Journal of Neurolinguistics 49 (2019) 214–223

Contents lists available at ScienceDirect

Journal of Neurolinguistics

journal homepage: www.elsevier.com/locate/jneuroling

# Neuroemergentism: A framework for studying cognition and the brain

Arturo E. Hernandez[a,*], Hannah L. Claussenius-Kalman[a], Juliana Ronderos[a], Anny P. Castilla-Earls[b], Lichao Sun[a], Suzanne D. Weiss[a], David R. Young[c]

[a] Department of Psychology, University of Houston, United States
[b] Department of Communication Sciences and Disorders, University of Houston, United States
[c] Department of Health and Human Performance, University of Houston, United States

ARTICLE INFO

*Keywords:*
Neuronal recycling
Neural reuse
Language as shaped by the brain
Neuroconstructivism
Emergentism
Neuroemergentism

ABSTRACT

There has been virtual explosion of studies published in cognitive neuroscience primarily due to increased accessibility to neuroimaging methods, which has led to different approaches in interpretation. This review seeks to synthesize both developmental approaches and more recent views that consider neuroimaging. The ways in which Neuronal Recycling, Neural Reuse, and Language as Shaped by the Brain perspectives seek to clarify the brain bases of cognition will be addressed. Neuroconstructivism as an additional explanatory framework which seeks to bind brain and cognition to development will also be presented. Despite sharing similar goals, the four approaches to understanding how the brain is related to cognition have generally been considered separately. However, we propose that all four perspectives argue for a form of Emergentism in which combinations of smaller elements can lead to a greater whole. This discussion seeks to provide a synthesis of these approaches that leads to the emergence of a theory itself. We term this new synthesis Neurocomputational Emergentism (or Neuromergentism for short).

The notion that cognitive functions emerge from the combination of smaller parts has a long history within psychology (Bates, Benigni, Bretherton, Camaioni, & Volterra, 1979). Whereas research in the past focused on mental components, the past twenty years have seen a revolution in cognitive neuroscience with a virtual explosion in the number of papers that are published mostly due to the increased accessibility to imaging modalities such as fMRI (Logothetis, 2008). The large influx of data has led to newer research adopting an Emergentist approach, in which specific functions are built out of smaller parts that combine to create a greater whole. For example, Gauthier and colleagues have argued that neural activity in the fusiform face area (FFA) is not specialized for faces per se but is actually the by-product of extensive expertise that is developed over time (Gauthier, Skudlarski, Gore, & Anderson, 2000; Gauthier, Tarr, Anderson, Skudlarski, & Gore, 1999; Gauthier et al., 2014). To test this hypothesis, bird and car experts were asked to look at both cars and birds (Gauthier et al., 2000). Gauthier and colleagues found increased activity in the FFA when these groups of experts looked at objects in their area of expertise.

The work by Gauthier stands in contrast to work by Kanwisher and colleagues who have built on the findings of increased neural activity for faces in an area of the fusiform gyrus first described by Sergent (1992). Based on a series of studies, these researchers (Kanwisher & Yovel, 2006; Kanwisher, McDermott, & Chun, 1997; Kanwisher, Stanley, & Harris, 1999) proposed a more specific function for the fusiform face area (FFA) after confirming its role in the processing of faces. Thus, there is debate about the extent to

* Corresponding author. Department of Psychology, University of Houston, 4800 Calhoun Rd, Houston, TX 77004, United States.
*E-mail address:* aehernandez@uh.edu (A.E. Hernandez).

https://doi.org/10.1016/j.jneuroling.2017.12.010
Received 30 June 2017; Received in revised form 23 December 2017; Accepted 23 December 2017
0911-6044/ © 2018 Elsevier Ltd. All rights reserved.

which the FFA shows activity for faces not because of some tuning to specific categories but rather because of its responsiveness to categories of perceptual expertise.

Work spearheaded by Morton and Johnson further expands on these two theories by proposing the ways in which humans are initially drawn to faces shortly after birth (Johnson, Dziurawiec, Ellis, & Morton, 1991; Morton & Johnson, 1991). These authors then go on to show that face processing changes dramatically across age. As children grow older the perception of faces leads to the recruitment of more and more cortical areas resulting in increased specialization across time. This more developmentally oriented view of face processing shows us how the recognition of faces within a single area of the brain does not capture the ontogentic nature of face perception. In Johnson's interactive specialization view, across-development face processing (like other cognitive functions) is characterized by a complex interaction between biology and environmental experience (de Haan, Humphreys, & Johnson, 2002). In short, face recognition is an emergent function that combines lots of small pieces to create a greater whole.

A number of recent theories have sought to provide a similar account for the neural bases of cognition. The Neuronal Recycling hypothesis espoused by Dehaene proposes a similar view with regard to reading (Dehaene & Cohen, 2007). In his account, the visual word form area (VWFA), another region in the fusiform gyrus, is specialized for reading words. However, unlike faces, words are not visual categories that are part of some evolutionary specialization for reading. Dehaene, like Gauthier, argues that reading arises from extensive experience with the environment, in this case through instruction in reading. In his view, visual word forms involve the matching of a visual object with a sound. The matching of objects to sounds occurs naturally for humans and leads to the specialization of the fusiform gyrus. The VWFA emerges because it is well suited to the properties of reading.

The Neuronal Recycling hypothesis is not the only theory to suggest that the brain is reconfigured to deal with more complex functions. In a similar vein, Anderson argues for Neural Reuse, where newer cognitive functions (i.e. language) will overtake areas that are dedicated to phylogenetically older (i.e. vision or audition) functions (Anderson & Penner-Wilger, 2013; Anderson, 2010, 2016). Along similar lines, Christiansen, Chater and colleagues (Christiansen & Chater, 2008a; Christiansen & Chater, 2016a) have provided a rich conceptual framework that seeks to bring together evolution, acquisition and language processing, termed language as shaped by the brain. These three theories will be discussed in more detail later.

The merging of computational modeling and brain science, as seen in recent theoretical and empirical work, is strongly resonant with a line of work carried out by a group of researchers which culminated in a jointly authored book entitled *Rethinking Innateness* (Elman, Bates, Johnson, & Karmiloff-Smith, 1996). Using computational modeling, psychological experiments and brain science, Elman and colleagues sketch out a blueprint for a developmental cognitive neuroscience approach which considers how cognition transforms across development. This developmental cognitive neuroscience approach can also be seen in the work of Annette Karmiloff-Smith, a co-author of *Rethinking Innateness*, who proposed a neuroconstructivist approach (D'Souza & Karmiloff-Smith, 2016; Karmiloff-Smith, 2009). Her line of research sought to trace the way in which very simple cognitive functions led to higher-level cognitive deficits. For example, one hypothesis of hers is that children with Williams Syndrome who have a deficit in spatial processing with relative sparing in both the social and language domains, might be caused by much simpler deficits one of which is unusual eye movements. These small deficits, propagated over time, result in widespread pervasive effects across the lifespan. However, her approach and that of other researchers who have sought to use dynamic-systems frameworks to study cognition are not always considered in more recent models of cognition and the brain (D'Souza & Karmiloff-Smith, 2016).

