hard to process. Westermann and Ruh (2012) presented a neural network model that implemented these principles by starting out with a minimal architecture and gradually adding (and deleting) hidden units as learning progressed.

This neuroconstructivist model accounted for a broad range of empirical data in past tense processing. First, while learning past tense forms the model made characteristic overregularization errors (such as *comed* and *eated*). In accordance with empirical work, error rates were lower for high-frequency verbs and for verbs that were classified as 'vulnerable' on the basis of statistical properties such as few phonological friends and more phonological enemies (Marchman, 1997).

A second set of results concerned modeling adults' generalization to non-words. After the model had acquired the past tense forms of all verbs, it generalized to new forms in ways that closely matched data from experiments with adults (Prasada & Pinker, 1993). Specifically, the model produced regular inflections for highly unusual novel verbs such as *ploamph*. Previous static connectionist models did not show such default generalization behavior that has been seen as strong evidence for dual mechanism accounts with default application of the rule (Prasada & Pinker, 1993). Importantly, the neuroconstructivist model even displayed default generalization when it was trained with an artificially low proportion of regular verbs so that the regular case became a minority default. Again, the productive application of a minority regular form had been taken as strong evidence for dual mechanism accounts (Marcus, Brinkmann, Clahsen, Wiese, & Pinker, 1995).

Third, when the fully trained model was then damaged by removing units and connections, it accounted for breakdown profiles after brain damage in aphasic patients. Specifically, when some or all of the inserted hidden units were removed the model's performance reflected data from agrammatic aphasic patients with a breakdown in irregular performance with preservation of regular inflections. Conversely, when the initial pathway was lesioned and only the inserted structure was preserved, the model showed a decline for both regular and irregular forms that was more marked for regulars. This impairment profile matched in its range the profiles of aphasic patients with more preserved irregular performance (Faroqi-Shah, 2007).

Fourth, Westermann and Ruh (under revision) showed that the model successfully simulated activation patterns from brain imaging studies of past tense inflection. Using synthetic MRI—the analysis of activation patterns in the model when processing different verbs—the model showed dissociations between regular and irregular forms, but deeper analysis revealed that these dissociations were graded and were based on the distributional statistical properties of verbs.

The neuroconstructivist theory instantiated in this model therefore suggests that the structure of the past tense task, expressed in the relationship of verbs to all other verbs in the language (e.g., phonological overlap with other forms) together with the statistical properties of the verbs themselves (e.g., frequency), manifests itself in the brain's processing structures through development. This process leads to the emergence of areas of functional specialization that superficially appear like modules but that, on closer inspection, reveal graded differences between verbs. A prediction of this view is that processing differences between verbs in the intact brain, and patterns of breakdown in the lesioned brain, are an outcome of the statistical properties of verbs in the language that have become internalized into the brain's neural circuits.

In sum, by integrating experience-dependent structural development into a model of past tense learning and processing, the model was able to provide an integrated account of development, adult generalization, brain activation patterns, and breakdown profiles after brain damage. Much of the empirical data accounted for by the model had previously been taken as evidence of a modular dual mechanisms account of inflection processing, but the model employed only a single associative mechanism. In its comprehensive and detailed accounting for empirical data the neuroconstructivist model went beyond other existing neurocomputational models that relied on static prespecified architectures and that usually only simulated one aspect of past tense processing.

Although the experience-dependent overproduction and deletion of structure in the model agree with proposals about the mechanisms of gradual elaboration of neural circuits in the brain (Greenough, Black & Wallace, 1987), the specific mechanisms employed in the model are unlikely to find direct correspondence at the neural level. In this sense the model is an abstraction from the low-level neural development in the cortex investigating the impact of experience-dependent structural learning on a more general level.

