

# COMPONENTS OF THE LANGUAGE-READY BRAIN

EDITED BY: Cedric Boeckx and Antonio Benítez-Burraco  
PUBLISHED IN: Frontiers in Psychology



# frontiers

## Frontiers Copyright Statement

© Copyright 2007-2016 Frontiers Media SA. All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

ISSN 1664-8714

ISBN 978-2-88919-893-1

DOI 10.3389/978-2-88919-893-1

## About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

## Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

## Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.

By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

## What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: [researchtopics@frontiersin.org](mailto:researchtopics@frontiersin.org)

# COMPONENTS OF THE LANGUAGE-READY BRAIN

Topic Editors:

**Cedric Boeckx**, University of Barcelona & Catalan Institute for Research and Advanced Studies (ICREA), Spain

**Antonio Benítez-Burraco**, University of Huelva, Spain

This volume highlights new avenues of research in the language sciences, and particularly, in the neurobiology of language. The term “language-ready brain” stresses, on the one hand, the importance of a brain-based description of our species’ linguistic capacity, and, on the other, the need to appreciate the crucial role culture plays in shaping the linguistic systems children acquire and adults use. For this reason, the focus is not put on language per se, but on our learning biases and cognitive pre-dispositions toward language. Both brain and culture are considered at two crucial levels of inquiry: phylogeny and ontogeny. In a fast-growing field like the language sciences and specifically, language evolution studies, this book has tried to capture several of the most exciting topics explored currently, sowing seeds for future investigations.

**Citation:** Boeckx, C., Benítez-Burraco, A., eds. (2016). Components of the Language-Ready Brain. Lausanne: Frontiers Media. doi: 10.3389/978-2-88919-893-1

# Table of Contents

## **04 Editorial: Components of the Language-Ready Brain**

Cedric Boeckx and Antonio Benítez-Burraco

## **Section 1: Genetics**

### **05 Retinoic Acid Signaling: A New Piece in the Spoken Language Puzzle**

Jon-Ruben van Rhijn and Sonja C. Vernes

## **Section 2: The brain**

### **13 Merge in the Human Brain: A Sub-Region Based Functional Investigation in the Left Pars Opercularis**

Emiliano Zaccarella and Angela D. Friederici

### **22 'Syntactic Perturbation' During Production Activates the Right IFG, but not Broca's Area or the ATL**

William Matchin and Gregory Hickok

### **36 Brain asymmetry in the white matter making and globularity**

Constantina Theofanopoulou

### **49 Corrigendum: Brain asymmetry in the white matter making and globularity**

Constantina Theofanopoulou

### **50 The brain dynamics of linguistic computation**

Elliot Murphy

## **Section 3: Development**

### **69 The "Globularization Hypothesis" of the Language-ready Brain as a Developmental Frame for Prosodic Bootstrapping Theories of Language Acquisition**

Aritz Irurtzun

### **76 Temporal Attention as a Scaffold for Language Development**

Ruth de Diego-Balaguer, Anna Martinez-Alvarez and Ferran Pons

## **Section 4: Evolution**

### **91 Can a bird brain do phonology?**

Bridget D. Samuels

### **101 Evolution of speech-specific cognitive adaptations**

Bart de Boer

### **105 Why language really is not a communication system: a cognitive view of language evolution**

Anne C. Reboul

### **117 Brain readiness and the nature of language**

Denis Bouchard





# Editorial: Components of the Language-Ready Brain

Cedric Boeckx<sup>1\*</sup> and Antonio Benítez-Burraco<sup>2</sup>

<sup>1</sup> Catalan Institute for Research and Advanced Studies (ICREA)/Department of General Linguistics, Universitat de Barcelona, Barcelona, Spain, <sup>2</sup> Department of Philology and its Didactics, University of Huelva, Huelva, Spain

**Keywords:** language, language development, evolution, neurolinguistics, evolutionary biology

## The Editorial on the Research Topic

### Components of the Language-Ready Brain

Our intention in putting together this volume was to exemplify and highlight new avenues of research in the language sciences concerning the neurobiology of language. We chose the term “language-ready brain” for our Research Topic, like we did for Boeckx and Benítez-Burraco, because we think it is high time to stress, on the one hand, the importance of a brain-based description of our species’ linguistic capacity, and, on the other, the need to appreciate the crucial role culture plays in shaping the linguistic systems children acquire and adults use. In this sense, the focus of neurobiological investigations should not be “language,” but our learning biases and cognitive pre-dispositions toward language (i.e., “language-readiness”). Both brain and culture considerations ought to shape research at all levels of inquiry: phylogeny and ontogeny.

The contributions to this research topic break new grounds, by either revisiting long-standing issues (such as the role of Broca’s region, the relevance of lateralization, the evolutionary origins of phonology, the role of basic cognitive and perceptive abilities in language acquisition, or the functions performed by language), or by examining closely issues that we are sure will rise to prominence in the near future (like the translational models of language processing into specific patterns of brain oscillations or the nature of the gene networks in which known “language genes” are found integrated). Taken together, the papers collected here shed light on language at the level of the genetics (van Rhijn and Vernes), brain connectivity (Murphy; Theofanopoulou), and physiology (Matchin and Hickok; Zaccarella and Friederici), cognition (de Boer; de Diego-Balaguer et al.), and behavior (Bouchard; Irurtzun; Reboul; Samuels).

In a fast-growing field like the language sciences, Research topics cannot hope to capture all relevant aspects of the field, but we hope that the present volume offers a snapshot that some of the most exciting research taking place today, sowing seeds for future investigations.

## OPEN ACCESS

### Edited and reviewed by:

Manuel Carreiras,  
Basque Center on Cognition, Brain  
and Language, Spain

### \*Correspondence:

Cedric Boeckx  
cedric.boeckx@ub.edu

### Specialty section:

This article was submitted to  
Language Sciences,  
a section of the journal  
Frontiers in Psychology

**Received:** 25 April 2016

**Accepted:** 06 May 2016

**Published:** 23 May 2016

### Citation:

Boeckx C and Benítez-Burraco A  
(2016) Editorial: Components of the  
Language-Ready Brain.  
Front. Psychol. 7:762.  
doi: 10.3389/fpsyg.2016.00762

## AUTHOR CONTRIBUTIONS

Both authors wrote the editorial.

## FUNDING

Preparation of this work was supported by funds from the Spanish Ministry of Economy and Competitiveness (grants FFI2013-43823-P and FFI2014-61888-EXP), as well as funds from the Generalitat de Catalunya (2014-SGR-200).

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Boeckx and Benítez-Burraco. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Retinoic Acid Signaling: A New Piece in the Spoken Language Puzzle

Jon-Ruben van Rhijn<sup>1,2</sup> and Sonja C. Vernes<sup>1,3\*</sup>

<sup>1</sup> Department of Language and Genetics, Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands, <sup>2</sup> Molecular Neurophysiology Group, Department of Cognitive Neuroscience, Radboud University Medical Center, Nijmegen, Netherlands, <sup>3</sup> Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, Netherlands

Speech requires precise motor control and rapid sequencing of highly complex vocal musculature. Despite its complexity, most people produce spoken language effortlessly. This is due to activity in distributed neuronal circuitry including cortico-striato-thalamic loops that control speech-motor output. Understanding the neuro-genetic mechanisms involved in the correct development and function of these pathways will shed light on how humans can effortlessly and innately use spoken language and help to elucidate what goes wrong in speech-language disorders. *FOXP2* was the first single gene identified to cause speech and language disorder. Individuals with *FOXP2* mutations display a severe speech deficit that includes receptive and expressive language impairments. The neuro-molecular mechanisms controlled by *FOXP2* will give insight into our capacity for speech-motor control, but are only beginning to be unraveled. Recently *FOXP2* was found to regulate genes involved in retinoic acid (RA) signaling and to modify the cellular response to RA, a key regulator of brain development. Here we explore evidence that *FOXP2* and RA function in overlapping pathways. We summate evidence at molecular, cellular, and behavioral levels that suggest an interplay between *FOXP2* and RA that may be important for fine motor control and speech-motor output. We propose RA signaling is an exciting new angle from which to investigate how neuro-genetic mechanisms can contribute to the (spoken) language ready brain.

**Keywords:** retinoic acid, FoxP2, synaptic plasticity, development, motor skills, striatum, dopamine receptor

## OPEN ACCESS

### Edited by:

Antonio Benítez-Burraco,  
University of Huelva, Spain

### Reviewed by:

Constance Scharff,  
Freie Universität Berlin, Germany  
Tomokazu Tomo Fukuda,  
Tohoku University, Japan

Christina Roeske contributed  
to the review of Constance Scharff

### \*Correspondence:

Sonja C. Vernes  
sonja.vernes@mpi.nl

### Specialty section:

This article was submitted to  
Language Sciences,  
a section of the journal  
Frontiers in Psychology

**Received:** 01 September 2015

**Accepted:** 10 November 2015

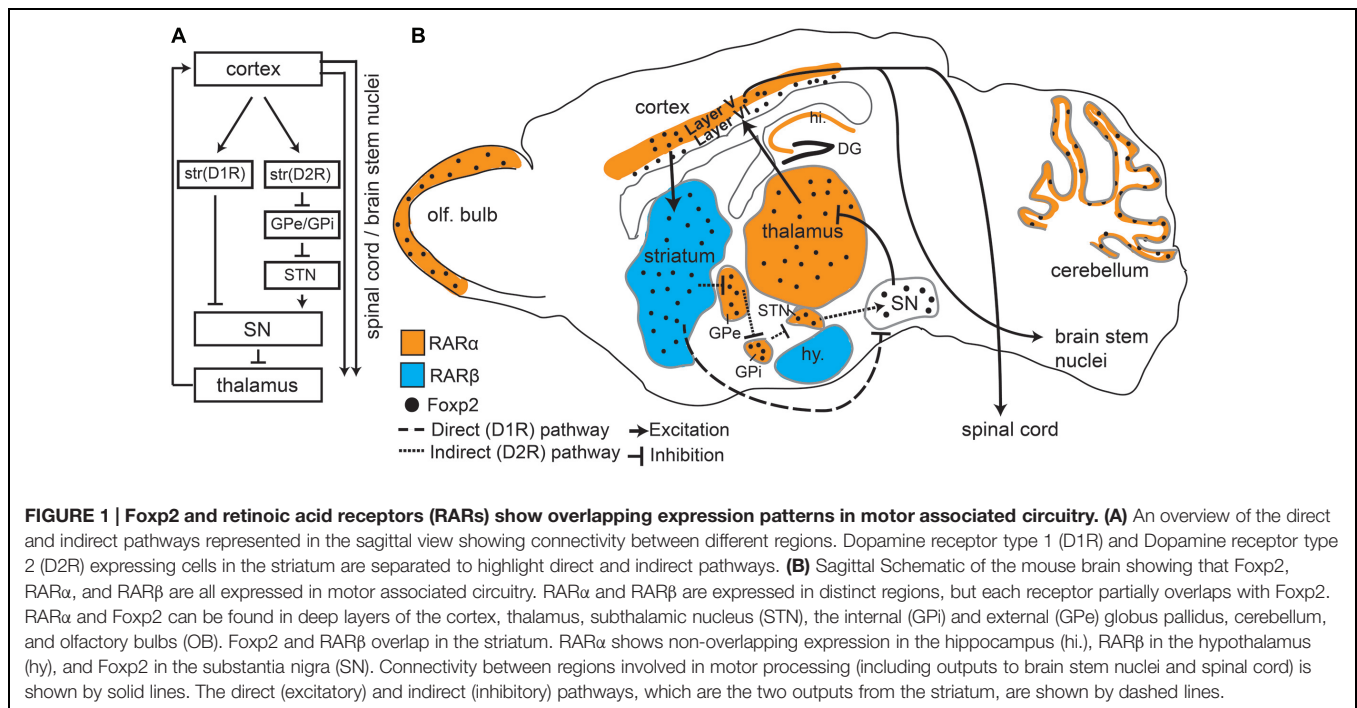
**Published:** 26 November 2015

### Citation:

van Rhijn JR and Vernes SC (2015)  
Retinoic Acid Signaling: A New Piece  
in the Spoken Language Puzzle.  
*Front. Psychol.* 6:1816.  
doi: 10.3389/fpsyg.2015.01816

## SPEECH AND SPOKEN LANGUAGE

Speech is the primary modality by which humans use language, and human orofacial morphology is uniquely suited to the production of intricate vocalizations needed for spoken language (Lieberman, 2007). The orofacial musculature is one of the most complex muscle systems in the body and in order to successfully produce meaningful speech these muscles must be controlled and coordinated in rapid sequences involving distributed neuronal circuitry. This motor activity is generated in several neural loops that select appropriate actions and generate the necessary motor patterns. One crucial circuit, the cortico-basal ganglia loop, sends activity from the motor cortex to the striatum (a component of the basal ganglia) where activity is integrated. Subsequently, outputs from here modulate activity in several thalamic nuclei. Activity from the thalamus is then sent back to the motor cortex, where a specialized population of output neurons organizes the complex thalamocortical inputs (Kravitz and Kreitzer, 2012; Calabresi et al., 2014). These cortical output neurons send the information, via the pyramidal tract, to motor neurons directly controlling muscle tissue. These neurons are either located in the spinal cord (controlling limb and body movements), or in the brainstem's cranial nerve nuclei



(controlling facial and vocal tract movements). An illustration of the cortico-basal ganglia loop (in the rodent brain) is given in **Figure 1A**. Proper connectivity within this pathway is necessary to enable the precise outputs needed for orofacial muscle control.

The striatum can be seen as a central hub within the motor pathway, making it one of the most intriguing regions in which to investigate properties of motor circuitry and orofacial control. Striatal activity is especially important for fine motor behavior and motor skill learning (Doyon et al., 2003) and cortical and subcortical circuitry, including the striatum, has been established as highly important for speech–motor control (Lieberman, 2002). Furthermore, increased activation of the basal ganglia (which incorporates the striatum) has been shown via functional brain imaging (fMRI) in specific speech–motor language tasks (Wildgruber et al., 2001; Booth et al., 2007). Lastly, morphological changes in the striatum have been described in individuals with speech problems such as stuttering (Craig-McQuaide et al., 2014) and non-fluent aphasia (Ogar et al., 2007).