Given the amount of work on this topic both in the past and more recent endeavors, a view which seeks to synthesize both developmental approaches and more recent views that consider neuroimaging is strongly needed. This review seeks to provide a new synthesis of this work. The ways in which Neuronal Recycling, Neural Reuse, and Language as Shaped by the Brain perspectives seek to clarify the brain bases of cognition will be addressed. Neuroconstructivism as an additional explanatory framework which seeks to bind brain and cognition to development will also be presented. Despite sharing similar goals, the four approaches to understanding how the brain is related to cognition have generally been considered separately. However, all four perspectives argue for a form of Emergentism in which brain areas involved in relatively simple functions can be recombined to handle more complex cognitive processing. This discussion seeks to provide a synthesis of these approaches which we term Neurocomputational Emergentism, or Neuromergentism for short. The review will end by considering a recent controversy in the field of bilingualism, the presence or absence of an advantage on cognitive control tasks for speakers of two languages. Specifically, we will point out ways in which a Neuroemergentist approach elucidates how language acquisition in bilinguals both affects and is affected by individual differences and ontogenetic changes in cognitive control.

## 1. Emergentism

The notion of Emergentism has its origin in the philosophy of science. A particularly fruitful argumentation was set forth by John Stuart Mill (Mill, Robson, McRae, & RoutledgeKegan PUniversity of Toronto P, 1974). Stuart Mill began by describing a system in which there are a *composition of causes* that lead to an additive or algebraic outcome. He describes cases in which one force" properly speaking, defeats or frustrates another." A body propelled by two separate forces, for example, east and north, respectively, will land in the same spot as if it had been acted upon by each force separately. He then proceeded to describe how chemistry can show a violation of the combinatorial principles seen in mechanics leading to an emergent property:

> *The chemical combination of two substances produces, as is well known, a third substance with properties different from those of either of the two substances separately, or both of them taken together. Not a trace of the properties of hydrogen or oxygen is observable in those of their compound, water*. (p. 371)

The notion of Emergentism was also brought to the fore by Elizabeth Bates when considering language acquisition. In her book

*The Emergence of Symbols*, Bates sought to clarify the ways in which symbols are built up from much smaller parts (Bates et al., 1979). Liz, as everyone called her, drew inspiration from D'Arcy Thompson's book *On Growth and Form* (Thompson, 1917). The classic text stepped through the clear examples of how beehive cells became hexagonal and why soap bubbles are spheres. In both cases, there was nothing about the ingredients (wind, soap and gravity or many round heads) that was programmed to a particular outcome. Rather the outcome occurred because of the interaction of these pieces to create a different whole. In short, Bates and colleagues (Bates et al., 1979) suggested that "language is a new machine built out of old parts" (p. 3). Examples of this can be seen in the strong relationship between the lexicon and grammatical complexity (Bates & Goodman, 1997). One stunning finding was what appears to be a "critical mass" hypothesis, the notion that across various populations, both typically and non-typically developing, there is a point at which a certain accumulation of words leads to the emergence of grammar.

The work of Elizabeth Bates was rooted in the notion of cognition and language as being biologically based. As such her choice of analogies were firmly rooted in biology:

> *Giraffes have the same 24 neck bones that you and I have, but they are elongated to solve the peculiar problems that giraffes are specialized for (i.e., eating leaves high up in the tree). As a result of this particular adaptation, other adaptations were necessary as well, including cardiovascular changes (to pump blood all the way up to the giraffe's brain), shortening of the hind legs relative to the forelegs (to ensure that the giraffe does not topple over), and so on. Should we conclude that the giraffe's neck is a "high-leaf-eating organ"? Not exactly. The giraffe's neck is still a neck, built out of the same basic blueprint that is used over and over in vertebrates, but with some quantitative adjustments. It still does other kinds of "neck work", just like the work that necks do in less specialized species, but it has some extra potential for reaching up high in the tree that other necks do not provide. If we insist that the neck is a leaf-reaching organ, then we have to include the rest of the giraffe in that category, including the cardiovascular changes, adjustments in leg length, and so on. I believe that we will ultimately come to see our "language organ" as the result of quantitative adjustments in neural mechanisms that exist in other mammals, permitting us to walk into a problem space that other animals cannot perceive much less solve*. (Bates, 1999, p. 20, p. 20)

By using the analogy of a giraffe, Bates is able to trace out how very small changes in a problem space (eating leaves high up in a tree) lead to a cascading set of other changes across the entire organism which result in substantial changes in its anatomy and physiology. This type of analogy has more recently been employed in the discussion of bilingualism as a non-linear dynamical system (Hernandez, 2013; Hernandez, Li, & MacWhinney, 2005; Li, Legault, & Litcofsky, 2014). The notion of language as a combination of many different pieces that are stretched, pulled and adapted is also synergistic with Christiansen and Chater's view of Language as Shaped by the Brain and this metaphor can also be applied to other cognitive functions in the way that Neuronal Recycling and Neural Reuse seek to do (Christiansen & Charter, 2016a). The view of physical adaptation is also in line with Neuroconstructivism in highlighting the need to trace the change in the relationship between cognition and the underlying brain substrate. In a sense, the synthesis of these four approaches themselves leads to the emergence of a theory itself, we term this new synthesis Neurocomputational Emergentism or Neuroemergentism.

## 2. Neuronal recycling

Stanislas Dehaene's Neuronal Recycling hypothesis aims to answer the question of how the human brain deals with cultural inventions (i.e., reading and arithmetic) that have not existed long enough to have influenced evolution (Dehaene & Cohen, 2007). A cultural invention is defined as any cognitive skill that permits a learned pattern of thinking but is not necessary to human survival. The hypothesis posits that the brain learns new cultural inventions by reconfiguring cortical areas that initially evolved for other more basic functions. If a cultural tool must be placed somewhere in the brain, it finds the region or neural "niche" whose properties make it best suited to its requirements; however, this does not mean that the brain evolved for that ability due to adaptive pressure. In fact, Dehaene deems the hypothesis similar to "exaptation" in biology (coined by Gould & Vrba, 1982), which is the counterpart to adaptation; many features of organisms are non-adapted but turned out to be useful. For example, feathers when they first appeared on dinosaurs were likely for thermogenesis and/or attracting a mate, but they turned out to permit flying later in the course of evolution (Gould & Vrba, 1982). While the brain must maintain enough plasticity to adopt these functions in the first place, the neuroanatomical organization of these structures constrains the resulting product. Rather than "genetic reshuffling," Neuronal Recycling is an experience-dependent process that occurs throughout a lifetime.

Reading, a cultural invention that is not old enough for a brain region to have evolved to support it, requires neuronal recycling (Dehaene & Fyssen, 2005). Dehaene et al. (2010) state that during reading acquisition a cortical area in the left occipitotemporal sulcus called the visual word form area (VWFA) is recycled and takes on the new function of processing print. The VWFA, part of the object recognition system, is an area found in approximately the same location across literate subjects that reliably becomes active when processing print and develops with adequate environmental input (print exposure). They found VWFA activation in response to written letter strings in adults who learned to read either in childhood or adulthood, but not in illiterate adults. Instead, in illiterate adults the VWFA showed activation in response to faces and checkerboards. The VWFA may develop in the ventral visual stream because reading is a form of object recognition. Additionally, the exact location of the VWFA in an individual can be predicted before reading acquisition based on connectivity patterns (Saygin et al., 2016). Dehaene and Dehaene-Lambertz (2016) embrace the view that the location of the VWFA takes advantage of a pre-existing brain circuit, in line with Neuronal Recycling.