Discussion

The EvoDevo approach to language evolution and development has argued against the notion of evolved and functionally pre-specified modules as an outcome of evolutionary pressures and has shifted the focus on interactions between epigenetic processes and phenotypic development. I have here discussed the importance of experience-dependent structural brain changes as a central aspect of phenotypic cognitive development and I have argued that functional specialization can emerge as the outcome of experience-dependent neural change in a structured environment. I then described a recent neuroconstructivist model of the acquisition and adult processing of the English past tense that implemented these considerations (Westermann & Ruh, 2012; Westermann & Ruh, under revision). The English past tense is particularly well suited for evaluating EvoDevo approaches to language because first, the past tense stands as a model system for language as a whole with a combination of regular (rule-like) and irregular cases, second, it has been studied empirically in great detail and so it is clear what data a successful theory has to account for, and third, one of the main theories explaining past tense processing is based on those massive modularity views arising from evolutionary psychology that have been criticized in EvoDevo approaches.

In the neurocontructivist model, experience-dependent learning mechanisms extract statistical regularities from the language environment and internalize these regularities as regions of functional specialization in the model structure. This process enabled the model to account for a broad range of empirical data from acquisition, adult generalization, brain activation patterns obtained in imaging studies, and patterns of breakdown after brain damage. Much of these data have previously been taken as evidence for a modular account of past tense processing. The model further suggested that explaining dissociations in past tense processing as between regular and irregular verbs is a post-hoc abstraction of in reality graded dissociations that are based on the distributional statistical properties of verbs. At least in this case, therefore, the assumption of modularity arises out of an abstracted view of the data that exaggerates dissociations and then attributes separate mechanisms to the dissociated processes.

The neuroconstructivist view presented here is radically different from the claim that functionally specialized modules are selected for through evolutionary pressures while de-emphasizing phenotypic development. It is more in line with empirical evidence for initially uncommitted cortical structures and experience-dependent structural changes in the brain. In line with the EvoDevo approach it suggests that selection pressures operate less on specialized brain structures than on development mechanisms that enable the effective learning from the statistics of the environment and the construction of brain structures that can process this information efficiently. While the neuroconstructivist model described here focuses on the ‘devo’ aspect of EvoDevo, it is worth considering at what level evolutionary selection might take place. One possibility is that evolution favors systems with ‘modularized’ areas of functional specialization for specific tasks because they facilitate processing of separate tasks by reducing interference (Bullinaria, 2007). From this perspective a genotype could be selected for that favors a developmental outcome of modularized structures. On another perspective one could assume that the modularized structure of the developed past tense system is an emergent outcome of very general learning mechanisms enabling experience-dependent structural development. Under this view the human brain could have evolved to develop in a protracted way in order to maximize the contribution of experience to its structural elaboration (Johnson, 2009) but without a bias to favor certain developmental outcomes. This latter view might be more plausible, given that in such a scenario modularized structures can still develop when they are beneficial for processing, but non-modular systems would develop for cases when they confer a processing advantage (see Bullinaria, 2007).

While the pathways by which genes affect developmental and thus adult cognitive processing structures are currently unknown, at least in developmental disorders a valid initial assumption is that some genetic abnormalities manifest themselves in some alterations of a typical developmental process with a phenotypic outcome of atypical processing structures. Initial work has been done to explore variations on the parameters in neural networks (e.g., numbers of units, connection patterns, activation functions) (Thomas, Ronald, & Forrester, 2009) with promising results, but this work has focused on static networks with a pre-determined architecture. It will be beneficial to extend such work to models that show experience-dependent structural development as a core aspect of their learning. In these types of models additional variations that can be assumed to be an outcome of genetic variation are possible, such as in the rate of structural growth (e.g., overproliferation of new structure, excessive neural pruning) and it is reasonable to expect that changes in these parameters will interact with changes in other aspects of the models' functioning, leading to a richer picture of how genes and cognitive development interact in the development of language.

Acknowledgements

This work was supported by the Economic and Social Research Council (UK) International Centre for Language and Communicative Development (LuCiD; ES/L008955/1).

References

Barrett, H. C., & Kurzban, R. (2006). Modularity in cognition: framing the debate. *Psychological Review, 113*(3), 628-647.