The principal cell type in the striatum is the medium spiny neuron (MSN), which makes up approximately 98% of all striatal cells (Kemp and Powell, 1971; Huang et al., 1992; for review, see Kreitzer and Malenka, 2008). MSNs can be further divided into two categories of neurons that have different connectivity and opposing functions: dopamine receptor type 1 (D1R) and dopamine receptor type 2 (D2R) expressing cells (**Figure 1A**). D1R expressing MSNs connect to thalamic nuclei via the “direct pathway” which results in excitation of the motor cortex. D2R expressing MSNs form an “indirect pathway” that connects to the thalamus via multiple subcortical regions leading to inhibition of the thalamus and thus reduced cortical input (**Figure 1A**), (Albin et al., 1989; Kravitz and Kreitzer, 2012; Calabresi et al., 2014).

This balance between excitation (resulting in more movement) and inhibition (less movement) is crucial for coordinated motor function (Calabresi et al., 2014) including fine orofacial motor control.

In order to unravel the fundamental components that enable humans to effortlessly use spoken language, we will need to understand the neuro-genetic mechanisms involved in establishment, function, and maintenance of speech–motor pathways.

## SPOKEN LANGUAGE AND FOXP2

A breakthrough in speech and language genetics came with the identification of the first gene to cause a speech/language disorder: *FOXP2* (Lai et al., 2001). Mutations in *FOXP2* were found in a large pedigree known as the KE family (Hurst et al., 1990; Fisher et al., 1998; Lai et al., 2001). Affected family members were diagnosed with a severe speech impairment known as developmental verbal dyspraxia (also known as childhood apraxia of speech; OMIM: 602081) and carried a mutation in one copy of their *FOXP2* gene. In addition to speech impairments, affected family members demonstrated receptive and expressive language problems (Watkins et al., 2002a). Although rare, *FOXP2* mutations have been found in a number of unrelated families and individuals with similar speech/language phenotypes (MacDermot et al., 2005; Feuk et al., 2006; Shriberg et al., 2006; Lennon et al., 2007; Palka et al., 2012; Rice et al., 2012; Zilina et al., 2012; for review, see Bacon and Rappold, 2012). In depth investigations of the KE family phenotype indicated a severe impairment in orofacial praxis tasks (Vargha-Khadem et al., 1995; Lai et al., 2001; Watkins

et al., 2002a). In addition, impairments in language production tasks (e.g., phoneme addition, word repetition) were found between control and affected individuals (Vargha-Khadem et al., 1995). Different aspects of speech are thus impaired in KE family members (Watkins et al., 2002a). Orofacial praxis deficits underlie impaired lexicon building and subvocal (internal) speech representations which can affect irregular verb grammar (Doyon et al., 2003) and rule based grammar learning (Ullman, 2001). Thus, some of the language impairments in the KE family could be related to the core speech production deficits observed.

FOXP2, and its murine homolog *Foxp2*, are found across many regions of the developing and postnatal brain (FoxP2 will be used when referring to both species). Intriguing is the high expression of FoxP2 throughout the mouse and human cortico-striato-thalamic motor circuitry (Lai et al., 2003). During early development FoxP2 is broadly expressed in these regions, but in later developmental and postnatal stages expression becomes more restricted (Figure 1B depicts *Foxp2* expression in the postnatal mouse brain). In adults, *Foxp2* is limited to deep layer cortical neurons (layer 5 motor cortex and layer 6 throughout; Ferland et al., 2003; Morikawa et al., 2009; Hisaoka et al., 2010; Tomassy et al., 2010; Reimers-Kipping et al., 2011; Tsui et al., 2013). Within the striatum, *Foxp2* is highly expressed in both types of MSN, though more commonly in D1R MSNs compared to D2R neurons (Vernes et al., 2011). Corresponding with its expression pattern, imaging studies have shown humans with *FOXP2* mutations display structural and functional differences in motor areas. Affected members of the KE family showed structural gray matter volume differences in the motor cortex and striatum (Watkins et al., 2002b). Furthermore, functional imaging studies showed an underactivation of the striatum and altered cortical activation (including speech/motor areas such as the left anterior insular cortex) during word generation and word repetition tasks (Liegeois et al., 2003).

Converging evidence from *Foxp2* expression pattern studies and phenotypic characterization of human mutations suggests that *FOXP2* may play an important role in the development of the speech-motor pathway. The high expression of *Foxp2* in a specific subset of neurons (D1R MSNs) in the striatum indicates a functional specificity related to motor tasks requiring the striato-thalamic connections of the direct pathway. Malfunctions within this pathway could ultimately affect aspects of the motor circuitry related to fine motor control and contribute to the observed speech-motor deficit in humans.

## FOXP2 AS A MOLECULAR ENTRY POINT INTO SPEECH-MOTOR PATHWAYS

FoxP2 is a transcription factor; its molecular function is to regulate the expression of other genes, switching them on or off in a temporally and spatially controlled manner. FoxP2 has been shown to regulate 100s of different genes involved in processes crucial to brain development and function, ranging from neurogenesis and migration, to neurite outgrowth and

synaptic activity (Spiteri et al., 2007; Vernes et al., 2007, 2011; Konopka et al., 2009; Devanna et al., 2014). Recently, evidence has suggested that *FOXP2* regulates a number of genes involved in the retinoic acid (RA) signaling pathway (Devanna et al., 2014). RA is a vitamin-A derivative essential to mammalian development. Disruption of the RA signaling pathway (caused by genetic disruptions or dietary deficiencies) can have severe consequences during development and adulthood (Holson et al., 1997; Krezel et al., 1998).

Retinoic acid induces genetic and morphological changes in cells. When neuronal precursors (cells that generate neurons during development) differentiate into neurons they switch on genes normally found in mature neurons, stop dividing and grow long processes known as neurites (Siegenthaler et al., 2009; Korecka et al., 2013). We previously compared how neuron-like cells with or without *FOXP2* responded to RA and found that cells showed stronger genetic and morphological changes in response to RA if *FOXP2* was present (Devanna et al., 2014). In addition we discovered that *FOXP2* changed the expression of RA receptors – proteins that directly control the cellular response to RA (Devanna et al., 2014). Of particular interest, *FOXP2* upregulated retinoic acid receptor  $\beta$  (*RAR $\beta$* ) and a number of other genes involved in transport or modification of RA were also transcriptionally regulated (e.g., *ROR $\beta$* , *CRABP2*, and *ASCL1*). These experiments suggest an intriguing link between *FOXP2* and the RA pathway, in which *FOXP2* seems to contribute to or modify the cellular response to RA.

Given the importance of the RA pathway for development, this raises new questions about how *FOXP2* might mediate its effects on brain and neural circuit development. Could the relationship between *FOXP2* and the RA pathway be relevant for (1) normal motor circuitry development and function, and/or (2) effects of *FOXP2* dysfunction in patients? To address these questions, we need to understand how *FoxP2* and the RA pathway might interact, and in what way *FoxP2* mutations might affect the RA pathway on a cellular, functional and behavioral level.

## RA, FOXP2, AND MOTOR BEHAVIOR

Retinoic acid is a key compound during embryogenesis, affecting a multitude of critical developmental pathways. Precise control of RA levels is essential for normal brain development as either an excess or a deficiency of RA results in widespread adverse effects on the brain.

Gestational treatment of rats with excess RA results in behavioral deficits in learning, memory and motor function (Holson et al., 1997). Rats treated with excess RA displayed poor generalized motor control including impairments in the 'righting reflex' (the ability to return to upright position), and the ability to sit only on the back paws. In addition, gestationally treated adult rats showed problems with learning and memory, such as decreased learning rates in a water filled T maze (Butcher et al., 1972; Holson et al., 1997). Rats lacking dietary vitamin A (of which RA is a metabolite) also perform poorly on motor learning



and motor performance tasks (Carta et al., 2006). Furthermore, mice engineered to lack a key facilitator of RA signaling (RAR $\beta$ ) develop severe locomotion deficits and are highly impaired on motor learning tasks (Krezel et al., 1998).

The displayed motor deficits are similar to phenotypes observed in mouse models of Foxp2 dysfunction. Mouse models of two well characterized patient mutations of FOXP2 have been created that have comparable phenotypes. One mouse model reflects the R553H missense mutation found in the KE family (Lai et al., 2001). The second mouse model mirrors an early stop codon in exon 7 introduced by a non-sense mutation that leads to a loss of FOXP2 protein in an independent family with speech/language disorder (MacDermot et al., 2005; Groszer et al., 2008). Mice that have a homozygous Foxp2 mutation show severe general motor impairments, reminiscent of animals treated with excess RA. However these Foxp2 homozygous mutants do not survive beyond 3–4 weeks after birth, possibly due to a requirement for Foxp2 in other organs such as the lungs or heart (Groszer et al., 2008). In mice where a single copy of Foxp2 is affected (as per the heterozygous state of the mutations observed in patients) general motor control is normal but motor learning is impaired (Groszer et al., 2008; French et al., 2012). This more subtle phenotype closely resembles the motor learning phenotype observed in RA deprived rats (Carta et al., 2006). For an overview of the different phenotypes exhibited by Foxp2 mutation, RAR mutation, and RA treatment, see Table 1.

## FOXP2 AND RA SIGNALING AFFECT NEURONAL FUNCTION

In addition to the behavioral deficits, vitamin A deprivation/supplementation adversely affects striatal development and function. Cells in the developing lateral ganglionic eminence (the precursor region of the striatum) do not differentiate into the appropriate neuronal subtypes when RA signaling is blocked (Toresson et al., 1999; Chatzi et al., 2011). However restoring RA levels rescued this phenotype and resulted in normal differentiation into appropriate neuronal cell types (Chatzi et al., 2011). Separately, mice engineered to knockout the RAR $\beta$  gene display gross morphological striatal defects including impaired neurogenesis and deficits in acquiring proper neuronal identities (Liao et al., 2008). Lastly, chronic postnatal vitamin A supplementation has been linked to oxidative cell toxicity in the striatum (de Oliveira et al., 2007).

Foxp2 also contributes to striatal cell morphology and function. Foxp2 mutant neurons exhibit reduced neurite growth and branching in primary striatal cultures (Vernes et al., 2011) and the *in vivo* striatum displays aberrant neuronal activity. Mice with a heterozygous Foxp2 mutation showed unusually high activity in the dorsomedial striatum during active motor behavior (French et al., 2012). This suggests striatal cells can no longer properly modulate their activity following input from motor areas when lacking Foxp2. Moreover, the increased striatal activity normally seen when animals perform motor learning tasks was absent in mutant mice. Instead, a decrease

in firing rate was seen, again suggesting aberrant modulation of responses to cortical and/or thalamic input (French et al., 2012). Additionally, extracellular measurements on striatal brain slices from heterozygous Foxp2 mutant animals show these cells fail to respond to induction of long term depression (LTD; Groszer et al., 2008). An inability to induce long term plasticity [either LTD or long term potentiation (LTP)] has debilitating consequences as scaled activity (plasticity) is necessary for circuits to properly regulate their input and output. Synaptic long term plasticity changes underlie information storage and are necessary for learning and memory (Novkovic et al., 2015; Zhu et al., 2015). Interestingly, in the striatum, synaptic plasticity has been strongly linked to motor learning (Dang et al., 2006; Kreitzer and Malenka, 2007). Defects specifically related to striatal LTD and LTP are known to affect procedural motor learning and the acquisition of new motor paradigms (Gubellini et al., 2004).

Aberrant induction of synaptic scaling has also been found in mice following acute RA depletion, which results in a complete lack of hippocampal LTP or LTD (Misner et al., 2001). This phenotype was specific to RA depletion and was reversible, as vitamin A supplementation rapidly restored normal synaptic plasticity (Misner et al., 2001). At a molecular level, RA signaling is mediated by the action of RA receptors (RARs; RAR $\alpha$ , RAR $\beta$ , and RAR $\gamma$ ) and similar plasticity defects have been shown for mice lacking RAR $\alpha$  (Sarti et al., 2012) or RAR $\beta$  (Chiang et al., 1998). Hippocampal cells from these mice fail to establish LTD when subjected to low frequency stimulation – the paradigm necessary to induce LTD in the hippocampus. By contrast, excess RA induced the reverse effect in cultured hippocampal slices, where increased excitatory activity was observed (Aoto et al., 2008). It is not yet known if RA signaling affects synaptic plasticity in the striatum. However, the similarity in synaptic activity phenotypes between Foxp2-, RAR $\alpha$ -, and RAR $\beta$ -deficient animals (albeit focusing on different brain regions) does indicate these transcription factors may play a role in similar intracellular pathways regulating neuronal activity and synaptic plasticity.

The aforementioned plasticity (LTD and/or LTP) deficits in Foxp2, RAR $\alpha$ , and RAR $\beta$  mutant animals suggests an improper reaction of neuronal circuits to changes in external input. Induction of LTD or LTP leads to a decrease or an increase, respectively, in the amount of glutamate receptors (of the AMPA-receptor class) at the synaptic membrane (Seidenman et al., 2003; Briand et al., 2014; for review, see Luscher and Huber, 2010). This change in AMPA receptor abundance modifies the response strength of a cell when it is excited. The change in stimulus–response strength is transient, and in time the normal AMPA receptor distribution will be restored, returning synaptic responses to normal levels. RA treatment of hippocampal cultures has shown an increase of AMPA receptors on the cell surface (Aoto et al., 2008), but no data on the striatum is currently present. The shared synaptic plasticity defect following disruption of RA signaling pathways or Foxp2 mutation does suggest that they both may influence receptor abundance or localization at the synapse in the striatum, an intriguing area for further study.