The hypothesis also provides an explanation for how the brain grapples with arithmetic. The visual number form area (VNFA) lies approximately 1 cm lateral and slightly anterior to the VWFA. Letters and digits are visually similar; however, while letter processing requires connectivity to language areas, digit processing requires connectivity to the horizontal segment of the intraparietal sulcus (HIPS) and the parietal lobes, which the location of the VNFA permits. Dehaene and Fyssen (2005) argue that the human ability to

symbolically represent numbers rests on an innate understanding of numbers, and that children learn to attach digits to their pre-existing intraparietal representation of numerosity. For example, when trained to order numbers 1 through 4, macaque monkeys they were able to generalize to numbers 5 through 9 (Brannon & Terrace, 1998). If humans do have an evolved "number sense" (Dehaene, 1997), then by recycling an area of the ventral visual stream that shares connectivity with the parietal lobes the human brain is able to attach symbols to number intuition.

The Neuronal Recycling hypothesis states that cortical areas are recycled within individuals to process new cultural inventions that are relevant to the functions the areas evolved to support. For example, the HIPS evolved to support understanding of numerosity. In learning arithmetic, this area is recycled and called on to support this new function (Dehaene & Fyssen, 2005). The VWFA evolved for object recognition and is adapted for gaining meaning from written words, a very specific form of object recognition. The fact that these cultural inventions use the same areas across individuals indicates that there must be a relationship between the original and the new function of the area. The VWFA and VNFA are not unlimitedly plastic.

While Dehaene provides a compelling account for two specific cognitive functions, the question remains whether this could be developed into a theory that captures a greater number of functions. An understanding of a larger number of skills, and the nature of human consciousness and the human experience, would require a more holistic explanation. If Emergentism is characterized by a non-linear dynamical process in which a process emerges from the combination of simpler parts into a more complex cognitive process, then the Neuronal Recycling hypothesis is only mildly resonant with it. While Dehaene does focus on the VWFA and VNFA, the focus is on their functions without consideration of the separate processes that occur across time leading up to their recruitment in math and reading. The process of learning to read is mediated by brain regions that likely adapted for object-recognition abilities. The previously discussed regions of the ventral visual system went from object-recognition to highly-specialized object-recognition (i.e., letters and numbers). This trajectory is more direct and does not result in the type of non-linear dynamical change observed within the Emergentist approach. By taking a highly plastic brain and changing the external input (language), we arrive at a specialized system, but not one that strays from its original intention: object recognition.

## 3. Neural reuse

In considering the brain, Anderson (2010) proposes the notion of Neural Reuse, taking a broader perspective that fits more closely within an Emergentist view. This hypothesis is similar to Neuronal Recycling as it puts forward the idea that neuronal structures that have served in one capacity may be borrowed for use in recently developed abilities, such as reading. However, Neural Reuse expands the idea of recycling to all novel skills, rather than focusing solely on reading and arithmetic. Finally, these adaptations depend on the similarity of novel and old functions. The adaptation is more successful when the new function has similar properties to a previous function of that particular neural region.

The Visual Word Form Area (VWFA) identified by Dehaene is a prime example of Neural Reuse. First, the VWFA is located in roughly the same brain area across speakers of a multitude of different languages. Second, there is no evolutionary reason for the VWFA's existence, given that reading is a recent function. The presence of a clearly defined and consistently located VWFA in humans provides evidence this area adapted functions from one set of areas for a very different purpose. The original area was likely involved in more elementary functions such as detecting lines, curves, and intersections. The fact that it was dedicated to such basic functions makes it, according to Anderson, an ideal circuit for neural reuse and the development of human language and reading. Anderson points out that areas like the VWFA show how the brain can take evolutionarily older areas and reuse them for newer purposes.

To account for Neural Reuse, Anderson theorizes that evolutionarily older brain areas are activated by a wider variety of cognitive functions. Anderson asserts that areas in the front of the brain are evolutionarily newer while those in the back of the brain are older. This is consistent with his observation that regions closer to the back of the brain are activated by a larger number of tasks than those closer to the front of the brain (for a similar discussion with regard to working memory see Christophel, Klink, Spitzer, Roelfsema, & Haynes, 2017). However, the data are less clear in this respect (for further discussion of an alternative framework see Miller, 2013). For example, Brodmann areas 39 and 40, which are closer to the back of the brain, are likely more recent as they do not appear in macaques. Similarly, Brodmann areas 4 and 6 are likely older based on their impact on motor control, yet are relatively close to the front of the brain. Thus, while older areas tend to be in the back of the brain, Anderson states that it would be better to arrive at a more accurate measure of a brain areas' evolutionary age rather than relying on generalized locations in the brain.

One corollary of Neural Reuse is the notion of massive redeployment. This hypothesis suggests that while functions may not be spatially organized in the most efficient way, using existing brain circuits is more efficient than creating a brand new pathway. A typical brain area is employed in a multitude of diverse cognitive processes. Based on a meta-analysis of functional neuroimaging studies, Anderson found that while 86 brain regions were activated by at least one task, 66% of those were activated by tasks from at least three categories. Only 17% of brain regions were active during the tasks performed in a single category, suggesting highly specified and single-function regions are the exception, rather than the rule.

Anderson provides further support for the massive redeployment hypothesis. Specifically, he asserts that cognitive functions that have been developed more recently will activate a higher quantity of brain regions, and the regions will be more scattered compared to older functions. The categories of tasks used in the fMRI and PET studies Anderson reviewed included relatively newer categories, like language-based tasks, and relatively older categories such as attention and perception-based tasks. Anderson found that language-based tasks indeed activated a more scattered area than attention or perception-based tasks.

## 4. Language as shaped by the brain

Christiansen and Chater, using a similar conceptual framework as Neural Reuse, analyze language processing from a phylogenetic perspective while augmenting this approach with an ontogenetic perspective. To achieve this, they adopt a construction-based approach to language acquisition which is primarily concerned with the production of utterances made up of multiple constructions (Goldberg, 2006; Tomasello, 2005). Language acquisition includes both the numerous attempts at processing individual utterances as well as systematic learning of specific combinations of constructions from the language environment. Thus, language processing and acquisition are intertwined with each other, and both are constrained by the domain-general learning mechanism derived from cultural transmission, perceptuomotor processing, and higher-order cognitive functions that are not language-specific. Christiansen and Chater (2016b) view the evolution of language as the gradual selection of constructions that better fit and adapt from pre-existing neural substrates in the human brain. This evolutionary perspective explains the close fit between the evolution of linguistic structure and the mechanism involved in language acquisition and processing. That is, language has evolved through the selective pressures from language users driven by the properties of a domain-general learning mechanism which impacts language acquisition and processing.

Much like the Emergentist proposal from Bates and colleagues as well as Neural Reuse and Neuronal Recycling, Christiansen and Chater propose that language is an evolutionary system subject to adaptive pressures from the human brain (Christiansen & Chater, 2008a, 2008b). That is, language has evolved over time to fit the human brain, rather than vice versa. Although language promotes communicative opportunities and enhances human survival, the adaption toward language is only one of many selective pressures working on human evolution. It is possible for humans to survive without language, but not the reverse. Natural languages exist only because humans can produce, learn, and process them.

One particularly fruitful metaphor used by Christiansen (1994) is that of language as a beneficial parasite engaged in a symbiotic relationship with its human hosts, without whom it cannot survive (Deacon, 1997). Similar to Darwin's theory of biological evolution for animal species (Darwin, 1871), language has also been adapted through natural selection, which is constrained by cultural transmission as well as the domain-general learning mechanism of language acquisition and processing. Thus, the features of language that are easiest for language learners to acquire and process are more likely to keep passing on to future generations, while those which prove difficult to acquire tend to disappear. For a constructivist grammar view of language, the basic units of selection are those individual constructions of words or combinations associated with syntactic and semantic information. The selective survival of an individual construction is primarily determined by its individual properties (e.g., individual's usage preference) as well as its relationship with other constructions in the linguistic system (e.g., syntactic, semantic, or pragmatic components). For example, English-language learners tend to have an easier time identifying the preterite tense of go (went) rather than lie (lay) due to the higher frequency of the former.