Baum, E. B. (1989). A proposal for more powerful learning algorithms. *Neural Computation, 1*, 201-207.

Baum, E. B., & Haussler, D. (1989). What Size Net Gives Valid Generalizations? *Neural Computation, 1*, 151-160.

Bird, H., Lambon-Ralph, M. A., Seidenberg, M. S., McClelland, J. L., & Patterson, K. (2003). Deficits in phonology and past-tense morphology: What's the connection? *Journal of Memory and Language, 48*(3), 502-526.

Bullinaria, J. A. (2007). Understanding the emergence of modularity in neural systems. *Cognitive Science: a Multidisciplinary Journal*, *31*(4), 673–695.

Burton, H. (2003). Visual cortex activity in early and late blind people. *Journal of Neuroscience*, *23*(10), 4005–4011.

Carruthers, P. (2006). *The Architecture of the Mind*. Oxford, UK: Oxford University Press.

Casey, B. J., Giedd, J. N., & Thomas, K. M. (2000). Structural and functional brain development and its relation to cognitive development. *Biological Psychology, 54*(1-3), 241-257.

Casey, B. J., Tottenham, N., Liston, C., & Durston, S. (2005). Imaging the developing brain: what have we learned about cognitive development? *Trends in Cognitive Sciences, 9*(3), 104-110.

Chen, C., Xue, G., Mei, L., Chen, C., Dong, Q., & Joan, Y. C. (2009). Cultural neurolinguistics. *Progress in Brain Research, 178*, 159-171.

Clahsen, H. (1999). Lexical entries and rules of language: a multidisciplinary study of German inflection. *Behavioral and Brain Sciences, 22*, 991-1013.

Cosmides, L., & Tooby, J. (1994). Origins of domain specificity: The evolution of functional specialization. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping* *the Mind: Domain Specificity in Cognition and Culture*. Cambridge, UK: Cambride University Press.

Dehaene, S. (1997). *The Number Sense: How the Mind Creates Mathematics*. Oxford, UK: Oxford University Press.

de Haan, M. (2001). The neuropsychology of face processing during infancy and childhood. In C. A. Nelson & M. Luciana (Eds.), *Handbook of Developmental Cognitive Neuroscience* (pp. 381–398). Cambridge, MA: MIT Press.

Desai, R., Conant, L. L., Waldron, E., & Binder, J. R. (2006). FMRI of past tense processing: The effects of phonological complexity and task difficulty. *Journal of Cognitive Neuroscience, 18*(2), 278-297.

Dhond, R. P., Marinkovic, K., Dale, A. M., Witzel, T., & Halgren, E. (2003). Spatiotemporal maps of past-tense verb inflection. *NeuroImage, 19*(1), 91-100.

Faroqi-Shah, Y. (2007). Are regular and irregular verbs dissociated in non-fluent aphasia?: A meta-analysis. *Brain Research Bulletin, 74*(1-3), 1-13.

Fodor, J. A. (1981). *The Modularity of Mind*. Cambridge, MA: MIT Press.

Goh, J. O., & Park, D. C. (2009). Culture sculpts the perceptual brain. *Progress in Brain Research, 178*, 95-111.

Gold, E. M. (1967). Language Identification in the Limit. *Information and Control, 10*, 447-474.

Gottlieb, G. (2007). Probabilistic epigenesis. *Developmental Science, 10*(1), 1-11.

Greenough, W. T., Black, J. E., & Wallace, C. S. (1987). Experience and Brain Development. *Child Development*, *58*, 539–229.

Jaeger, J. J., Lockwood, A. H., Kemmerer, D. L., VanValin, R. D., Murphy, B. W., & Khalak, H. G. (1996). A positron emission tomographic study of regular and irregular verb morphology in English. *Language, 72*, 451-497.

Joanisse, M. F., & Seidenberg, M. S. (1999). Impairments in verb morphology after brain injury: A connectionist model. *Proceedings of the National Academy of Sciences of the United States of America, 96*(13), 7592-7597.