**TABLE 1 | Overview of phenotypes described in Foxp2 mutation, retinoic acid receptor (RAR) mutation and RA excess/depletion treatments.**

	Deficit	Foxp2 mutation		RA receptor mutation	RA excess/depletion
		Homozygous	Heterozygous		
<b>Development</b>	Embryogenesis defects	–	–	–	+
	Lethality	++	–	+	++
	Aberrant basal ganglia development	++	+	++	++
<b>Cellular</b>	Basal ganglia cell identity defects	NT	–	+	++
	Decreased neurite growth and branching	++	NT	NT	NT
	Aberrant neuronal activity in striatum	++	++	–	+
	Unable to induce LTD	NT	++	++	NT
	Unable to induce LTP	NT	NT	++	NT
<b>Behavior</b>	General motor control deficits	++	–	++	++
	Motor learning deficits	N/A	+	++	++
		(postnatal lethality)			(postnatal treatment)
	Spatial learning deficits	N/A	NT	+	+
		(postnatal lethality)			

–, no effect; +, mild effect; ++, strong effect; N/A, not applicable; NT, not tested.

A thorough investigation of the mechanisms leading to LTD and LTP deficits resulting from RA/RAR and Foxp2 malfunction will be necessary to understand if they function in the same pathways. Understanding the molecular mechanisms underlying striatal function, especially related to complex motor circuitry function, will lead to a better understanding of striatal speech-motor control.

## MOLECULAR LINKS BETWEEN RARs AND FOXP2

Retinoic acid receptors canonically function as transcription factors, regulating genes responsible for directing normal embryogenesis and brain development. Interestingly, FoxP2 and RARs share some of the same target genes (Balmer and Blomhoff, 2002; Delacroix et al., 2010; Devanna et al., 2014). RARs are highly expressed in the brain (Krezel et al., 1999) and are present throughout embryonal development (Mollard et al., 2000), postnatal development (Wei et al., 2011), and in adults (Krezel et al., 1999; Zetterstrom et al., 1999). Notably high expression of RARs can be found throughout the motor circuitry, including cortical, striatal, and multiple thalamic regions (Krezel et al., 1999), (**Figure 1B**). We focus on two key receptors found in the motor circuitry: RAR $\alpha$  and RAR $\beta$ . RAR $\alpha$  is found in layer 5 of the cortex and in the thalamus – both regions that overlap with murine Foxp2 expression (Krezel et al., 1999; Zetterstrom et al., 1999; Ferland et al., 2003; Lai et al., 2003; Hisaoka et al., 2010). Interestingly, Foxp2 only overlaps with RAR $\alpha$  in the motor cortex layer 5, because Foxp2 expression is largely restricted to layer 6 of other mature cortical areas. RAR $\beta$  is strongly expressed only in the striatum, another site where Foxp2 expression is highest (**Figure 1B**). Notably, FOXP2 has been shown to directly drive RAR $\beta$  expression in human cells (Vernes et al., 2007; Devanna

et al., 2014), although this is yet to be shown in the striatum. This high level of overlap, combined with shared target genes and molecular interactions, strongly supports interplay between FoxP2 and RARs in motor pathways.

## CONCLUDING REMARKS

In addition to its canonical role during embryogenesis, studies described here suggest RA signaling plays a specific role in the development and function of striatal motor circuitry and may link to FoxP2 function. Disruption of the RA pathway results in strikingly similar phenotypes to FoxP2 mutation on multiple levels, which suggests a potential mechanistic interaction. FoxP2 and RARs can regulate some common target genes, affect similar cellular phenotypes and show highly overlapping expression patterns in the cortico-striato-thalamic motor circuitry. In the striatum, aberrant function of Foxp2 and RA signaling contributes to altered development and, in the case of mutations of mouse Foxp2, altered synaptic plasticity similar to that seen in the hippocampus of RAR $\alpha$  mutant animals. Given that RAR $\beta$  is predominantly expressed in the postnatal striatum, it seems likely that its disruption will also affect striatal plasticity, however, this is yet to be experimentally determined. Lastly, animals with mutated Foxp2 or RA signaling defects show comparable motor control/learning impairments. Thus at multiple levels (molecular, cellular, circuit, and behavioral) there is evidence that interplay between FoxP2 and RA signaling may facilitate proper development and function of motor circuitry. This evidence from mice is strengthened by findings in songbirds which show both FoxP2 and RA influence song learning by acting in circuits that have parallels with human vocal-motor pathways (Haesler et al., 2007; Wood et al., 2008). In the future it will be of great value to understand if these signaling cascades interact to influence neuronal mechanisms related to

song learning or speech-motor control, and if RA signaling deficits are involved in aberrant speech-motor development in humans. The capacity for human speech and spoken language is dependent on multiple molecular and neural building blocks. With the link between FoxP2 and RA signaling, a new block has been suggested, giving us new opportunities to investigate the evolution and development of the (spoken) language ready brain.

## REFERENCES

- Albin, R. L., Young, A. B., and Penney, J. B. (1989). The functional-anatomy of basal ganglia disorders. *Trends Neurosci.* 12, 366–375. doi: 10.1016/0166-2236(89)90074-X
- Aoto, J., Nam, C. I., Poon, M. M., Ting, P., and Chen, L. (2008). Synaptic signaling by all-trans retinoic acid in homeostatic synaptic plasticity. *Neuron* 60, 308–320. doi: 10.1016/j.neuron.2008.08.012
- Bacon, C., and Rappold, G. A. (2012). The distinct and overlapping phenotypic spectra of FOXP1 and FOXP2 in cognitive disorders. *Hum. Genet.* 131, 1687–1698. doi: 10.1007/s00439-012-1193-z
- Balmer, J. E., and Blomhoff, R. (2002). Gene expression regulation by retinoic acid. *J. Lipid Res.* 43, 1773–1808. doi: 10.1194/jlr.R100015-JLR200
- Booth, J. R., Wood, L., Lu, D., Houk, J. C., and Bitan, T. (2007). The role of the basal ganglia and cerebellum in language processing. *Brain Res.* 1133, 136–144. doi: 10.1016/j.brainres.2006.11.074
- Briand, L. A., Kimmey, B. A., Ortinski, P. I., Haganir, R. L., and Pierce, R. C. (2014). Disruption of glutamate receptor-interacting protein in nucleus accumbens enhances vulnerability to cocaine relapse. *Neuropsychopharmacology* 39, 759–769. doi: 10.1038/npp.2013.265
- Butcher, R. E., Brunner, R. L., Roth, T., and Kimmel, C. A. (1972). Learning impairment associated with maternal hypervitaminosis-a in rats. *Life Sci.* 11, 141–145. doi: 10.1016/0024-3205(72)90227-5
- Calabresi, P., Picconi, B., Tozzi, A., Ghiglieri, V., and Di Filippo, M. (2014). Direct and indirect pathways of basal ganglia: a critical reappraisal. *Nat. Neurosci.* 17, 1022–1030. doi: 10.1038/nn.3743
- Carta, M., Stancampiano, R., Tronci, E., Collu, M., Usiello, A., Morelli, M., et al. (2006). Vitamin A deficiency induces motor impairments and striatal cholinergic dysfunction in rats. *Neuroscience* 139, 1163–1172. doi: 10.1016/j.neuroscience.2006.01.027
- Chatzi, C., Brade, T., and Duester, G. (2011). Retinoic acid functions as a key GABAergic differentiation signal in the basal ganglia. *PLoS Biol.* 9:e1000609. doi: 10.1371/journal.pbio.1000609
- Chiang, M. Y., Misner, D., Kempermann, G., Schikorski, T., Giguere, V., Sucov, H. M., et al. (1998). An essential role for retinoid receptors RAR beta and RXR gamma in long-term potentiation and depression. *Neuron* 21, 1353–1361. doi: 10.1016/S0896-6273(00)80654-6
- Craig-McQuaide, A., Akram, H., Zrinzo, L., and Tripoliti, E. (2014). A review of brain circuitries involved in stuttering. *Front. Hum. Neurosci.* 8:884. doi: 10.3389/fnhum.2014.00884
- Dang, M. T., Yokoi, F., Yin, H. H., Lovinger, D. M., Wang, Y. Y., and Li, Y. Q. (2006). Disrupted motor learning and long-term synaptic plasticity in mice lacking NMDAR1 in the striatum. *Proc. Natl. Acad. Sci. U.S.A.* 103, 15254–15259. doi: 10.1073/pnas.0601758103
- Delacroix, L., Moutier, E., Altobelli, G., Legras, S., Poch, O., Choukallah, M. A., et al. (2010). Cell-specific interaction of retinoic acid receptors with target genes in mouse embryonic fibroblasts and embryonic stem cells. *Mol. Cell. Biol.* 30, 231–244. doi: 10.1128/MCB.00756-09
- de Oliveira, M. R., De Bittencourt Pasquali, M. A., Silvestrin, R. B., Mello E Souza, T., and Moreira, J. C. (2007). Vitamin A supplementation induces a prooxidative state in the striatum and impairs locomotory and exploratory activity of adult rats. *Brain Res.* 1169, 112–119. doi: 10.1016/j.brainres.2007.07.008
- Devanna, P., Middelbeek, J., and Vernes, S. C. (2014). FOXP2 drives neuronal differentiation by interacting with retinoic acid signaling pathways. *Front. Cell. Neurosci.* 8:305. doi: 10.3389/fncel.2014.00305
- Doyon, J., Penhune, V., and Ungerleider, L. G. (2003). Distinct contribution of the cortico-striatal and cortico-cerebellar systems to motor skill learning. *Neuropsychologia* 41, 252–262. doi: 10.1016/S0028-3932(02)00158-6
- Ferland, R. J., Cherry, T. J., Preware, P. O., Morrissey, E. E., and Walsh, C. A. (2003). Characterization of Foxp2 and Foxp1 mRNA and protein in the developing and mature brain. *J. Compar. Neurol.* 460, 266–279. doi: 10.1002/cne.10654
- Feuk, L., Kalervo, A., Lipsanen-Nyman, M., Skaug, J., Nakabayashi, K., Finucane, B., et al. (2006). Absence of a paternally inherited FOXP2 gene in developmental verbal dyspraxia. *Am. J. Hum. Genet.* 79, 965–972. doi: 10.1086/508902
- Fisher, S. E., Vargha-Khadem, F., Watkins, K. E., Monaco, A. P., and Pembrey, M. E. (1998). Localisation of a gene implicated in a severe speech and language disorder. *Nat. Genet.* 18, 168–170. doi: 10.1038/ng0298-168
- French, C. A., Jin, X., Campbell, T. G., Gerfen, E., Groszer, M., Fisher, S. E., et al. (2012). An aetiological Foxp2 mutation causes aberrant striatal activity and alters plasticity during skill learning. *Mol. Psychiatry* 17, 1077–1085. doi: 10.1038/mp.2011.105
- Groszer, M., Keays, D. A., Deacon, R. M. J., De Bono, J. P., Prasad-Mulcare, S., Gaub, S., et al. (2008). Impaired synaptic plasticity and motor learning in mice with a point mutation implicated in human speech deficits. *Curr. Biol.* 18, 354–362. doi: 10.1016/j.cub.2008.01.060
- Gubellini, P., Pisani, A., Centonze, D., Bernardi, G., and Calabresi, P. (2004). Metabotropic glutamate receptors and striatal synaptic plasticity: implications for neurological diseases. *Prog. Neurobiol.* 74, 271–300. doi: 10.1016/j.pneurobio.2004.09.005
- Haesler, S., Rochefort, C., Georgi, B., Licznarski, P., Osten, P., and Scharff, C. (2007). Incomplete and inaccurate vocal imitation after knockdown of FoxP2 in songbird basal ganglia nucleus Area X. *PLoS Biol.* 5:e321. doi: 10.1371/journal.pbio.0050321
- Hisaoka, T., Nakamura, Y., Senba, E., and Morikawa, Y. (2010). The forkhead transcription factors, Foxp1 and Foxp2, identify different subpopulations of projection neurons in the mouse cerebral cortex. *Neuroscience* 166, 551–563. doi: 10.1016/j.neuroscience.2009.12.055
- Holson, R. R., Gazzara, R. A., Ferguson, S. A., and Adams, J. (1997). Behavioral effects of low-dose gestational day 11–13 retinoic acid exposure. *Neurotoxicol. Teratol.* 19, 355–362. doi: 10.1016/S0892-0362(97)00041-X
- Huang, Q., Zhou, D., Chase, K., Gusella, J. F., Aronin, N., and Difiglia, M. (1992). Immunohistochemical localization of the D1 dopamine receptor in rat-brain reveals its axonal-transport, presynaptic and postsynaptic localization, and prevalence in the basal ganglia, limbic system, and thalamic reticular nucleus. *Proc. Natl. Acad. Sci. U.S.A.* 89, 11988–11992. doi: 10.1073/pnas.89.24.11988
- Hurst, J. A., Baraitser, M., Auger, E., Graham, F., and Norell, S. (1990). An extended family with a dominantly inherited speech disorder. *Dev. Med. Child Neurol.* 32, 352–355. doi: 10.1111/j.1469-8749.1990.tb16948.x
- Kemp, J. M., and Powell, T. P. S. (1971). Termination of fibres from cerebral cortex and thalamus upon dendritic spines in caudate nucleus - study with golgi method. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 262, 429–439. doi: 10.1098/rstb.1971.0105
- Konopka, G., Bomar, J. M., Winden, K., Coppola, G., Jonsson, Z. O., Gao, F. Y., et al. (2009). Human-specific transcriptional regulation of CNS development genes by FOXP2. *Nature* 462, U213–U289. doi: 10.1038/nature08549
- Korecka, J. A., Van Kesteren, R. E., Blaas, E., Spitzer, S. O., Kamstra, J. H., Smit, A. B., et al. (2013). Phenotypic characterization of retinoic acid differentiated sh-sy5y cells by transcriptional profiling. *PLoS ONE* 8:e63862. doi: 10.1371/journal.pone.0063862

## ACKNOWLEDGMENTS

This work was supported by a grant from the Donders Institute, Radboud University Nijmegen, a Marie Curie career development grant awarded to SCV and by the Max Planck Society. We would like to thank Moritz Negwer for valuable comments on the manuscript.