If language had been shaped to fit the human brain, the properties of linguistic structure should reflect the neural and cognitive machinery of language users (Christiansen & Chater, 2016b). Theoretical argumentation used by Anderson (2010) supports the view of linguistic structure proposed by Language as Shaped by the Brain. In order to accomplish language acquisition and processing, a variety of pre-existing neural circuits are redeployed for accommodating an emerging neuronal network for language. According to Anderson's review on 472 fMRI studies of the activation of Brodmann areas (Anderson, 2008), neural co-activation regions were widely distributed across the whole brain during language processing. Moreover, the neural circuit of language was generally active across cognitive tasks ranging from language, executive functioning, vision, memory, attention, and other non-language specific competencies (Anderson, 2010). These findings suggest that language is a product of recycling or reusing of pre-existing neural substrates for other cognitive functions much like the view espoused by Bates and colleagues. Although Christiansen and Chater (2016b) point out the important role of brain structures and functions in the emergence of language, they do not illustrate the relationship clearly as other neurobiological approaches do.

## 5. Neuroconstructivism

Like Language as Shaped by the Brain, Neuroconstructivism considers the importance of ontogenesis, but it further develops the role of multidirectional interactions between genes, neural components, cognition, and the environment during the lifespan (Karmiloff-Smith, 2009). In neuroconstructivist approaches, human abilities, and in particular intelligence, are believed to be the result of a dynamic process rather than a state. At birth, the brain is more highly interconnected and not as specialized, and over time, the environment leads to more specificity in the brain (Karmiloff-Smith, 2006). Both associations and dissociations, as well as genetic mutations that alter ontogenesis, can provide a better insight into development of higher cognitive functions in the brain.

The Neuroconstructivist approach argues that an in-depth analysis of developmental disorders provides the strongest evidence against the modular organization of the brain. As researchers study adult neuropsychological patients, the tendency has been to create a link between cognitive deficits and particular brain regions. Work looking at disorders linked to particular genes have also followed a similar logic. Specifically, researchers who examine certain genetic disorders that present uneven profiles, which manifest themselves in certain cognitive domains while other domains seem intact, are then motivated to search for specific genes that map onto these modules. This overlooks subtle deficits in the intact domains. Neuroconstructivists argue that the brain cannot be selectively impaired in specific areas while intact or relatively intact in other areas. In this sense, Karmiloff-Smith (2009) augments arguments made by Dehaene and Anderson. Specifically, small changes that occur early in life affect the entire cognitive system, leading to changes across both simple and complex cognitive domains.

Examples of uneven profiles in cognitive domains are seen in studies of Williams Syndrome (WS), a genetic developmental

disorder caused by the deletion of about 28 genes from the long arm of chromosome 7 (Donnai & Karmiloff-Smith, 2000). Individuals with WS exhibit proficient language and facial recognition processing behavior (falling within the normal range on norm-referenced face processing tasks) coexisting with a mean IQ of 56 (Mervis et al., 2000) and with very impaired spatial and numerical cognition (Karmiloff-Smith et al., 2004). The brain in individuals with WS is different than a typical adult brain. It is generally smaller and it has overall reduced curvature, increased gyral complexity, in addition to reductions in parietal and limbic regional volumes (Eisenberg, Jabbi, & Berman, 2010). Furthermore, while the typical adult brain develops gradual specialization and hemispheric localization of function over ontogenetic development (i.e. left hemisphere for language), this tends not to occur in the mature brain of individuals with WS, which continues instead to process faces and language bilaterally (Karmiloff-Smith et al., 2004).

In adult neuropsychology studies, researchers have found a double dissociation between numerical abilities affecting two intraparietal units, one for exact number computing and another for approximate number computing (Butterworth, 2010; Dehaene, 1997; Demeyere, Lestou, & Humphreys, 2010). Research on typically developing infants also shows two systems that develop at different rates (Izard, Dehaene-Lambertz, & Dehaene, 2008; Lipton & Spelke, 2003; Xu, 2003). A series of studies examining small exact number discrimination and large approximate number discrimination in infants with WS and with Down Syndrome (DS) show that infants with WS performed like typically developing infants in small number discrimination but had difficulty with large number approximations. The opposite was found for the infants with DS; they performed poorly with small number discrimination but were significantly better with large number approximation (Karmiloff-Smith et al., 2012; Paterson, Brown, Gsodl, Johnson, & Karmiloff-Smith, 1999; Van Herwegen, Ansari, Xu, & Karmiloff-Smith, 2008). These results were then compared to data on sustained attention and saccadic eye movement planning in both infants with WS and infants with DS, which identified deficits in attention shifting in infants with WS and deficits in sustained attention in infants with DS (Brown et al., 2003). The infants with DS tended to scan the overall array and the infants with WS tended to fixate on specific aspects. Deficits and proficiencies in sustained attention and saccadic eye movement planning, both basic-level skills, explain both the numerical deficits and proficiencies exhibited by each syndrome. The Neuroconstructivist perspective concludes that these differences are likely the result of these basic-level deficits and proficiencies in the visual and attention systems in development which impact cognitive-level phenotypes over ontogenetic time (Brown et al., 2003).

Neuroconstructivists have also responded to the assertions made for Neural Reuse. For example, Dekker and Karmiloff-Smith (2010) explicitly questioned whether Anderson (2010) asked how development played a role in the internal workings of a particular cognitive function. Furthermore, Dekker and Karmiloff-Smith (2010) argued that it was unclear if Anderson (2010) implied that a recombination of connections between regions was the only source of complexity relevant to Neural Reuse or whether, similar to Dehaene's Neuronal Recycling, Anderson considered that neural circuits used for old skills could be optimized for a related but new skill which may potentially impact the old skills. This prediction has been supported by Curby and Gauthier (2014) who found that developing expertise for particular objects in a visual domain (i.e. cars) leads to changes in face processing. One possible implication of this criticism could be viewed within the perspective of Emergentism. Specifically, Dekker and Karmiloff-Smith suggest that in order to evaluate a cognitive function it is important to consider non-linear dynamical change.

A neuroconstructivist perspective would clarify this uncertainty by suggesting that despite early evolutionary differentiation rendering certain areas of the cortex biased towards performing specific functions, it is the environmental influence and competition between regions which ultimately gives shape to the adult brain with domain-specific areas (Karmiloff-Smith, 2009). In this way, the neuroconstructivist approach would attempt to provide a full picture for brain development across ontogeny that encompasses both typical and atypical development (for further discussions along these lines see Johnson, 2011).

Neuroconstructivism accounts for the developmental aspect that the preceding three theories leave out. However, Christiansen and Chater's view provides the computational aspect of evolution and language acquisition which is not described in Neuroconstructivism. In this way, each of the four theories described pull from aspects of the merging of computational modeling and brain science seen in *Rethinking Innateness* (Elman et al., 1996).