Joanisse, M. F., & Seidenberg, M. S. (2005). Imaging the past: Neural activation in frontal and temporal regions during regular and irregular past-tense processing. *Cognitive Affective & Behavioral Neuroscience, 5*(3), 282-296.

Johnson, M. H. (2009). The Human Social Brain: An “Evo-Devo” Perspective. In L. Tommasi, M. A. Peterson, & L. Nadel (Eds.), *Cognitive Biology: Evolutionary* *and Developmental Perspectives on Mind, Brain, and Behavior* (pp. 309–319). Cambridge, MA: MIT Press.

Johnson, M. H. (2011). Interactive Specialization: A domain-general framework for human functional brain development? *Developmental Cognitive Neuroscience, 1*(1), 7-21.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17*(11), 4302-4311.

Kitayama, S., & Uskul, A. K. (2011). Culture, Mind, and the Brain: Current Evidence and Future Directions. *Annual Review of Psychology, 62*(1), 419-449.

Kochunov, P., Fox, P., Lancaster, J., Tan, L. H., Amunts, K., Zilles, K., . . . Gao, J. H. (2003). Localized morphological brain differences between English-speaking Caucasians and Chinese-speaking Asians: new evidence of anatomical plasticity. *NeuroReport, 14*(7), 961-964.

Lickliter, R., & Honeycutt, H. (2003). Developmental dynamics: Toward a biologically plausible evolutionary psychology. *Psychological Bulletin, 129*(6), 819-835.

Marchman, V. A. (1997). Children's productivity in the English past tense: The role of frequency, phonology, and neighborhood structure. *Cognitive Science, 21*(3), 283-303.

Marcus, G. F., Brinkmann, U., Clahsen, H., Wiese, R., & Pinker, S. (1995). German inflection: The exception that proves the rule. *Cognitive Psychology, 29*, 189-256.

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* (Vol. 57, pp. 1-182).

Mareschal, D., Johnson, M. H., Sirois, S., Spratling, M. W., Thomas, M., & Westermann, G. (2007). *Neuroconstructivism: How the Brain Constructs Cognition*. Oxford, UK: Oxford University Press.

Marslen-Wilson, W., & Tyler, L. K. (1997). Dissociating types of mental computation. *Nature, 387*, 592-594.

Mårtensson, J., Eriksson, J., Bodammer, N. C., Lindgren, M., Johansson, M., Nyberg, L., & Lövdén, M. (2012). Growth of language-related brain areas after foreign language learning. *NeuroImage, 63*(1), 240-244.

McClelland, J. L., & Patterson, K. (2002). Rules or connections in past-tense inflections: what does the evidence rule out? *Trends in Cognitive Sciences, 6*(11), 465-472.

Mechelli, A., Crinion, J. T., Noppeney, U., O'Doherty, J., Ashburner, J., Frackowiak, R. S., & Price, C. J. (2004). Structural plasticity in the bilingual brain. *Nature, 431*(7010), 757-757.

Minsky, M. L., & Papert, S. (1969). *Perceptrons*. Cambridge, MA: MIT Press.

Moore, D. S. (2015). *The Developing Genome: An Introduction to Behavioral Epigenetics.* Oxford: Oxford University Press.

O'Leary, D. D. M. (1989). Do Cortical Areas Emerge from a Protocortex? *Trends in Neurosciences*, *12*, 400–406.

Pinker, S. (1991). Rules of language. *Science, 253*, 530-535.

Pinker, S. (1994). *The Language Instinct*. New York: Morrow.

Pinker, S. (1997). Words and rules in the human brain. *Nature, 387*, 547-548.

Pinker, S., & Ullman, M. T. (2002). The past and future of the past tense. *Trends in Cognitive Sciences, 6*(11), 456-463.

Pinker, S., & Ullman, M. T. (2003). Beyond one model per phenomenon. *Trends in Cognitive Sciences, 7*(3), 108-109.