- Kravitz, A. V., and Kreitzer, A. C. (2012). Striatal mechanisms underlying movement, reinforcement, and punishment. *Physiology* 27, 167–177. doi: 10.1152/physiol.00004.2012
- Kreitzer, A. C., and Malenka, R. C. (2007). Endocannabinoid-mediated rescue of striatal LTD and motor deficits in Parkinson's disease models. *Nature* 445, 643–647. doi: 10.1038/nature05506
- Kreitzer, A. C., and Malenka, R. C. (2008). Striatal plasticity and basal ganglia circuit function. *Neuron* 60, 543–554. doi: 10.1016/j.neuron.2008.11.005
- Krezel, W., Ghyselinck, N., Samad, T. A., Dupe, V., Kastner, P., Borrelli, E., et al. (1998). Impaired locomotion and dopamine signaling in retinoid receptor mutant mice. *Science* 279, 863–867. doi: 10.1126/science.279.5352.863
- Krezel, W., Kastner, P., and Chambon, P. (1999). Differential expression of retinoid receptors in the adult mouse central nervous system. *Neuroscience* 89, 1291–1300. doi: 10.1016/S0306-4522(98)00342-X
- Lai, C. S. L., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F., and Monaco, A. P. (2001). A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* 413, 519–523. doi: 10.1038/35097076
- Lai, C. S., Gerrelli, D., Monaco, A. P., Fisher, S. E., and Copp, A. J. (2003). FOXP2 expression during brain development coincides with adult sites of pathology in a severe speech and language disorder. *Brain* 126, 2455–2462. doi: 10.1093/brain/awg247
- Lennon, P. A., Cooper, M. L., Peiffer, D. A., Gunderson, K. L., Patel, A., Peters, S., et al. (2007). Deletion of 7q31-1 supports involvement of FOXP2 in language impairment: clinical report and review. *Am. J. Med. Genet. Part A* 143a, 791–798. doi: 10.1002/ajmg.a.31632
- Liao, W. L., Tsai, H. C., Wang, H. F., Chang, J., Lu, K. M., Wu, H. L., et al. (2008). Modular patterning of structure and function of the striatum by retinoid receptor signaling. *Proc. Natl. Acad. Sci. U.S.A.* 105, 6765–6770. doi: 10.1073/pnas.0802109105
- Lieberman, P. (2002). On the nature and evolution of the neural bases of human language. *Am. J. Phys. Anthropol.* 45, 36–62. doi: 10.1002/ajpa.10171
- Lieberman, P. (2007). The evolution of human speech - Its anatomical and neural bases. *Curr. Anthropol.* 48, 39–66. doi: 10.1098/rstb.2011.0099
- Liegeois, F., Baldeweg, T., Connelly, A., Gadian, D. G., Mishkin, M., and Vargha-Khadem, F. (2003). Language fMRI abnormalities associated with FOXP2 gene mutation. *Nat. Neurosci.* 6, 1230–1237. doi: 10.1038/nn1138
- Luscher, C., and Huber, K. M. (2010). Group 1 mGluR-dependent synaptic long-term depression: mechanisms and implications for circuitry and disease. *Neuron* 65, 445–459. doi: 10.1016/j.neuron.2010.01.016
- MacDermot, K. D., Bonora, E., Sykes, N., Coupe, A. M., Lai, C. S. L., Vernes, S. C., et al. (2005). Identification of FOXP2 truncation as a novel cause of developmental speech and language deficits. *Am. J. Hum. Genet.* 76, 1074–1080. doi: 10.1086/430841
- Misner, D. L., Jacobs, S., Shimizu, Y., De Urquiza, A. M., Solomin, L., Perlmann, T., et al. (2001). Vitamin A deprivation results in reversible loss of hippocampal long-term synaptic plasticity. *Proc. Natl. Acad. Sci. U.S.A.* 98, 11714–11719. doi: 10.1073/pnas.191369798
- Mollard, R., Viville, S., Ward, S. J., Decimo, D., Chambon, P., and Dolle, P. (2000). Tissue-specific expression of retinoic acid receptor isoform transcripts in the mouse embryo. *Mech. Dev.* 94, 223–232. doi: 10.1016/S0925-4773(00)00303-8
- Morikawa, Y., Hisaoka, T., and Senba, E. (2009). Characterization of Foxp2-expressing cells in the developing spinal cord. *Neuroscience* 162, 1150–1162. doi: 10.1016/j.neuroscience.2009.05.022
- Novkovic, T., Shchyglo, O., Gold, R., and Manahan-Vaughan, D. (2015). Hippocampal function is compromised in an animal model of multiple sclerosis. *Neuroscience* 309, 100–112. doi: 10.1016/j.neuroscience.2015.03.008
- Ogar, J. M., Dronkers, N. F., Brambati, S. M., Miller, B. L., and Gorno-Tempini, M. L. (2007). Progressive nonfluent aphasia and its characteristic motor speech deficits. *Alzheimer Dis. Assoc. Disord.* 21, S23–S30. doi: 10.1097/WAD.0b013e31815d19fe
- Palka, C., Alfonsi, M., Mohn, A., Cerbo, R., Franchi, P. G., Fantasia, D., et al. (2012). Mosaic 7q31 deletion involving FOXP2 gene associated with language impairment. *Pediatrics* 129, E183–E188. doi: 10.1542/peds.2010-2094
- Reimers-Kipping, S., Hevers, W., Paabo, S., and Enard, W. (2011). Humanized Foxp2 specifically affects cortico-basal ganglia circuits. *Neuroscience* 175, 75–84. doi: 10.1016/j.neuroscience.2010.11.042
- Rice, G. M., Raca, G., Jakielski, K. J., Laffin, J. J., Iyama-Kurtycz, C. M., Hartley, S. L., et al. (2012). Phenotype of FOXP2 haploinsufficiency in a mother and son. *Am. J. Med. Genet. Part A* 158a, 174–181. doi: 10.1002/ajmg.a.34354
- Sarti, F., Schroeder, J., Aoto, J., and Chen, L. (2012). Conditional RAR alpha knockout mice reveal acute requirement for retinoic acid and RAR alpha in homeostatic plasticity. *Front. Mol. Neurosci.* 5:16. doi: 10.3389/fnmol.2012.00016
- Seidenman, K. J., Steinberg, J. P., Hagan, R., and Malinow, R. (2003). Glutamate receptor subunit 2 serine 880 phosphorylation modulates synaptic transmission and mediates plasticity in CA1 pyramidal cells. *J. Neurosci.* 23, 9220–9228.
- Shriberg, L. D., Ballard, K. J., Tomblin, J. B., Duffy, J. R., Odell, K. H., and Williams, C. A. (2006). Speech, prosody, and voice characteristics of a mother and daughter with a 7;13 translocation affecting FOXP2. *J. Speech Lang. Hear. Res.* 49, 500–525. doi: 10.1044/1092-4388(2006)038
- Siegenthaler, J. A., Ashique, A. M., Zarbalis, K., Patterson, K. P., Hecht, J. H., Kane, M. A., et al. (2009). Retinoic acid from the meninges regulates cortical neuron generation. *Cell* 139, 597–609. doi: 10.1016/j.cell.2009.10.004
- Spiteri, E., Konopka, G., Coppola, G., Bomar, J., Oldham, M., Ou, J., et al. (2007). Identification of the transcriptional targets of FOXP2, a gene linked to speech and language, in developing human brain. *Am. J. Hum. Genet.* 81, 1144–1157. doi: 10.1086/522237
- Tomassy, G. S., De Leonibus, E., Jabaudon, D., Lodato, S., Alfano, C., Mele, A., et al. (2010). Area-specific temporal control of corticospinal motor neuron differentiation by COUP-TFI. *Proc. Natl. Acad. Sci. U.S.A.* 107, 3576–3581. doi: 10.1073/pnas.0911792107
- Torresson, H., Mata De Urquiza, A., Fagerstrom, C., Perlmann, T., and Campbell, K. (1999). Retinoids are produced by glia in the lateral ganglionic eminence and regulate striatal neuron differentiation. *Development* 126, 1317–1326.
- Tsui, D., Vessey, J. P., Tomita, H., Kaplan, D. R., and Miller, F. D. (2013). FoxP2 Regulates Neurogenesis during Embryonic Cortical Development. *J. Neurosci.* 33, 244–258. doi: 10.1523/JNEUROSCI.1665-12.2013
- Ullman, M. T. (2001). The declarative/procedural model of lexicon and grammar. *J. Psycholinguist. Res.* 30, 37–69. doi: 10.1023/A:1005204207369
- Vargha-Khadem, F., Watkins, K., Alcock, K., Fletcher, P., and Passingham, R. (1995). Praxic and nonverbal cognitive deficits in a large family with a genetically transmitted speech and language disorder. *Proc. Natl. Acad. Sci. U.S.A.* 92, 930–933. doi: 10.1073/pnas.92.3.930
- Vernes, S. C., Oliver, P. L., Spiteri, E., Lockstone, H. E., Puliyadi, R., Taylor, J. M., et al. (2011). Foxp2 regulates gene networks implicated in neurite outgrowth in the developing brain. *PLoS Genet.* 7:e1002145. doi: 10.1371/journal.pgen.1002145
- Vernes, S. C., Spiteri, E., Nicod, J., Groszer, M., Taylor, J. M., Davies, K. E., et al. (2007). High-throughput analysis of promoter occupancy reveals direct neural targets of FOXP2, a gene mutated in speech and language disorders. *Am. J. Hum. Genet.* 81, 1232–1250. doi: 10.1086/522238
- Watkins, K. E., Dronkers, N. F., and Vargha-Khadem, F. (2002a). Behavioural analysis of an inherited speech and language disorder: comparison with acquired aphasia. *Brain* 125(pt 3), 452–464. doi: 10.1093/brain/awf058
- Watkins, K. E., Vargha-Khadem, F., Ashburner, J., Passingham, R. E., Connelly, A., Friston, K. J., et al. (2002b). MRI analysis of an inherited speech and language disorder: structural brain abnormalities. *Brain* 125(pt 3), 465–478. doi: 10.1093/brain/awf057
- Wei, J. A., Yi, W. E., Min, G., Yuan, S., Li, C., Yang, B., et al. (2011). The pattern of retinoic acid receptor expression and subcellular, anatomic and functional area translocation during the postnatal development of the rat cerebral cortex and white matter. *Brain Res.* 1382, 77–87. doi: 10.1007/s00221-011-2873-y
- Wildgruber, D., Ackermann, H., and Grodd, W. (2001). Differential contributions of motor cortex, basal ganglia, and cerebellum to speech motor control: effects of syllable repetition rate evaluated by fMRI. *Neuroimage* 13, 101–109. doi: 10.1006/nimg.2000.0672
- Wood, W. E., Olson, C. R., Lovell, P. V., and Mello, C. V. (2008). Dietary retinoic acid affects song maturation and gene expression in the song system of the zebra finch. *Dev. Neurobiol.* 68, 1213–1224. doi: 10.1002/dneu.20642
- Zetterstrom, R. H., Lindqvist, E., De Urquiza, A. M., Tomac, A., Eriksson, U., Perlmann, T., et al. (1999). Role of retinoids in the CNS: differential expression



- of retinoid binding proteins and receptors and evidence for presence of retinoic acid. *Euro. J. Neurosci.* 11, 407–416. doi: 10.1046/j.1460-9568.1999.00444.x
- Zhu, D., Li, C. C., Swanson, A. M., Villalba, R. M., Guo, J. D., Zhang, Z. B., et al. (2015). BAI1 regulates spatial learning and synaptic plasticity in the hippocampus. *J. Clin. Investig.* 125, 1497–1508. doi: 10.1172/JCI74603
- Zilina, O., Reimand, T., Zjablovskaja, P., Mannik, K., Mannamaa, M., Traat, A., et al. (2012). Maternally and paternally inherited deletion of 7q31 involving the FOXP2 gene in two families. *Am. J. Med. Genet. Part A* 158a, 254–256. doi: 10.1002/ajmg.a.34378

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 van Rhijn and Vernes. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Merge in the Human Brain: A Sub-Region Based Functional Investigation in the Left Pars Opercularis

Emiliano Zaccarella<sup>1,2\*</sup> and Angela D. Friederici<sup>1,2</sup>

<sup>1</sup> Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany, <sup>2</sup> Berlin School of Mind and Brain, Humboldt-Universität zu Berlin, Berlin, Germany

## OPEN ACCESS

### Edited by:

Cedric Boeckx,  
Catalan Institute for Research and  
Advanced Studies (ICREA) and  
Universitat de Barcelona, Spain

### Reviewed by:

Narly Golestani,  
Université de Genève, Switzerland  
Evie Malaia,  
University of Texas at Arlington, USA

### \*Correspondence:

Emiliano Zaccarella  
zaccarella@cbs.mpg.de

### Specialty section:

This article was submitted to  
Language Sciences,  
a section of the journal  
Frontiers in Psychology

**Received:** 15 September 2015

**Accepted:** 10 November 2015

**Published:** 27 November 2015

### Citation:

Zaccarella E and Friederici AD (2015)  
Merge in the Human Brain:  
A Sub-Region Based Functional  
Investigation in the Left Pars  
Opercularis. *Front. Psychol.* 6:1818.  
doi: 10.3389/fpsyg.2015.01818

Language is thought to represent one of the most complex cognitive functions in humans. Here we break down complexity of language to its most basic syntactic computation which hierarchically binds single words together to form larger phrases and sentences. So far, the neural implementation of this basic operation has only been inferred indirectly from studies investigating more complex linguistic phenomena. In the present sub-region based functional magnetic resonance imaging (fMRI) study we directly assessed the neuroanatomical nature of this process. Our results showed that syntactic phrases—compared to word-list sequences—corresponded to increased neural activity in the ventral-anterior portion of the left pars opercularis [Brodmann Area (BA) 44], whereas the adjacently located deep frontal operculum/anterior insula (FOP/aINS), a phylogenetically older and less specialized region, was found to be equally active for both conditions. Crucially, the functional activity of syntactic binding was confined to one out of five clusters proposed by a recent fine-grained sub-anatomical parcellation for BA 44, with consistency across individuals. Neuroanatomically, the present results call for a redefinition of BA 44 as a region with internal functional specializations. Neurocomputationally, they support the idea of invariance within BA 44 in the location of activation across participants for basic syntactic building processing.