## 6. Neuroemergentism

In this review, theoretical frameworks that are synergistic with an Emergentist view in which a greater whole arises from the dynamic interaction of smaller parts were presented. In the case of Neuronal Recycling, reading emerges from the combination of sights and sounds which is adapted from the natural world to the written domain. Anderson extends this view by suggesting that evolutionarily newer functions are handled by more disparate areas of the brain whereas older functions tend to be more localized. Christiansen and Chater draw inspiration from the brain as a metaphor with which to conceptualize a theory of language acquisition, evolution and processing. Their view has also been further expanded by incorporating computational modeling. Discussion of Karmiloff-Smith's framework brings us back to the brain and to a dynamic view in which generalized deficits lead to more specific impairments. In her view, the restructuring of cognition has a cascading effect that can later lead to what appear to be very specific deficits. However, the causes are not specific in and of themselves, a point which can only be understood by tracking changes that occur from infancy across child and adolescent development and into adulthood.

One specific aspect that is made clear is the need to consider ontogenesis. For example, Anderson (2010) proposes that specialization is differentiated on a posterior-to-anterior gradient with the latter being less specialized and the former being more specialized. However, as he notes the data do not fit this pattern very cleanly. Primary auditory, visual and somatosensory cortex lie in the posterior portions of the brain. Primary motor cortex is in the frontal lobe. Whereas Anderson's view takes into account results from a large number of neuroimaging studies, the exception to the posterior-to-anterior rule becomes evident when taking development into account. For example, Hernandez and Li (2007) proposed the sensorimotor hypothesis which helps to explain age of

acquisition (AoA) effects across non-verbal and language domains, including the learning of a second language. These sensorimotor hubs develop from basic primary processing (i.e. topographic visual perception) to more specialized processing (i.e. motion or color mapping). Over time there is an additional increase in the interconnectivity between sensorimotor and association areas. These AoA effects are also driven by changes in the subcortical connections between the basal ganglia and the frontal lobe which lead to a more "cognitive" form of processing in adulthood. In a similar vein, Neuroconstructivism offers a more developmental view that is resonant with the sensorimotor hypothesis. Thus, Emergentist theories, such as Neural Reuse, would be further bolstered by the incorporation of a developmental view such as that proposed by Neuroemergentism.

The proponents of the various perspectives presented here have all sought to tackle higher-level functions using an Emergentist approach. At the outset of this discussion, several analogies were offered including the metaphors of D'Arcy Thompson (Thompson, 1917) and the extensions by Elizabeth Bates and colleagues (Bates et al., 1979). Face perception was also presented as an example of a Neuroemergentist approach to questions in cognitive psychology. As noted in the introduction, the notion of a fusiform face area, has been the center of discussion in the adult neuropsychological and neuroimaging literature for many years (Kanwisher et al., 1997, 1999). The interpretation by some is that this area is dedicated purely to face processing. However, work by Gauthier and colleagues suggests that it is more likely to be considered a product of perceptual expertise (Gauthier et al., 1999, 2000, 2014).

Face processing as an example of Neuroemergentism can be seen in the work of Johnson, Morton and Colleagues (Johnson et al., 1991; Morton & Johnson, 1991). Their two-process theory of face processing was based on a model of chick imprinting from Horn which identified two areas involved in imprinting (Horn, 1981, 1986; Horn, Nicol, & Brown, 2001). The first area, the optic tectum, which is equivalent to the superior colliculus in humans, draws the newborn visual system to the neck area of the hen. Since the mother hen is very likely to be the first moving object the chick observes, this leads to focusing attention on it. The second system, which is part of the forebrain (equivalent to the cerebral cortex in humans), the Intermediate and medial part of the Mesopallium, is involved in linking visual objects with motor responses. This system will lead to imprinting. Morton and Johnson expanded on this view by proposing two systems in humans. Conspec initially draws attention to face-like configurations, and Conlearn, serves as a trainer for the lower level system. Computational modeling of the process has also led to novel predictions (O'Reilly & Johnson, 1994). Face perception serves as a model for other cognitive domains, especially in light of the use of a neurobiologically plausible computational model.

It is also important to note that Johnson has proposed the interactive specialization view which is also highly resonant with the Neuroemergentist approach discussed here (Johnson, 2011). In this view, the specialization that occurs across time is due to a pattern of competition across brain regions. Areas of the brain will have initial biases that are dependent on interconnections with other areas of the brain. Some areas may begin with poorly defined functions but become specialized across time as they take on other functions. In the case of face perception, Conspec draws attention to face-like configurations while other areas in the cortex such as Conlearn help to further refine its function. This competitive process works itself out across development. This notion of interactive specialization is one that is resonant with the Neuroemergentist view that is outlined here.

Although a Neuroemergentist approach can be used for many different cognitive functions, it may be particularly well suited for higher-level cognitive functions that develop slowly over time. This view is one that was championed by both Karmiloff-Smith using Neuroconstructivism and Bates in her cognitive neurodevelopmental view of language as "a new machine built out of old parts." Recent work, in the bilingual literature, has begun to appear that might also be well suited for a Neuroemergentist approach. In the following section, we will briefly consider how this approach might help inform a currently hotly debated topic in the literature, the cognitive advantages that arise because of the use of two languages.

## 7. Neuroemergentism and the bilingual cognitive advantage

Recent controversies have erupted over the presence or absence of a bilingual cognitive advantage (Abutalebi et al., 2012; Anton et al., 2014; de Bruin, Treccani, & Della Sala, 2015; Duñabeitia & Carreiras; Dunabeitia et al., 2014; Hernández, Martin, Barceló, & Costa, 2013; Hilchey & Klein, 2011; Paap & Greenberg, 2013; Paap, Johnson, & Sawi, 2014; Paap & Sawi, 2014). Whereas the discussion over its presence or absence continues, there is still considerable debate over a possible cause for this putative advantage (Costa, Pannunzi, Deco, & Pickering, 2016). The notion of cognitive differences has also been expanded to views of the brain bases of bilingualism (García-Pentón, Fernández García, Costello, Duñabeitia, & Carreiras, 2015; Li et al., 2014). For example, the adaptive control hypothesis posits that the use of two languages engages a general cognitive control mechanism to a greater extent than the use of one language (Green & Abutalebi, 2013). In a similar vein, the bilingual brain training model of Stocco, Prat and colleagues posits that the use of two languages results in neural adaptations within the basal ganglia (Stocco, Lebiere, O'Reilly, & Anderson, 2012; Stocco & Prat, 2014; Stocco, Yamasaki, Natalenko, & Prat, 2014). In this framework, bilinguals are particularly adept at learning new rules resulting in differences in brain activity in the basal ganglia relative to monolinguals (Stocco & Prat, 2014). Furthermore, the framework derives its predictions from a neurocomputational model of the basal ganglia (Stocco et al., 2012, 2014). However, neither model takes a developmental view or incorporates data from developmental populations.

What mechanism could potentially lead to alterations of the basal ganglia across development? Work by Krizman and Marian (2015) suggests that bilingualism involves the use of two phonologies which helps to tune the basal ganglia via the auditory cortex. Given that early bilingualism occurs at a period of greater plasticity, it is likely that differences between bilinguals and monolinguals will involve the basal ganglia. This view is resonant with work by Petitto, Kovelman and colleagues who propose a perceptual wedge hypothesis, where exposure to two languages early in life would lead to additional flexibility with regard to the recognition of non-native phonemes (Petitto et al., 2012).