Plunkett, K., & Juola, P. (1999). A connectionist model of English past tense and plural morphology. *Cognitive Science, 23*(4), 463-490.

Prasada, S., & Pinker, S. (1993). Generalization of regular and irregular morphological patterns. *Language and Cognitive Processes, 8*, 1-56.

Quartz, S. R. (1993). Neural networks, nativism, and the plausibility of constructivism. *Cognition, 48*, 223-242.

Quartz, S. R. (1999). The constructivist brain. *Trends in Cognitive Sciences, 3*(2), 48-57.

Quartz, S. R., & Sejnowski, T. J. (1997). The neural basis of cognitive development: A constructivist manifesto. *Behavioral and Brain Sciences, 20*(4), 537-596.

Rumelhart, D., & McClelland, J. L. (1986). On learning the past tense of English verbs: implicit rules or parallel distributed processing? In J. L. McClelland, D. Rumelhart, & the PDP Research Group (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Cambridge, MA: MIT Press.

Samuels, R. (2005). The Complexity of Cognition. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The innate mind: Structure and contents* (pp. 107-121). Oxford, UK: Oxford University Press.

Seidenberg, M. S., & Plaut, D. C. (2014). Quasiregularity and Its Discontents: The Legacy of the Past Tense Debate. *Cognitive Science*, *38*(6), 1190–1228.

Scholl, B. J., & Leslie, A. M. (1999). Modularity, development and 'theory of mind'. *Mind & Language, 14*(1), 131-153.

Shultz, T. R. (2003). *Computational Developmental Psychology*. Cambridge, MA: MIT Press.

Thomas, M. S. C., Ronald, A., & Forrester, N. (2009). A simulated twin study exploring the heritability of past tense acquisition in a population of neural network models. *Poster presented at the Meeting of the Society for Research in Child Development, April 2-4, 2009*. Denver, USA.

Tyler, L., deMornay-Davies, P., Anokhina, R., Longworth, C., Randall, B., & Marslen-Wilson, W. (2002). Dissociations in processing past tense morphology: Neuropathology and behavioral studies. *Journal of Cognitive Neuroscience, 14*(1), 79-94.

Ullman, M. T. (2004). Contributions of memory circuits to language: the declarative/procedural model. *Cognition, 92*(1-2), 231-270.

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*(2), 266-276.

Ullman, M. T., Pancheva, R., Love, T., Yee, E., Swinney, D., & Hickok, G. (2005). Neural correlates of lexicon and grammar: Evidence from the production, reading, and judgment of inflection in aphasia. *Brain and Language, 93*(2), 185-238.

Westermann, G., Mareschal, D., Johnson, M. H., Sirois, S., Spratling, M. W., & Thomas, M. S. C. (2007). Neuroconstructivism. *Developmental Science, 10*(1), 75-83.

Westermann, G., & Ruh, N. (2012). A neuroconstructivist model of past tense development and processing. *Psychological Review, 119*, 649-667.

Westermann, G., & Ruh, N. (under revision). Origins of Dissociations in Imaging the English Past Tense: A Synthetic Brain Imaging Model.

Westermann, G., Ruh, N., & Plunkett, K. (2009). Connectionist approaches to language learning. *Linguistics, 47*(2), 413-452.

Westermann, G., Sirois, S., Shultz, T. R., & Mareschal, D. (2006). Modeling developmental cognitive neuroscience. *Trends in Cognitive Sciences, 10*(5), 227-233.

Westermann, G., Thomas, M. S. C., & Karmiloff-Smith, A. (2010). Neuroconstructivism. In U. Goswami (Ed.), *The Wiley-Blackwell Handbook of Cognitive Development* (2nd ed., pp. 723-748). Chichester: Wiley-Blackwell.

White, H. (1990). Connectionist Nonparametric Regression: Multilayer Feedforward Networks Can Learn Arbitrary Mappings. *Neural Networks, 3*, 535-549.