**Keywords:** pars opercularis, clusters, syntax, merge, fMRI

## INTRODUCTION

Traditionally language is thought of as one of the most complex cognitive functions. Recently, it has been claimed, however, that the human capacity to process complex syntactic structures is based on a very basic binary process which syntactically binds words together hierarchically to form larger structures. Because of the fundamental nature of this computation, called *merge* in theoretical linguistics (Chomsky, 1999; Adger, 2003), the determination of its neural implementation would constitute the neurobiological basis of a process which is at the root of any complex syntactic structure (Berwick et al., 2013). Up to now, the operation has almost never been directly studied in isolation, as syntax usually has been studied in more complex sentential contexts (Just et al., 1996; Stromswold et al., 1996; Moro et al., 2001; Cooke et al., 2002; Röder et al., 2002; Ben-Shachar et al., 2003, 2004; Constable et al., 2004; Bornkessel et al., 2005; Fiebach et al., 2005; Grewe et al., 2005; Friederici et al., 2006b; Santi and Grodzinsky, 2007, 2010; Caplan et al., 2008; Kinno et al., 2008;

Newman et al., 2010). These and other studies across different languages indicate that the larger region in and around Broca's area in the inferior frontal cortex (IFG) supports syntactic processes (for reviews see Vigneau et al., 2006; Friederici, 2011).

A second region to be considered in the frontal cortex is the frontal operculum (FOP) which is a phylogenetically older than Brodmann Area (BA) 44 (Sanides, 1962; Friederici, 2006) and has been shown to be involved in syntactic classification as well as word based processing (Grasby et al., 1994; Stowe et al., 1998; Friederici et al., 2000b, 2006a). The adjacent anterior insula (aINS) is involved in the processing of short two-word sequences independent of whether they constitute a phrase or not (Zaccarella and Friederici, 2015). Other studies trying to localize syntactic processes reported the involvement of temporal regions, i.e., the left posterior superior and anterior temporal lobe (ATL) regions rather than inferior frontal regions (Bottini et al., 1994; Stowe et al., 1999; Vandenberghe et al., 2002; Humphries et al., 2005, 2006). These studies, however, used long sentences and compared these to word-lists often allowing minimal syntactic processes. The lack of any ATL activity in studies comparing sentences that only differed in syntactic complexity (Friederici et al., 2006b) speaks in favor of a compositional semantic role rather than a syntactic role of this area (Barsalou, 1982; Humphries et al., 2007). Indeed, a series of recent magnetoencephalography (MEG) experiments looking at conceptual compositionality effects at the phrasal level, found the ATL to be active during the construction of complex semantic representations, when color concepts had to be combined together with real objects, or when the same colors had to be combined with nouns (object labels) carrying semantic information (Bemis and Pykkänen, 2011, 2013; Del Prato and Pykkänen, 2014). Semantic sensitivity for the area is strongly confirmed by a recent fMRI study showing that activity change in the ATL varied as a function of the presence of lexico-semantic information, but not of syntactic variables, during the construction of progressively increasing linguistic structures (Pallier et al., 2011).

The goal of the present study was to identify the neural basis of the most basic syntactic computation, upon which any more complex hierarchical structure can be derived. By this computation two words, i.e., *this* and *ship*, are bound together to a hierarchical phrase containing both words—i.e., *this ship* with *this* dominating *ship*. By applying the same mechanism again, we then recombine this phrase with the closest element occurring in the sentence, to form increasing syntactic hierarchies—i.e., *this ship sinks*. Phrases of two-word length—like *this ship*—are the ideal level to investigate this most basic process of syntactic binding, as the amount of cognitive load required to process such small constructions is very limited. This means that—after classifying *this* as a determiner and *ship* as a noun—only the operation of merge is necessary to make it a phrase. Crucially, at the very same two-word level, it is also possible to create contexts consisting of simple lists of words—like *stone, ship*—in which no phrase can be created, as no syntactic relationship holds. Because of this minimal opposition, two-word manipulations are an ideal level to identify the most basic and essential syntactic computation of *merge*. We hypothesize that this computation

should be located in Broca's area, as this region has been found to support syntactic processes (Vigneau et al., 2006; Friederici, 2011; Hagoort and Indefrey, 2014) and further that if its assumed fundamental nature holds, it should be localizable (a) in a very confined subregion within this area, and (b) with little variance across individuals.

Neuroanatomically, Broca's area can be subdivided into BA 45 and BA 44, whose borders were first mapped based on the cellular organization of their regional tissues (Brodmann, 1909), and then redefined using observer-independent cell density profiles over histological slices of postmortem brains (Amunts et al., 1999). More recently, a multireceptor-based analysis separated BA 44 into an anterior dorsal part, and a posterior ventral part (Zilles and Amunts, 2009; Amunts et al., 2010). A very recent meta-analytic functional connectivity-based parcellation (CBP) approach even proposed a decomposition of BA 44 into five separate subregions, called clusters (Cs) two of which located in its more posterior part (C1 and C4), another two in the more anterior part (C2 and C3), and a third one in the inferior frontal junction (C5; Clos et al., 2013). The CBP approach first identifies the whole-brain co-activation pattern for each voxel contained in BA 44 across several thousands fMRI studies, and then groups together those voxels into distinct clusters, according to the similarity of their co-activation patterns across the brain. While the functional specificity of these five clusters as indicated by the metaanalysis is low, with each of the clusters' functional domain ranging from action, working memory, switching to other cognitive tasks, the fine-grained subdivision of BA 44 may allow a precise localization of the most fundamental syntactic process assumed for any natural language. Here we hypothesize, that if the high functional specificity of *merge* as a fundamental syntactic computation holds, this mechanism should to be localizable in one of the sub-clusters within BA 44, with little inter-individual variance.

This hypothesis was investigated in an fMRI study using two-word sequences either allowing hierarchical syntactic binding to apply (phrase trials) or not (list trials), which were presented visually. Phrase trials and list trials were constructed as parallel as possible, only varying in the possible application of syntactic binding. They differed in their first element, which was either a determiner (e.g., *this*) or a noun (*apple*), while the second element was always a phonotactically legal pseudoword (*flirk*) to drastically reduce conceptual-semantic processing in both conditions (Bemis and Pykkänen, 2011). Two corresponding one-word conditions were also included, in which the pseudowords were substituted with a series of X's (e.g., XXXXX) to explore the effect of number of words. Thus, the experiment included two factors: type of STRUCTURE ["phrase" (PH), allowing syntactic binding, vs. "list" (LS) not allowing syntactic binding] and number of WORDS ("2-words" vs. "1 word plus Xs"). We employed three progressively region-stringent levels of data analysis to localize syntactic binding: (i) a whole-brain analysis to know whether BA 44 and FOP and/or the aINS show activity during two-word processing; (ii) a more restricted volume-of-interest analysis for BA 44 to directly contrast phrase processing vs. word-list processing; and (iii) a cluster-of-interest analysis within BA 44 to localize syntactic

binding at the individual subject level. These three analyses allow us to test whether: (i) BA 44 is highly sensitive to structure formation even at the lowest level of phrase structure building; (ii) the fundamental nature of syntactic binding is expressed by a stringent localization in a subregion within BA 44, using a cluster-based approach in which the five clusters by Clos et al. (2013) are used to test sub-regional sensitivity for merge within BA 44; (iii) the invariant character of syntactic binding assuming little variance of the localization within this subcluster of BA 44 across individuals.

## MATERIALS AND METHODS

### Participants

We tested 27 right-handed subjects (Oldfield, 1971), but only 22 subjects (11 female; mean age 28.5 years, standard deviation (SD) 3.62 years; all native German speakers) were included in the analysis. Four subjects were excluded because of poor behavioral performance. One additional subject was excluded because the trial list file was corrupted. The local ethics committee of the University of Leipzig approved all procedures used during the experiment. Written informed consent was obtained from all subjects.

### Stimuli Construction

At the 2-words level, phrasal syntactic contexts (2-PH) comprised eight adjectival determiners of two-syllables in length to appear as first word—*jede/jedes* (each), *eure/euer* (your), *jene/jenes* (that), *diese/dieses* (this) followed by 48 different pseudowords (e.g., DIESE FLIRK). List contexts (2-LS) comprised eight nouns selected from the CELEX corpus for German (Baayen et al., 1995) which were matched to the determiners for syllabic length, letter length, and syllabic stress—*Apfel* (apple), *Käse* (cheese), *Ofen* (oven), *Efeu* (ivy), *Motor* (motor), *Kiwi* (kiwi), *Haken* (hook), *Koffer* (suitcase). A corresponding example of list context was APFEL, FLIRK. These words and the pseudowords were constructed and controlled according to all relevant psycholinguistic parameters, and for the pseudowords we strongly avoided associative effects with real words by using an automatized screening procedure based on the same CELEX corpus, followed by a final filtering selection done by three mother-tongue German speakers (see Appendix A—Stimuli construction for detailed information). Importantly, the use of pseudowords in either phrasal or list contexts at two-word level, was directly intended to reduce potential interactions due to semantic activity, given indeed that we were explicitly interested in finding neural correlates of syntactic processing in the brain. In doing so, we took advantage of the intrinsic linguistic distinction between the syntactically prominent functional lexicon (e.g., determiners, prepositions, conjunctions) and the semantically prominent contentive lexicon (e.g., nouns, verbs, adjectives), and coupled it together with the semantic-free nature of the pseudowords themselves. In this respect, since determiners have less semantic content than do nouns, the choice to use pseudowords instead of real words helped us to: (1) keep syntactic activity at work in determiner-pseudoword contexts, while removing semantic information; (2) remove syntactic and

semantic information in the noun-pseudoword contexts, by reducing compounding effects via head de-lexicalization (see Supplementary Material); (3) shed further light on the role of BA 44 as syntactic-sensitive area. Finally, at the 1-word level, the pseudowords were substituted with a series of X's (e.g., XXXXX) to obtain 1-word phrasal contexts (1-PH: DIESE XXXXX) and 1-word list contexts (APFEL XXXXX).

### Procedure

Before entering the scanning room, participants performed a short practice session of the actual experiment on a desktop computer located just outside the MR unit area. None of the stimuli used in the instruction session were used during the experimental session. Once in the scanner, stimuli were presented visually using the software package Presentation® (Neurobehavioral Systems, Inc., Albany, CA, USA) with a Sanyo PLC-XP50L LCD XGA (Sanyo Electric Co., Ltd., Moriguchi, Japan; pixels = 1024 × 768; refresh rate = 100 Hz) back-mirror system mounted on the head-coil. Although the projector was already adjusted for minimal luminance, a white font/gray background was used, and preferred by all subjects. A mono-spaced font (Courier) was used (capitalized letters; 45 pt.). A single trial consisted of a white fixation cross which remained at the center of the screen until a random jitter of either 0 or 1000 ms after volume acquisition started the visual stimulation. Stimulus-onset-asynchrony was 8.6 s on average. All trials had a total duration of 900 ms and the items were presented sequentially on the screen one after the other (Supplementary Figure 1). Given that our stimulus construction was syllable-constrained, the first bi-syllabic word remained on the screen for 600 ms, while the second monosyllabic word/X string lasted 300 ms. As soon as the fixation cross reappeared, immediately after the second item within the trial had been shown, subjects were requested to perform a simple sequence judgment task similar to the one used in Friederici et al. (2000a), as quickly as possible, by indicating via triple-choice button-pressing whether the two words together formed a phrase (e.g., DIESE FLIRK = yes), whether they did not form a phrase, but just a list of two nouns (APFEL FLIRK = no), or whether it was a trash trial with X strings (DIESE XXXXX/APFEL XXXXX = trash). We used a fully counter-balanced stimulus exposure across conditions, such that in half of the cases the determiner (or the noun) was followed by a pseudoword, and in the other half of the cases by a sequence of Xs. Therefore, subjects could not discriminate between conditions on occurrence of the first word of the trial, rather they were forced to pay attention to the second word to solve the task. Subjects were requested to use the right index finger, the right middle finger, and the right ring finger to accomplish the task. The order of both buttons and trials were fully randomized across subjects. Each experimental dataset collection lasted ~42 min.

### Behavioral Data Analysis

Mean reaction times for correct responses (RTs) and accuracy rates were calculated for each condition of each participant and were analyzed using a Two-way within-subject analysis of variance (ANOVA), with factors STRUCTURE (phrase: PH

vs. list: LS) and number of WORDS (2- vs. 1-word). Missing responses were counted as non-correct responses.

## fMRI Data Acquisition

Functional images were acquired with a 3T whole-body Bruker Medspec 3000 Scanner. The functional data were acquired using a T2\*-weighted gradient-echo echo-planar-imaging (EPI) sequence, with the following parameters: TR = 2.0 s, TE = 30 ms, flip angle = 90°, FOV = 19.2 × 19.2 cm<sup>2</sup>, in-plane resolution = 3 × 3 mm<sup>2</sup>; data matrix = 64 × 64; slice thickness = 3 mm; interslice gap = 1 mm; number of slices = 30 (axial slices, parallel to AC-PC line/whole-brain coverage, ascending direction), number of volumes = 1270 volumes. T1-weighted 3D MP-RAGE (magnetization-prepared rapid gradient echo) images (Mugler and Brookeman, 1990)—TI = 650 ms; TR = 1300 ms; alpha = 10°; FOV = 256 × 240 mm—were previously acquired with a non-selective inversion pulse to be used for preprocessing of the functional data.

## Functional Imaging Data Analysis

Functional data were analyzed using the SPM8 software package (<http://www.fil.ion.ucl.ac.uk/spm/>). In the pre-processing session, subject-specific functional volumes were co-registered with corresponding structural T1-weighted images. Functional time series were further realigned to the first image to correct for motion artifacts, and resliced for timing correction. A gray-matter segmentation-based procedure was used for normalization to the standard MR template included in the SPM software package. A Gaussian filter of 8 mm<sup>3</sup> FWHM was used to smooth the data. A high pass filter of 128 s was used to attenuate slow global signal changes. These data entered in a number of analyses described below, they were also used for an additional analysis focusing on the insula and its subregions presented in a separate article (Zaccarella and Friederici, 2015).