The flexible nature of processing is also considered by Li and colleagues who suggest that bilingualism can be viewed as a non-

linear dynamical system in which language experience affects the brain (Li et al., 2014). In fact, Li et al. (2014) highlight the ways in which bilingual and second language experience are resonant with the dynamic Emergentist view espoused by Bates (1999). In a similar vein, Hernandez (2013) proposes that "two languages live inside one brain almost as two species live in an ecosystem. For the most part they peacefully coexist and often share resources. But they also compete for resources especially when under stress, as occurs when there is brain damage." (p. 12) The metaphor of two species in an ecosystem adds a biological and evolutionary perspective to bilingualism. This approach is similar to the one used by Elizabeth Bates when discussing adaptation and the Giraffe's neck as an analogy for understanding how language is represented in the brain.

The notion that learning two languages has a non-linear dynamical flavor to it is very much in line with a Neuroemergentist view as proposed in the present review. One way forward would be to incorporate more neurocomputational models that account for changes across development. This developmental approach would extend the view proposed by Hernandez (2013) that languages can compete for and share resources by pointing out how this changes across time. For example, early in development competition in bilinguals might involve phonological information as proposed by Krizman and Marian. Later in development the competition might come from higher-levels of language (i.e. semantics, syntax, etc.) and hence involve abstract rules as proposed by Stocco and Prat.

Another way in which the field could move forward is by looking at individual differences. The role of genetics has also only recently come to the fore as a potential explanatory variable for both the acquisition of a second language (Mamiya, Richards, Coe, Eichler, & Kuhl, 2016; Wong, Morgan-Short, Ettlinger, & Zheng, 2012) and the use of cognitive and language control in bilinguals (Hernandez, Greene, Vaughn, Francis, & Grigorenko, 2015; Vaughn et al., 2016). The use of a more neurobiological and computational approach would help researchers to more precisely capture the particular factors of bilingual use across development that may or may not contribute to changes in cognitive control. Of course, it is also quite possible that this effect is non-linear and that using two languages changes the brain but that individual differences including genetic status may also render some people more amenable to the changes that occur because of the extended use of two languages (Hernandez et al., 2015; Mamiya et al., 2016). In short, we propose Neuroemergentism as a way to refine the theories with regard to the brain bases of bilingualism as well as the particular effects that these experiences might have on cognitive control.

## 8. Conclusion

The notion that cognition emerges due to a dynamic interplay between the individual and the environment over a long period of development has a long history in the cognitive and neuroscience literature. Recent theoretical frameworks reviewed here, Neuronal Recycling, Neural Reuse, and Language as Shaped by the Brain, are resonant with this view. Neuroconstructivism adds development to the argument. Taken together, aspects of each of these four frameworks lead to a synthesis which we term Neurocomputational Emergentism or Neuroemergentism for short. The incorporation of neurally-based modeling as well as neuroimaging and genetics has the potential to transform various subfields within cognitive neuroscience.

Finally, Neuroemergentism represents an extension of an Emergentist framework as proposed by Elizabeth Bates, John Stuart Mill and many others. As noted earlier, Bates suggests that a Giraffe's neck which is adapted to eating leaves up high leads to a number of changes across the entire organism. In a similar vein, Hernandez suggests an ecological view for the competition and cooperation of two species-like language systems that interact across time. These evolutionary analogies lend themselves to more biologically- and neurally-based approaches. It is worth ending on the note that face processing is one of the few phenomena that approximates a Neuroemergentist account. We propose that the adoption of this level of description across a wider set of domains is paramount in order to move the study of the neural bases of language and cognition forward.

## Acknowledgements

This research was supported in part by grant # 5R03HD079873-02, Effects of genetic differences and bilingual status on cognitive control, to AEH.

## Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jneuroling.2017.12.010.

## References

Abutalebi, J., Della Rosa, P. A., Green, D. W., Hernandez, M., Scifo, P., Keim, R., ... Costa, A. (2012). Bilingualism tunes the anterior cingulate cortex for conflict monitoring. *Cereb Cortex, 22*(9), 2076–2086. http://dx.doi.org/10.1093/cercor/bhr287.

Anderson, M. L. (2008). Circuit sharing and the implementation of intelligent systems. *Connection Science, 20*(4), 239–251.

Anderson, M. L. (2010). Neural reuse: A fundamental organizational principle of the brain. *Behavioral and Brain Sciences, 33*, 245–266. http://dx.doi.org/10.1017/S0140525X10000853 discussion 266-313.

Anderson, M. L. (2016). Neural reuse in the organization and development of the brain. *Developmental Medicine and Child Neurology, 58*, 3–6. http://dx.doi.org/10.1111/dmcn.13039.

Anderson, M. L., & Penner-Wilger, M. (2013). Neural reuse in the evolution and development of the brain: Evidence for developmental homology? *Developmental Psychobiology, 55*, 42–51. http://dx.doi.org/10.1002/dev.21055.

Anton, E., Dunabeitia, J. A., Estevez, A., Hernandez, J. A., Castillo, A., Fuentes, L. J., ... Carreiras, M. (2014). Is there a bilingual advantage in the ANT task? Evidence from children. *Fronters in Psychology, 5*, 398. http://dx.doi.org/10.3389/fpsyg.2014.00398.

Bates, E. A. (1999). Plasticity, localization, and language development. In S. H. Broman, J. M. Fletcher, S. H. Broman, & J. M. Fletcher (Eds.). *The changing nervous*

*system: Neurobehavioral consequences of early brain disorders* (pp. 214–253). New York, NY, US: Oxford University Press.
Bates, E. A., Benigni, L., Bretherton, I., Camaioni, L., & Volterra, V. (1979). *The emergence of symbols: Cognition and communication in infancy.* New York, NY: Academic Press.
Bates, E. A., & Goodman, J. C. (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia, and real-time processing. *Language & Cognitive Processes, 12*(5–6), 507–584.
Brannon, E. M., & Terrace, H. S. (1998). Ordering of the numerosities 1 to 9 by monkeys. *Science, 282*(5389), 746–749.
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), 1037–1046.
de Bruin, A., Treccani, B., & Della Sala, S. (2015). Cognitive advantage in bilingualism: An example of publication bias? *Psychological Science, 26*(1), 99–107. http://dx.doi.org/10.1177/0956797614557866.
Butterworth, B. (2010). Foundational numerical capacities and the origins of dyscalculia. *Trends in Cognitive Sciences, 14*(12), 534–541. http://dx.doi.org/10.1016/j.tics.2010.09.007.
Christiansen, M. H. (1994). *Infinite languages, finite minds: Connectionism, learning and linguistic structure.* Unpublished PhD thesisScotland: University of Edinburgh.
Christiansen, M. H., & Chater, N. (2008a). Brains, genes, and language evolution: A new synthesis. *Behavioral and Brain Sciences, 31*(5), 537–558. http://dx.doi.org/10.1017/S0140525x08005281.
Christiansen, M. H., & Chater, N. (2008b). Language as shaped by the brain. *Behavioral and Brain Sciences, 31*(5), http://dx.doi.org/10.1017/S0140525x08004998 489-+.
Christiansen, M. H., & Chater, N. (2016a). *Creating language: Integrating evolution, acquisition, and processing.* Cambridge, MA: MIT Press.
Christiansen, M. H., & Chater, N. (2016b). The Now-or-Never bottleneck: A fundamental constraint on language. *Behavioral & Brain Sciences, 39*, e62.
Christophel, T. B., Klink, P. C., Spitzer, B., Roelfsema, P. R., & Haynes, J.-D. (2017). The distributed nature of working memory. *Trends in Cognitive Sciences, 21*, 111–124. http://dx.doi.org/10.1016/j.tics.2016.12.007.
Costa, A., Pannunzi, M., Deco, G., & Pickering, M. J. (2016). Do bilinguals automatically activate their native language when they are not using it? *Cognitive Science*. http://dx.doi.org/10.1111/cogs.12434.
Curby, K. M., & Gauthier, I. (2014). Interference between face and non-face domains of perceptual expertise: A replication and extension. *Frontiers in Psychology, 5*(955)http://doi.org/10.3389/fpsyg.2014.00955.
D'Souza, D., & Karmiloff-Smith, A. (2016). Why a developmental perspective is critical for understanding human cognition. *Behavioral and Brain Sciences, 39*, e122. http://dx.doi.org/10.1017/S0140525X15001569.
Darwin, C. (1871). *The descent of man, and selection in relation to sex.* London, United Kingdom: John Murray.
Deacon, T. W. (1997). *The symbolic species: The co-evolution of language and the human brain.* London, United Kingdom: Penguin Press.
Dehaene, S. (1997). *The number sense: How mathematical knowledge is embedded in our brains.* New York, NY: Oxford University Press.
Dehaene, S., & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron, 56*(2), 384–398. http://dx.doi.org/10.1016/j.neuron.2007.10.004.
Dehaene, S., & Dehaene-Lambertz, G. (2016). Is the brain prewired for letters? *Nature Neuroscience, 19*(9), 1192–1193. http://dx.doi.org/10.1038/nn.4369.
Dehaene, S., & Fyssen, S. (2005). *From monkey brain to human brain: A fyssen foundation symposium.* Cambridge, MA: MIT Press.
Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Nunes Filho, G., Jobert, A., ... Cohen, L. (2010). How learning to read changes the cortical networks for vision and language. *Science, 330*(6009), 1359–1364. http://dx.doi.org/10.1126/science.1194140.
Dekker, T. M., & Karmiloff-Smith, A. (2010). The importance of ontogenetic change in typical and atypical development. *Behavioral and Brain Sciences, 33*(4), 271–272. http://dx.doi.org/10.1017/S0140525X10001019.
Demeyere, N., Lestou, V., & Humphreys, G. W. (2010). Neuropsychological evidence for a dissociation in counting and subitizing. *Neurocase, 16*(3), 219–237. http://dx.doi.org/10.1080/13554790903405719.
Donnai, D., & Karmiloff-Smith, A. (2000). Williams syndrome: from genotype through to the cognitive phenotype. *American Journal of Medical Genetics, 97*(2), 164–171.
Duñabeitia, J. A., Carreiras, M.. The bilingual advantage: Acta est fabula?. Cortex. doi: https://doi.org/10.1016/j.cortex.2015.06.009.
Dunabeitia, J. A., Hernandez, J. A., Anton, E., Macizo, P., Estevez, A., Fuentes, L. J., et al. (2014). The inhibitory advantage in bilingual children revisited. *Experimental Psychology, 61*(3), 234–251. http://dx.doi.org/10.1027/1618-3169/a000243.
Eisenberg, D. P., Jabbi, M., & Berman, K. F. (2010). Bridging the gene-behavior divide through neuroimaging deletion syndromes: Velocardiofacial (22q11.2 Deletion) and Williams (7q11.23 Deletion) syndromes. *NeuroImage, 53*(3), 857–869. http://dx.doi.org/10.1016/j.neuroimage.2010.02.070.
Elman, J. L., Bates, E. A., Johnson, M. H., & Karmiloff-Smith, A. (1996). *Rethinking innateness: A connectionist perspective on development.* Cambridge, MA: MIT Press.
García-Pentón, L., Fernández García, Y., Costello, B., Duñabeitia, J. A., & Carreiras, M. (2015). The neuroanatomy of bilingualism: How to turn a hazy view into the full picture. *Language, Cognition and Neuroscience,* 1–25. http://dx.doi.org/10.1080/23273798.2015.1068944.
Gauthier, I., McGugin, R. W., Richler, J. J., Herzmann, G., Speegle, M., & Van Gulick, A. E. (2014). Experience moderates overlap between object and face recognition, suggesting a common ability. *Journal of Vision, 14*(8), 7. http://dx.doi.org/10.1167/14.8.7.
Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience, 3*(2), 191–197. http://dx.doi.org/10.1038/72140.
Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. *Nature Neuroscience, 2*(6), 568–573. http://dx.doi.org/10.1038/9224.
Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language.* New York, NY: Oxford University Press.
Gould, S. J., & Vrba, E. S. (1982). Exaptation - a missing term in the science of form. *Paleobiology, 8*(1), 4–15.
Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology, 25*(5), 515–530. http://dx.doi.org/10.1080/20445911.2013.796377.
de Haan, M., Humphreys, K., & Johnson, M. H. (2002). Developing a brain specialized for face perception: A converging methods approach. *Developmental Psychobiology, 40*(3), 200–212.
Hernandez, A. E. (2013). *The bilingual brain.* New York, NY: Oxford University Press.
Hernandez, A. E., Greene, M. R., Vaughn, K. A., Francis, D. J., & Grigorenko, E. L. (2015). Beyond the bilingual advantage: The potential role of genes and environment on the development of cognitive control. *Journal of Neurolinguistics, 35*, 109–119. http://dx.doi.org/10.1016/j.jneuroling.2015.04.002.
Hernandez, A. E., & Li, P. (2007). Age of acquisition: Its neural and computational mechanisms. *Psychological Bulletin, 133*(4), 638–650.
Hernandez, A. E., Li, P., & MacWhinney, B. (2005). The emergence of competing modules in bilingualism. *Trends in Cognitive Sciences, 9*, 220–225.
Hernández, M., Martin, C. D., Barceló, F., & Costa, A. (2013). Where is the bilingual advantage in task-switching? *Journal of Memory and Language, 69*(3), 257–276. http://dx.doi.org/10.1016/j.jml.2013.06.004.
Hilchey, M., & Klein, R. (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychonomic Bulletin & Review, 18*(4), 625–658.
Horn, G. (1981). Neural mechanisms of learning: An analysis of imprinting in the domestic chick. Proceedings of the royal society of London. Series B. *Biological Sciences, 213*(1191), 101–137.
Horn, G. (1986). Imprinting, learning, and memory. *Behavioral Neuroscience, 100*(6), 825–832.
Horn, G., Nicol, A. U., & Brown, M. W. (2001). Tracking memory's trace. *Proceedings of the National Academy of Sciences, 98*(9), 5282–5287. http://dx.doi.org/10.1073/pnas.091094798.
Izard, V., Dehaene-Lambertz, G., & Dehaene, S. (2008). Distinct cerebral pathways for object identity and number in human infants. *PLoS Biology, 6*(2), e11. http://dx.doi.org/10.1371/journal.pbio.0060011.
Johnson, M. H. (2011). Interactive specialization: A domain-general framework for human functional brain development? *Developmental Cognitive Neuroscience, 1*(1), 7–21. http://dx.doi.org/10.1016/j.dcn.2010.07.003.