## fMRI Whole-brain Data Analysis

The SPM8 software package was then used to perform a two-stage random-effects analysis to ensure result generalizability over the population level (Penny and Holmes, 2004). The first five volumes from each dataset were excluded to allow for magnetic saturation effect. Subject-specific general linear models were assessed using the hemodynamic response function from the SPM software (Friston et al., 1995). Single stimulus functions were modeled according to their timing onsets. Error trials and fillers trials were modeled as distinct conditions, and movement parameters were treated as regressors of no interest. Contrast estimates for the four experimental conditions (compared against the global mean) were obtained using first-level statistics. The contrast estimates were then used in a second-level within-subjects ANOVA to assess group contrasts. Statistical inferences were drawn at  $P < 0.05$ , with a Family-Wise Error (FWE) correction.

## Volume-of-Interest Analysis

Following our initial hypothesis that BA 44 is responsible for merge processing, we focused on the left infero-frontal region alone, by performing a finer-grained analysis in BA

44 to assess the specific effect of phrases compared to lists, directly (2-PH > 2-LS). The cytoarchitectonically-defined BA 44 from the maximum probability map (MPM; Supplementary Figure 2) of the SPM Anatomy Toolbox (Eickhoff et al., 2005) served as an independent search space to avoid selection bias (Kriegeskorte et al., 2009; Vul and Kanwisher, 2010). A small-volume correction (SVC) was used to threshold the results, at  $P < 0.05$ , FWE-corrected.

## Cluster-of-Interest Analysis

The multi-modal CBP map for BA 44 proposed in Clos et al. (2013) served as mask for the Cluster-of-Interest (COI) analysis. Here we first wanted to simply localize the cortical distribution of the active voxels to the contrast 2-PH > 2-LS, which we found at  $P < 0.05$ , FWE-corrected in the SVC analysis discussed above. This map, which is bounded by the same cytoarchitectonic region (MPM) that was used in the above SVC analysis, consists of five sub-regional BA 44 clusters comprising a posterior-dorsal cluster (C1), an anterior-dorsal cluster (C2), an anterior-ventral cluster (C3), a posterior-ventral cluster (C4), and an inferior frontal junction cluster (C5; see also Supplementary Figure 3). For our purposes, the activation mass obtained from the SVC analysis was first transformed into a binary image of zeroes (not-active voxels) and ones (active voxels), and then dot-multiplied with each cluster volume from the BA 44 parcellation map described above. Following this procedure we then counted the total number of ones (active-voxels) falling within each cluster to determine the overlapping region. Additionally, the five CBP clusters were further used as seed sub-regions to extract signal intensity, to evaluate the mean activity distribution of the syntactic binding effect across the different clusters. Mean signal extraction from the five clusters was done using Marsbar 0.41 for SPM (available at <http://marsbar.sourceforge.net>).

## Individual Peak Activity Distribution Analysis

Finally, we were interested in assessing whether at the individual level, the peak distribution of neural activity across individual subjects was homogeneously spread within BA 44, or rather gathered around one of the CBP clusters described above, therefore showing little variance over space. To evaluate cluster sensitivity at the subject level, we again used the map for BA 44 as a binary searchable space, and dot-multiplied it with the subject-specific contrast we obtained as {T} maps from first-level statistics, for the contrast (2-PH > 2-LS). For each subject, we then extracted a unique 3D coordinate maximum corresponding to one voxel. Each 3D coordinate was in turn localized as belonging to a particular cluster, using each BA 44 sub-regional cluster as independent mask, following an analogous counting procedure of the one described in the COI analysis above. From the resulting distribution, we first performed a standard chi-square distribution test. We then employed a randomization test of goodness-of-fit to strengthen, or possibly weaken, the significance of our cluster-sensitivity. We therefore drew 10,000 random samples from a population with our known proportions we obtained from the data, re-calculated the chi-square for each replicate sample, and then counted how many times a larger chi-square value was obtained during randomization



(McDonald, 2009). The proportion of replicates with chi-square values equal to or greater than the first observed value was then taken as final  $p$ -value. A threshold of  $p = 0.05$  (5% of times) was chosen.

## RESULTS

### Behavioral Results

No significant effect for accuracy was found. A significant effect for STRUCTURE [ $F_{(1, 21)} = 25.003$ ;  $p < 0.0001$ ] and an interaction between WORDS and STRUCTURE [ $F_{(1, 21)} = 5.896$ ;  $p < 0.05$ ] were found for the reaction time data. A series of paired  $t$ -tests revealed that subjects were slower for 2-LS compared to 2-PH ( $t = 3.93$ ;  $p < 0.001$ ). An almost significant difference for LS: 2 > 1 was found ( $p = 0.059$ ), while there was no significant difference for the contrast PH: 2 > 1 ( $p > 0.1$ ; Supplementary Figure 4).

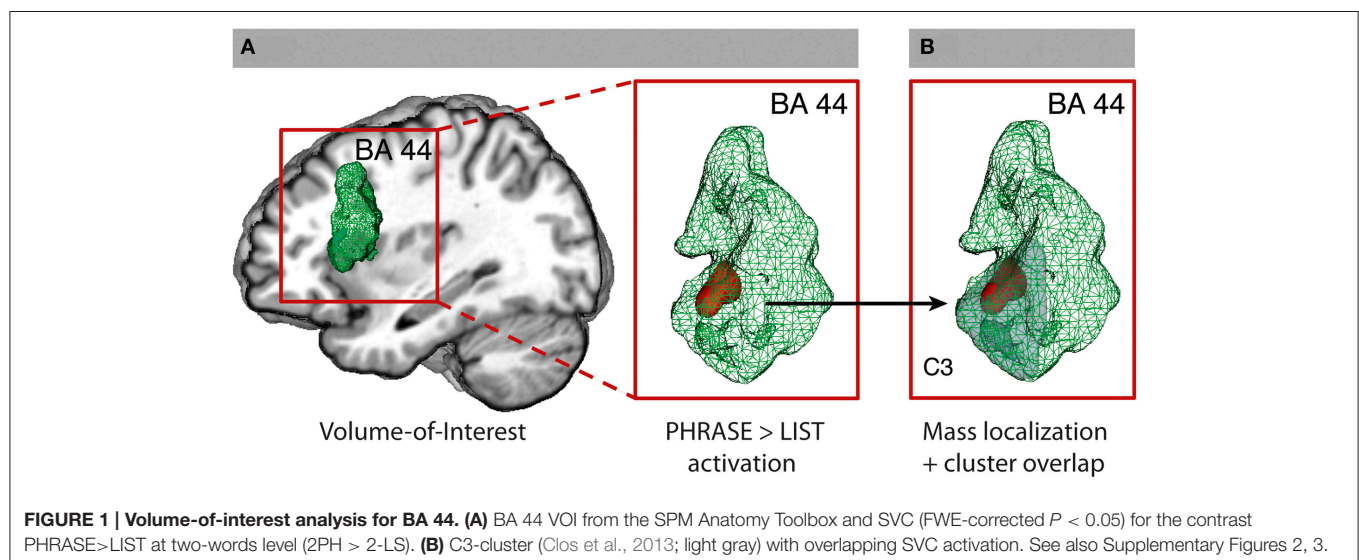
### Whole-brain Analysis

We found a main effect of WORDS in the left and right FOP/adINS at  $x = -33$ ;  $y = 23$ ;  $z = -2$  and  $x = 36$ ;  $y = 23$ ;  $z = -2$ , respectively, and in the left BA 44 at  $x = -48$ ;  $y = 11$ ;  $z = 7$ . The main effect of STRUCTURE, as well the WORDS  $\times$  STRUCTURE interaction did not yield any significant clusters that survived standard statistical thresholds (Supplementary Table 1).

### Volume-of-Interest Analysis

The hypothesis-driven analysis in BA 44 revealed a significant cluster for phrase compared to list at two-words level (2-PH > 2-LS) in the ventral anterior part of BA 44 at  $x = -48$ ;  $y = 17$ ;  $z = 16$ , using a small volume correction analysis (SVC) in the area. No significant voxel was found active for the opposite contrast (2-PH < 2-LS), even at more liberal thresholds. To note, prior direct comparison between phrase and list at two-words level, we run an ANOVA with factors WORDS and STRUCTURE within BA 44 to gain information about a possible interaction within the

region. Interestingly, we detected within the region the 3D voxel showing the highest peak by downloading the unthresholded WORDS  $\times$  STRUCTURE activation map we obtained from the SPM group-averaged output ( $x = -51$ ;  $y = 20$ ;  $z = 13$ ). From this 3D coordinates we then extracted signal intensity for all four conditions to verify whether an interaction between WORDS and STRUCTURE would have survived statistical control. To note, we found a significant interaction between the two factors at  $p = 0.039$  level [ $F_{(1, 21)} = 4.83$ ]. Direct comparison between 2-PH and 1-PH was significant at  $p < 0.001$  level ( $t = 3.87$ ), as it was between 2-PH and 2-LS at  $p = 0.007$  level ( $t = 2.94$ ). Direct comparison between 2-LS and 1-LS was not significant ( $t = 1.67$ ;  $p = 0.11$ ; see **Figure 1A**. See also Supplementary Table 1 and Supplementary Figure 5 for the interaction effect). We further performed additional independent analyses in the other regions that were found to be active for the main effect of WORDS, to evaluate whether syntactic binding was specifically performed in BA 44 alone, or whether additional portions of the cortex were also involved. There was no difference between phrases and lists in the other regions under analysis (see also Appendix B–Volume-of-interest Analysis for more information). To gain further exploratory indication on the relative functional contribution of phrases and lists to the main effect of WORDS in the inferior frontal regions, and to verify the results we obtained from the SVC analysis, we went then back to our full brain datasets and performed two distinct planned contrasts at the two levels of the STRUCTURE factor. For the contrast lists vs. one-word list condition (LS: 2 > 1) we found activity in the opercula only at  $x = -33$ ;  $y = 23$ ;  $z = -2$  and  $x = 36$ ;  $y = 23$ ;  $z = -2$ . For the contrast phrases vs. one-word phrase condition (PH: 2 > 1) we found additional recruitment of the ventral-anterior portion of left BA 44 at  $x = -51$ ;  $y = 11$ ;  $z = 7$ , which, together with the LS: 2 > 1 results, suggests a stronger involvement of BA 44 for phrases than for lists. At the whole-brain level, however there was no significant difference in activity for the contrast phrases and lists (2-PH > 2-LS) using standard threshold methods.



### Cluster-of-Interest Analysis within BA 44

All voxels active for the contrast phrase vs. list at two-words level (2-PH > 2-LS) in the SVC in BA 44 fell within the anterior-ventral cluster C3 (100% overlap; 12/12 voxels; see **Figure 1B**). Remarkably, paired *t*-tests for signal intensity revealed a significant difference in activity between phrases and lists (2-PH > 2-LS) in the C3 cluster only [ $t_{(21)} = 2.97, p = 0.007$ , surviving Bonferroni-correction for the number of tests [ $p = (0.05/5 \text{ tests}) = 0.01$ ], and in no other sub-region (see **Figure 2** and Supplementary Figure 6).

### Individual Peak Activity Distribution Analysis

A strong cluster-sensitive distribution of the individual peak activity for the contrast (2-PH > 2-LS) was found in C3, as compared to the other BA 44 sub-regions [ $\chi^2_{(4)} = 13.45, p = 0.009$ , confirmed after 10,000 randomization tests for goodness-of-fit; see **Figure 3**].

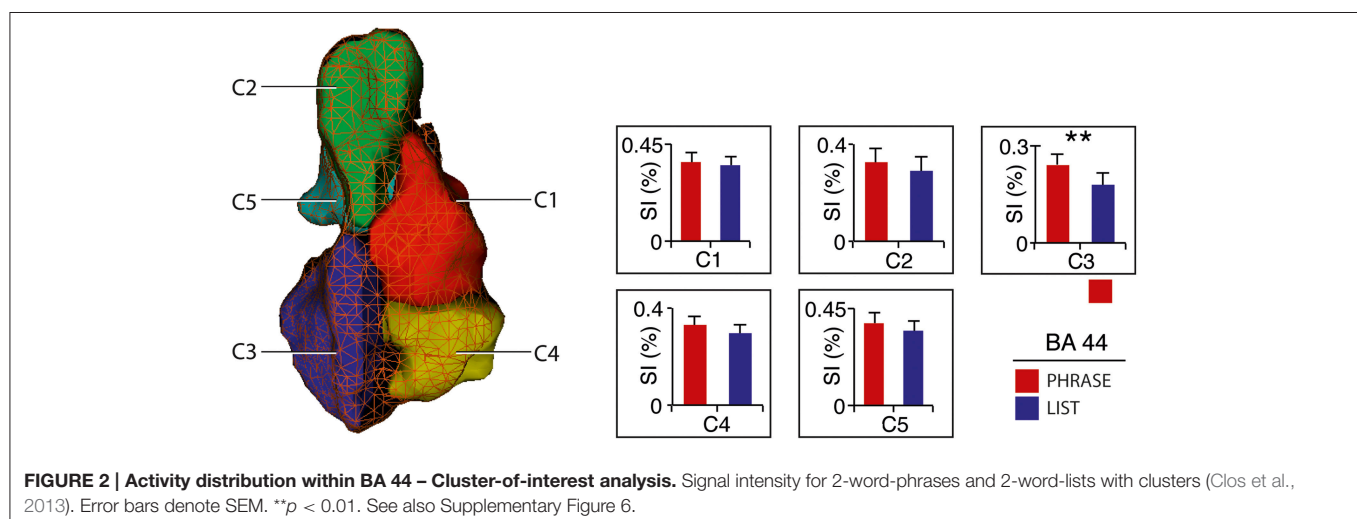
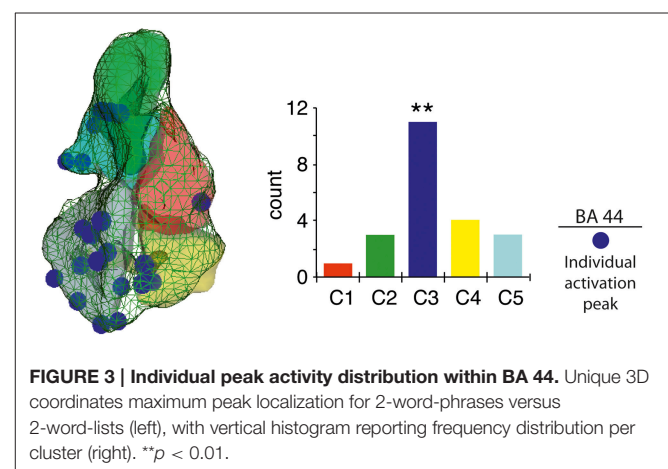
## DISCUSSION

The goal of the present fMRI study was to identify the neuroanatomical basis of the most fundamental syntactic computation, which is at the root of all natural languages (Chomsky, 1995; Berwick et al., 2013). This basic computation, called *merge*, which binds two words together syntactically, allows to build up syntactic structures with increasing hierarchy. Here we found that this most basic process of syntactic binding corresponded to increased activity in a most confined brain region, i.e., the anterior section of the ventral left pars opercularis, BA 44 at the posterior part of Broca's area. Conversely, a phylogenetically older area, the FOP/adINS, was found to be equally active for both phrasal structures and unstructured word-lists, not discriminating between these.