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition, 40*(1–2), 1–19.
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.
Kanwisher, N., Stanley, D., & Harris, A. (1999). The fusiform face area is selective for faces not animals. *NeuroReport, 10*(1), 183–187.
Kanwisher, N., & Yovel, G. (2006). The fusiform face area: A cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society B: Biological Sciences, 361*(1476), 2109–2128. http://dx.doi.org/10.1098/rstb.2006.1934.
Karmiloff-Smith, A. (2006). The tortuous route from genes to behavior: A neuroconstructivist approach. *Cognitive, Affective, & Behavioral Neuroscience, 6*(1), 9–17.
Karmiloff-Smith, A. (2009). Nativism versus neuroconstructivism: Rethinking the study of developmental disorders. *Developmental Psychology, 45*(1), 56–63. http://dx.doi.org/10.1037/a0014506.
Karmiloff-Smith, A., D'Souza, D., Dekker, T. M., Van Herwegen, J., Xu, F., Rodic, M., et al. (2012). Genetic and environmental vulnerabilities in children with neurodevelopmental disorders. *Proceedings of the National Academy of Sciences, 109*(Suppl 2), 17261–17265. http://dx.doi.org/10.1073/pnas.1121087109.
Karmiloff-Smith, A., Thomas, M., Annaz, D., Humphreys, K., Ewing, S., Brace, N., ... Campbell, R. (2004). Exploring the Williams syndrome face-processing debate: The importance of building developmental trajectories. *Journal of Child Psychological Psychiatry, 45*(7), 1258–1274. http://dx.doi.org/10.1111/j.1469-7610.2004.00322.x.
Krizman, J., & Marian, V. (2015). Neural consequences of bilingualism for cortical and subcortical function. In J. W. Schwieter (Ed.). *The cambridge handbook of bilingual processing*. Cambridge, United Kingdom: Cambridge University Press.
Li, P., Legault, J., & Litcofsky, K. A. (2014). Neuroplasticity as a function of second language learning: Anatomical changes in the human brain. *Cortex, 58*, 301–324. http://dx.doi.org/10.1016/j.cortex.2014.05.001.
Lipton, J. S., & Spelke, E. S. (2003). Origins of number sense. Large-number discrimination in human infants. *Psychological Science, 14*(5), 396–401. http://dx.doi.org/10.1111/1467-9280.01453.
Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. *Nature, 453*(7197), 869–878. http://www.nature.com/nature/journal/v453/n7197/suppinfo/nature06976_S1.html.
Mamiya, P. C., Richards, T. L., Coe, B. P., Eichler, E. E., & Kuhl, P. K. (2016). Brain white matter structure and COMT gene are linked to second-language learning in adults. *Proceedings of the National Academy of Sciences, 113*(26), 7249–7254. http://dx.doi.org/10.1073/pnas.1606602113.
Mervis, C. B., Robinson, B. F., Bertrand, J., Morris, C. A., Klein-Tasman, B. P., & Armstrong, S. C. (2000). The Williams syndrome cognitive profile. *Brain and Cognition, 44*(3), 604–628. http://dx.doi.org/10.1006/brcg.2000.1232.
Miller, E. K. (2013). The “working” of working memory. *Dialogues in Clinical Neuroscience, 15*(4), 411–418.
Mill, J. S., Robson, J. M., McRae, R. F., Routledge, Kegan, P, & University of Toronto, P (1974). *Collected works of John Stuart Mill. 7, 7*. Toronto, Canada: University of Toronto Press (London, United Kingdom : Routledge & Kegan Paul).
Morton, J., & Johnson, M. H. (1991). CONSPEC and CONLERN: A two-process theory of infant face recognition. *Psychological Review, 98*(2), 164–181.
O'Reilly, R. C., & Johnson, M. H. (1994). Object recognition and sensitive periods: A computational analysis of visual imprinting. *Neural Computation, 6*(3), 357–389. http://dx.doi.org/10.1162/neco.1994.6.3.357.
Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology, 66*(2), 232–258. http://dx.doi.org/10.1016/j.cogpsych.2012.12.002.
Paap, K. R., Johnson, H. A., & Sawi, O. (2014). Are bilingual advantages dependent upon specific tasks or specific bilingual experiences? *Journal of Cognitive Psychology, 26*(6), 615–639. http://dx.doi.org/10.1080/20445911.2014.944914.
Paap, K. R., & Sawi, O. (2014). Bilingual advantages in executive functioning: Problems in convergent validity, discriminant validity, and the identification of the theoretical constructs. *Frontiers in Psychology, 5*, 962. http://dx.doi.org/10.3389/fpsyg.2014.00962.
Paterson, S. J., Brown, J. H., Gsodl, M. K., Johnson, M. H., & Karmiloff-Smith, A. (1999). Cognitive modularity and genetic disorders. *Science, 286*(5448), 2355–2358.
Petitto, L. A., Berens, M. S., Kovelman, I., Dubins, M. H., Jasinska, K., & Shalinsky, M. (2012). The "Perceptual Wedge Hypothesis" as the basis for bilingual babies' phonetic processing advantage: New insights from fNIRS brain imaging. *Brain and Language, 121*(2), 130–143.
Saygin, Z. M., Osher, D. E., Norton, E. S., Youssoufian, D. A., Beach, S. D., Feather, J., ... Kanwisher, N. (2016). Connectivity precedes function in the development of the visual word form area. *Nature Neuroscience, 19*(9), 1250–1255. http://dx.doi.org/10.1038/nn.4354.
Sergent, J., Ohta, S., & Macdonald, B. (1992). Functional neuroanatomy of face and object processing: A positron emission tomography study. *Brain, 115*(1), 15–36. http://dx.doi.org/10.1093/brain/115.1.15.
Stocco, A., Lebiere, C., O'Reilly, R. C., & Anderson, J. R. (2012). Distinct contributions of the caudate nucleus, rostral prefrontal cortex, and parietal cortex to the execution of instructed tasks. *Cognitive, Affective, & Behavioral Neuroscience, 12*(4), 611–628. http://dx.doi.org/10.3758/s13415-012-0117-7.
Stocco, A., & Prat, C. S. (2014). Bilingualism trains specific brain circuits involved in flexible rule selection and application. *Brain and Language, 137*, 50–61. http://dx.doi.org/10.1016/j.bandl.2014.07.005.
Stocco, A., Yamasaki, B., Natalenko, R., & Prat, C. S. (2014). Bilingual brain training: A neurobiological framework of how bilingual experience improves executive function. *International Journal of Bilingualism, 18*(1), 67–92. http://dx.doi.org/10.1177/1367006912456617.
Thompson, D. A. W. (1917). *On growth and form*. Cambridge, United Kingdom: University Press.
Tomasello, M. (2005). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
Van Herwegen, J., Ansari, D., Xu, F., & Karmiloff-Smith, A. (2008). Small and large number processing in infants and toddlers with Williams syndrome. *Developmental Science, 11*(5), 637–643. http://dx.doi.org/10.1111/j.1467-7687.2008.00711.x.
Vaughn, K. A., Ramos Nunez, A. I., Greene, M. R., Munson, B. A., Grigorenko, E. L., & Hernandez, A. E. (2016). Individual differences in the bilingual brain: The role of language background and DRD2 genotype in verbal and non-verbal cognitive control. *Journal of Neurolinguistics, 40*, 112–127. http://dx.doi.org/10.1016/j.jneuroling.2016.06.008.
Wong, P. C. M., Morgan-Short, K., Ettlinger, M., & Zheng, J. (2012). Linking neurogenetics and individual differences in language learning: The dopamine hypothesis. *Cortex, 48*(9), 1091–1102. https://doi.org/10.1016/j.cortex.2012.03.017.
Xu, F. (2003). Numerosity discrimination in infants: Evidence for two systems of representations. *Cognition, 89*(1), B15–B25. https://doi.org/10.1016/S0010-0277(03)00050-7.