With respect to the FOP/aINS, the present analysis indicates that this area's function—previously identified by a region-of-interest analysis of the insula—is also identifiable at the whole brain analysis, thereby extending results from the same study (Zaccarella and Friederici, 2015). Its involvement revealed from

the contrast between two- and one-word stimuli may reflect word-accumulation processes during which the categorical information and the grammatical status of the word is first accessed (Friederici et al., 2000b) and then shortly maintained on hold (Grasby et al., 1994), before further processing takes place. A similar activity pattern found for both phrases and lists is not surprising, given the low degree of functional specialization that the FOP/adINS has inside the language processing system (Saygin et al., 2004; Mutschler et al., 2009). This lower specialization of the left FOP/adINS compared to BA 44 also finds support in another fMRI study which showed that this area, in contrast to BA 44, was not able to distinguish between grammar types, but only able to detect an error in the order of syllables in sequences (Friederici et al., 2006a).

In BA44 we found that the basic operation of syntactic merge was sensitive to a specific cluster within the region, such that only the anterior-ventral cluster as one among five sub-regions discriminated between phrases and lists. Crucially, we discovered that the localization of this activity within the same anterior-ventral cluster was highly consistent across participants. While a



previous functional connectivity-based meta-analysis described this subregion as being associated with all kinds of language processes (Clos et al., 2013), the present study is the first to delineate this subregion from other regions within BA 44 in its function in language. This particular subregion appears to be responsible for the most basic syntactic computation at the root of all syntactic hierarchies. Remarkably, within neurolinguistic literature, the contribution of parts of Broca's area, in particular BA 44, to syntactic processing has mostly been discussed in terms of syntactic complexity at the sentential level, since the area was found to be crucial for the processing of syntactically more complex sentential hierarchies, compared to simpler ones (Röder et al., 2002; Friederici et al., 2006a,b; Bahlmann et al., 2008). The present data are in line with the view of BA 44 being activated as a function of structural hierarchy, *but* they clearly go beyond this view by demonstrating that the most basic syntactic computation upon which more complex hierarchies are built, can be neuroanatomically located in a sub-region of BA 44. This means that because both complex and simpler linguistic hierarchies necessarily share the same computational merging algorithm, BA 44 activates as a function of structural hierarchy regardless of the linguistic complexity itself.

The sustained cluster-sensitivity of the merge computation in the anterior-ventral cluster of BA 44 both at the group and the individual level points toward a fundamental and constrained nature of *merge* at the neural level. While we acknowledge that the present work only tested for one single language, the low interindividual spatial variability we found across our representative set of subjects, might be taken as a first approximate indication in favor of the fundamental character of the computation itself. The present finding of a sub-regional specificity and invariability of the most basic process closely resembles the neural organization of other basic sensory processes, as for example in the visual system (Downing et al., 2006), while the low inter-individual variability of the basic processes gives rise to the assumption that their function to structure relation is predetermined. In this respect, because of the essential nature of *merge* as being the shared computation across all human languages, future studies should systematically focus their attention on the neural implementation of very basic linguistic processes in multiple languages. This would ultimately prove whether the universality of *merge* at theoretical level (Chomsky, 1995) adequately corresponds to some neuroanatomical generalizability at the neural level suggested here.

From an evolutionary perspective, no clear evidence for a human-like language syntax in nonhuman species has been presented so far (Fitch and Hauser, 2004; Bolhuis et al., 2014). Structural studies have shown that the FOP is a phylogenetically older cortex, which is fully represented in monkeys (Sanides, 1962), while BA 44 seems to be more expressed in humans than in monkeys, in which it plays a role in orofacial somatomotor processes (Petrides et al., 2005). The activation pattern we report in our study closely resembles the one proposed for human and

non-human artificial grammar processing in the adult brain, in which violations to transition probabilities are found to activate FOP, while violations to more rules activate BA 44 (Friederici et al., 2006a). This functional split between the labor of FOP/aINS and that of BA 44 is particularly intriguingly if put into relation with recent theoretical linguistic models, which propose that the specificity of *merge* in language should reside in the property that words have to create constituents where the lexical label of the single dominant word (e.g., determiner) is reflected as a hierarchical influence onto the newly created syntactic constituent (e.g., determiner phrase; Boeckx, 2010; Chomsky, 2013; Murphy, 2015). At the interface between linguistic theory and neurolinguistics, the *merge* mechanism would then consist of two phases: one in which linguistic elements are strung together without any hierarchical dimension in the FOP/aINS, and a labeling phase in which the dominant lexical element transforms the string into a hierarchically labeled syntactic structure in the anterior-dorsal BA 44. We believe that this speculation can constitute a testable model for the evolution of the language faculty, in which behavioral, functional and anatomical data can be put together in a comparative perspective within across human and animal species (Murphy, 2015).

## CONCLUSION

The sub-anatomical specificity for the process of syntactic binding, called *merge*, has a strict neural basis in the anterior-ventral cluster of BA 44. The profoundly constrained regional localization of this syntactic operation converges on the conclusion that the computation at the root of our syntactic knowledge has strict neural basis. The constraint localization of the activity and its consistency across the participants point toward the fundamental neurobiological nature of the operation of *merge* itself, thereby providing a novel view on the relation between linguistic theory and neurobiology.

## ACKNOWLEDGMENTS

We wish to thank all participants for their involvement, Mandy Jochemko, Anke Kummer, and Simone Wipper for MRI data acquisition, Lars Meyer for his valuable support during the analysis of the functional data, Michiru Makuuchi, Katharina von Kriegstein, and Isabell Wartenburger for their comments on earlier versions of the manuscript, Francie Manhardt for evaluating the stimulus material, and Kerstin Flake for helping us creating the figures. This work was supported by a grant of the European Research Council (ERC-2010-AdG 20100407 awarded to AF) and by the Berlin School of Mind and Brain (EZ).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpsyg.2015.01818>



## REFERENCES

- Adger, D. (2003). *Core Syntax: A Minimalist Approach*. New York, NY: Oxford University Press.
- Amunts, K., Lenzen, M., Friederici, A. D., Schleicher, A., Morosan, P., Palomero-Gallagher, N., et al. (2010). Broca's region: novel organizational principles and multiple receptor mapping. *PLoS Biol.* 8:e1000489. doi: 10.1371/journal.pbio.1000489
- Amunts, K., Schleicher, A., Bürgel, U., Mohlberg, H., Uylings, H. B., and Zilles, K. (1999). Broca's region revisited: cytoarchitecture and intersubject variability. *J. Comp. Neurol.* 412, 319–341.
- Baayen, R. H., Piepenbrock, R., and Gulikers, L. (1995). *The CELEX Lexical Database*. Philadelphia, PA: Linguistics Data Consortium; University of Pennsylvania.
- Bahlmann, J., Schubotz, R. I., and Friederici, A. D. (2008). Hierarchical artificial grammar processing engages Broca's area. *Neuroimage* 42, 525–534. doi: 10.1016/j.neuroimage.2008.04.249
- Barsalou, L. W. (1982). Context-independent and context-dependent information in concepts. *Mem. Cognit* 10, 82–93. doi: 10.3758/BF03197629
- Bemis, D. K., and Pykkänen, L. (2011). Simple composition: a magnetoencephalography investigation into the comprehension of minimal linguistic phrases. *J. Neurosci.* 31, 2801–2814. doi: 10.1523/JNEUROSCI.5003-10.2011
- Bemis, D. K., and Pykkänen, L. (2013). Flexible composition: MEG evidence for the deployment of basic combinatorial linguistic mechanisms in response to task demands. *PLoS ONE* 8:e73949. doi: 10.1371/journal.pone.0073949
- Ben-Shachar, M., Hendler, T., Kahn, I., Ben-Bashat, D., and Grodzinsky, Y. (2003). The neural reality of syntactic transformations: evidence from functional magnetic resonance imaging. *Psychol. Sci.* 14, 433–440. doi: 10.1111/1467-9280.01459
- Ben-Shachar, M., Palti, D., and Grodzinsky, Y. (2004). Neural correlates of syntactic movement: converging evidence from two fMRI experiments. *Neuroimage* 21, 1320–1336. doi: 10.1016/j.neuroimage.2003.11.027
- Berwick, R. C., Friederici, A. D., Chomsky, N., and Bolhuis, J. J. (2013). Evolution, brain, and the nature of language. *Trends Cogn. Sci.* 17, 89–98. doi: 10.1016/j.tics.2012.12.002
- Boeckx, C. (2010). *Language in Cognition: Uncovering Mental Structures and the Rules Behind them*. Hoboken, NJ: Wiley-Blackwell.
- Bolhuis, J. J., Tattersall, I., Chomsky, N., and Berwick, R. C. (2014). How could language have evolved? *PLoS Biol.* 12:e1001934. doi: 10.1371/journal.pbio.1001934
- Bornkessel, I., Zysset, S., Friederici, A. D., von Cramon, D. Y., and Schlesewsky, M. (2005). Who did what to whom? The neural basis of argument hierarchies during language comprehension. *Neuroimage* 26, 221–233. doi: 10.1016/j.neuroimage.2005.01.032
- Bottini, G., Corcoran, R., Sterzi, R., Paulesu, E., Schenone, P., Scarpa, P., et al. (1994). The role of the right hemisphere in the interpretation of figurative aspects of language. A positron emission tomography activation study. *Brain* 117, 1241–1253. doi: 10.1093/brain/117.6.1241
- Brodman, K. (1909). *Vergleichende Lokalisationslehre der Großhirnrinde*. Leipzig: Barth JA.
- Caplan, D., Chen, E., and Waters, G. (2008). Task-dependent and task-independent neurovascular responses to syntactic processing. *Cortex* 44, 257–275. doi: 10.1016/j.cortex.2006.06.005
- Chomsky, N. (1995). *The Minimalist Program*. Cambridge: MIT Press.
- Chomsky, N. (1999). *Derivation by Phase*. Cambridge: MIT Press.
- Chomsky, N. (2013). Problems of projection. *Lingua* 130, 33–49. doi: 10.1016/j.lingua.2012.12.003
- Clos, M., Amunts, K., Laird, A. R., Fox, P. T., and Eickhoff, S. B. (2013). Tackling the multifunctional nature of Broca's region meta-analytically: co-activation-based parcellation of area 44. *Neuroimage* 83, 174–188. doi: 10.1016/j.neuroimage.2013.06.041
- Constable, R. T., Pugh, K. R., Berroya, E., Mencl, W. E., Westerveld, M., Ni, W., et al. (2004). Sentence complexity and input modality effects in sentence comprehension: an fMRI study. *Neuroimage* 22, 11–21. doi: 10.1016/j.neuroimage.2004.01.001
- Cooke, A., Zurif, E. B., DeVita, C., Alsop, D., Koenig, P., Detre, J., et al. (2002). Neural basis for sentence comprehension: grammatical and short-term memory components. *Hum. Brain Mapp.* 15, 80–94. doi: 10.1002/hbm.10006
- Del Prado, P., and Pykkänen, L. (2014). MEG evidence for conceptual combination but not numeral quantification in the left anterior temporal lobe during language production. *Front. Psychol.* 5:524. doi: 10.3389/fpsyg.2014.00524
- Downing, P. E., Chan, A. W., Peelen, M. V., Dodds, C. M., and Kanwisher, N. (2006). Domain specificity in visual cortex. *Cereb. Cortex* 16, 1453–1461. doi: 10.1093/cercor/bhj086
- Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., et al. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage* 25, 1325–1335. doi: 10.1016/j.neuroimage.2004.12.034
- Fiebach, C. J., Schlesewsky, M., Lohmann, G., von Cramon, D. Y., and Friederici, A. D. (2005). Revisiting the role of Broca's area in sentence processing: syntactic integration versus syntactic working memory. *Hum. Brain Mapp.* 24, 79–91. doi: 10.1002/hbm.20070
- Fitch, W. T., and Hauser, M. D. (2004). Computational constraints on syntactic processing in a nonhuman primate. *Science* 303, 377–380. doi: 10.1126/science.1089401
- Friederici, A. D. (2006). Broca's area and the ventral premotor cortex in language: functional differentiation and specificity. *Cortex* 42, 472–475. doi: 10.1016/S0010-9452(08)70380-0
- Friederici, A. D. (2011). The brain basis of language processing: from structure to function. *Physiol. Rev.* 91, 1357–1392. doi: 10.1152/physrev.00006.2011
- Friederici, A. D., Bahlmann, J., Heim, S., Schubotz, R. I., and Anwander, A. (2006a). The brain differentiates human and non-human grammars: functional localization and structural connectivity. *Proc. Natl. Acad. Sci. U.S.A.* 103, 2458–2463. doi: 10.1073/pnas.0509389103
- Friederici, A. D., Fiebach, C. J., Schlesewsky, M., Bornkessel, I. D., and von Cramon, D. Y. (2006b). Processing linguistic complexity and grammaticality in the left frontal cortex. *Cereb. Cortex* 16, 1709–1717. doi: 10.1093/cercor/bhj106
- Friederici, A. D., Meyer, M., and von Cramon, D. Y. (2000a). Auditory language comprehension: an event-related fMRI study on the processing of syntactic and lexical information. *Brain Lang.* 75, 289–300. doi: 10.1006/brln.2000.2313
- Friederici, A. D., Opitz, B., and von Cramon, D. Y. (2000b). Segregating semantic and syntactic aspects of processing in the human brain: an fMRI investigation of different word types. *Cereb. Cortex* 10, 698–705. doi: 10.1093/cercor/10.7.698
- Friston, K. J., Holmes, A. P., Worsley, J.-P., Frith, C. D., and Frackowiak, R. S. (1995). Statistical parametric maps in functional imaging: a general linear approach. *Hum. Brain Mapp.* 2, 189–210. doi: 10.1002/hbm.460020402
- Grasby, P. M., Frith, C. D., Friston, K. J., Simpson, J., Fletcher, P. C., Frackowiak, R. S., et al. (1994). A graded task approach to the functional mapping of brain areas implicated in auditory-verbal memory. *Brain* 117, 1271–1282. doi: 10.1093/brain/117.6.1271
- Grewe, T., Bornkessel, I., Zysset, S., Wiese, R., von Cramon, D. Y., and Schlesewsky, M. (2005). The emergence of the unmarked: a new perspective on the language-specific function of Broca's area. *Hum. Brain Mapp.* 26, 178–190. doi: 10.1002/hbm.20154
- Hagoort, P., and Indefrey, P. (2014). The neurobiology of language beyond single words. *Annu. Rev. Neurosci.* 37, 347–362. doi: 10.1146/annurev-neuro-071013-013847
- Humphries, C., Binder, J. R., Medler, D. A., and Liebenthal, E. (2006). Syntactic and semantic modulation of neural activity during auditory sentence comprehension. *J. Cogn. Neurosci.* 18, 665–679. doi: 10.1162/jocn.2006.18.4.665
- Humphries, C., Binder, J. R., Medler, D. A., and Liebenthal, E. (2007). Time course of semantic processes during sentence comprehension: an fMRI study. *Neuroimage* 36, 924–932. doi: 10.1016/j.neuroimage.2007.03.059
- Humphries, C., Love, T., Swinney, D., and Hickok, G. (2005). Response of anterior temporal cortex to syntactic and prosodic manipulations during sentence processing. *Hum. Brain Mapp.* 26, 128–138. doi: 10.1002/hbm.20148
- Just, M. A., Carpenter, P. A., Keller, T. A., Eddy, W. F., and Thulborn, K. R. (1996). Brain activation modulated by sentence comprehension. *Science* 274, 114–116. doi: 10.1126/science.274.5284.114

- Kinno, R., Kawamura, M., Shioda, S., and Sakai, K. L. (2008). Neural correlates of noncanonical syntactic processing revealed by a picture-sentence matching task. *Hum. Brain Mapp.* 29, 1015–1027. doi: 10.1002/hbm.20441
- Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S., and Baker, C. I. (2009). Circular analysis in systems neuroscience: the dangers of double dipping. *Nat. Neurosci.* 12, 535–540. doi: 10.1038/nn.2303
- McDonald, J. H. (2009). *Handbook of Biological Statistics, 2nd Edn.* Baltimore, MD: Sparky House Publishing.
- Moro, A., Tettamanti, M., Perani, D., Donati, C., Cappa, S. F., and Fazio, F. (2001). Syntax and the brain: disentangling grammar by selective anomalies. *Neuroimage* 13, 110–118. doi: 10.1006/nimg.2000.0668
- Mugler, J. P. III, and Brookeman, J. R. (1990). Three-dimensional magnetization-prepared rapid gradient-echo imaging (3D MP RAGE). *Magn. Reson. Med.* 15, 152–157. doi: 10.1002/mrm.1910150117
- Murphy, E. (2015). Labels, cognomes, and cyclic computation: an ethological perspective. *Front. Psychol.* 6:715. doi: 10.3389/fpsyg.2015.00715
- Mutschler, I., Wieckhorst, B., Kowalewski, S., Derix, J., Wentlandt, J., Schulze-Bonhage, A., et al. (2009). Functional organization of the human anterior insular cortex. *Neurosci. Lett.* 457, 66–70. doi: 10.1016/j.neulet.2009.03.101
- Newman, S. D., Ikuta, T., and Burns, T. Jr. (2010). The effect of semantic relatedness on syntactic analysis: an fMRI study. *Brain Lang.* 113, 51–58. doi: 10.1016/j.bandl.2010.02.001
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4
- Pallier, C., Devauchelle, A. D., and Dehaene, S. (2011). Cortical representation of the constituent structure of sentences. *Proc. Natl. Acad. Sci. U.S.A.* 108, 2522–2527. doi: 10.1073/pnas.1018711108
- Penny, W. D., and Holmes, A. P. (2004). “Random effect analysis,” in *Human Brain Function*, eds R. S. J. Frackowiak, J. T. Ashburner, W. D. Penny, S. Zeki, K. J. Friston, C. D. Frith, R. J. Dolan, and C. J. Price (San Diego, CA: Elsevier), 843–850.
- Petrides, M., Cadoret, G., and Mackey, S. (2005). Orofacial somatomotor responses in the macaque monkey homologue of Broca’s area. *Nature* 435, 1235–1238. doi: 10.1038/nature03628
- Röder, B., Stock, O., Neville, H., Bien, S., and Rosler, F. (2002). Brain activation modulated by the comprehension of normal and pseudo-word sentences of different processing demands: a functional magnetic resonance imaging study. *Neuroimage* 15, 1003–1014. doi: 10.1006/nimg.2001.1026
- Sanides, F. (1962). *Die Architektonik des Menschlichen Stirnhirns*. Berlin: Springer.
- Santi, A., and Grodzinsky, Y. (2007). Working memory and syntax interact in Broca’s area. *Neuroimage* 37, 8–17. doi: 10.1016/j.neuroimage.2007.04.047
- Santi, A., and Grodzinsky, Y. (2010). fMRI adaptation dissociates syntactic complexity dimensions. *Neuroimage* 51, 1285–1293. doi: 10.1016/j.neuroimage.2010.03.034
- Saygin, A. P., Wilson, S. M., Dronkers, N. F., and Bates, E. (2004). Action comprehension in aphasia: linguistic and non-linguistic deficits and their lesion correlates. *Neuropsychologia* 42, 1788–1804. doi: 10.1016/j.neuropsychologia.2004.04.016
- Stowe, L. A., Broere, C. A., Paans, A. M., Wijers, A. A., Mulder, G., Vaalburg, W., et al. (1998). Localizing components of a complex task: sentence processing and working memory. *Neuroreport* 9, 2995–2999. doi: 10.1097/00001756-199809140-00014
- Stowe, L. A., Paans, A. M., Wijers, A. A., Zwarts, F., Mulder, G., and Vaalburg, W. (1999). Sentence comprehension and word repetition: a positron emission tomography investigation. *Psychophysiology* 36, 786–801. doi: 10.1111/1469-8986.3660786
- Stromswold, K., Caplan, D., Alpert, N., and Rauch, S. (1996). Localization of syntactic comprehension by positron emission tomography. *Brain Lang.* 52, 452–473. doi: 10.1006/brln.1996.0024
- Vandenberghe, R., Nobre, A. C., and Price, C. J. (2002). The response of left temporal cortex to sentences. *J. Cogn. Neurosci.* 14, 550–560. doi: 10.1162/08998290260045800
- Vigneau, M., Beaucousin, V., Hervé, P. Y., Duffau, H., Crivello, F., Houdé, O., et al. (2006). Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. *Neuroimage* 30, 1414–1432. doi: 10.1016/j.neuroimage.2005.11.002
- Vul, E., and Kanwisher, N. (2010). “Begging the question: the non-independence error in fMRI data analysis,” in *Foundational Issues in Human Brain Mapping*, eds S. J. Hanson and M. Bunzl (Cambridge: MIT Press), 71–91.
- Zaccarella, E., and Friederici, A. D. (2015). Reflections of word processing in the insular cortex: a sub-regional parcellation based functional assessment. *Brain Lang.* 142, 1–7. doi: 10.1016/j.bandl.2014.12.006
- Zilles, K., and Amunts, K. (2009). Receptor mapping: architecture of the human cerebral cortex. *Curr. Opin. Neurol.* 22, 331–339. doi: 10.1097/WCO.0b013e32832d95db

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Zaccarella and Friederici. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# ‘Syntactic Perturbation’ During Production Activates the Right IFG, but not Broca’s Area or the ATL

William Matchin<sup>1\*</sup> and Gregory Hickok<sup>2</sup>

<sup>1</sup> Cognitive Neuroscience of Language Laboratory, Department of Linguistics, University of Maryland, College Park, MD, USA, <sup>2</sup> Auditory and Language Neuroscience Laboratory, Department of Cognitive Sciences, University of California, Irvine, Irvine, CA, USA

## OPEN ACCESS

### Edited by:

Cedric Boeckx,  
Catalan Institution for Research  
and Advanced Studies/Universitat  
de Barcelona, Spain

### Reviewed by:

Olaf Hauk,  
MRC Cognition and Brain Sciences  
Unit, UK  
Ina Bornkessel-Schlesewsky,  
University of South Australia, Australia

### \*Correspondence:

William Matchin  
wmatchin@umd.edu

### Specialty section:

This article was submitted to  
Language Sciences,  
a section of the journal  
Frontiers in Psychology

**Received:** 12 December 2015

**Accepted:** 05 February 2016

**Published:** 23 February 2016

### Citation:

Matchin W and Hickok G (2016)  
‘Syntactic Perturbation’ During  
Production Activates the Right IFG,  
but not Broca’s Area or the ATL.  
Front. Psychol. 7:241.  
doi: 10.3389/fpsyg.2016.00241

Research on the neural organization of syntax – the core structure-building component of language – has focused on Broca’s area and the anterior temporal lobe (ATL) as the chief candidates for syntactic processing. However, these proposals have received considerable challenges. In order to better understand the neural basis of syntactic processing, we performed a functional magnetic resonance imaging experiment using a constrained sentence production task. We examined the BOLD response to sentence production for active and passive sentences, unstructured word lists, and *syntactic perturbation*. Perturbation involved cued restructuring of the planned syntax of a sentence mid utterance. Perturbation was designed to capture the effects of syntactic violations previously studied in sentence comprehension. Our experiment showed that Broca’s area and the ATL did not exhibit response profiles consistent with syntactic operations – we found no increase of activation in these areas for sentences > lists or for perturbation. Syntactic perturbation activated a cortical-subcortical network including robust activation of the right inferior frontal gyrus (RIFG). This network is similar to one previously shown to be involved in motor response inhibition. We hypothesize that RIFG activation in our study and in previous studies of sentence comprehension is due to an inhibition mechanism that may facilitate efficient syntactic restructuring.

**Keywords:** syntax, sentence processing, language, fMRI, inferior frontal gyrus, Broca’s area, ATL, production

## INTRODUCTION

Language can be analyzed as a cognitive faculty consisting of several components, including a core structure-building system – *syntax* – that operates over stored lexical atoms (Chomsky, 1982, 1995; Hauser et al., 2002). Much work attempting to localize syntactic operations has focused on Broca’s area (Stromswold et al., 1996; Hagoort, 2005; Grodzinsky and Santi, 2008; Friederici, 2011) in the left inferior frontal gyrus (LIFG; Brodmann areas 44 and 45) and the anterior temporal lobe (ATL; Rogalsky and Hickok, 2009; Bemis and Pytkkanen, 2011; Brennan et al., 2012). However, the response profile of Broca’s area during sentence comprehension appears to be more compatible with a domain-general function such as working memory or cognitive control than with syntax (Kaan and Swaab, 2002; Novick et al., 2005; Rogalsky and Hickok, 2011; Bornkessel-Schlesewsky and Schlewsky, 2013), although this is still a hotly debated issue (Hickok and Rogalsky, 2011; Fedorenko et al., 2012b). Similarly, recent neuroimaging and neuropsychological studies have implicated the ATL in semantic rather than syntactic processes (Rogalsky and Hickok, 2009; Pallier et al., 2011; Del Prato and Pytkkanen, 2014; Wilson et al., 2014).

We chose to perform a functional magnetic resonance imaging (fMRI) experiment during sentence production to contribute to this debate. Sentence production studies in fMRI and magnetoencephalography (MEG) have revealed large overlap with the activation patterns found in comprehension, suggesting that similar neural networks underlie sentence processing in both modalities (Braun et al., 2001; Blank et al., 2002; Haller et al., 2005; Golestani et al., 2006; Menenti et al., 2011; Segaert et al., 2012; Del Prato and Pykkänen, 2014; Pykkänen et al., 2014). These studies are informative with respect to the neurobiology of sentence production; however, few production studies have manipulated syntactic variables compared to the vast literature on syntactic processing in comprehension. Syntax production studies will be important to provide complementary evidence to the comprehension literature to better understand syntactic processing in the brain.

We attempted to parallel the effects of syntactic violations that several researchers have used to study syntax in comprehension (e.g., Embick et al., 2000; Moro et al., 2001; Friederici et al., 2003). However, there are significant obstacles in extending the violation approach to production. Instructing subjects to produce artificial syntactic errors means that subjects will expect the upcoming violation. This may eliminate the effect of interest because expectation lessens the strength of the neural response to syntactic violations (Lau et al., 2006). The short time constraints of fMRI make difficult capturing infrequent natural errors or using a paradigm to induce subjects to produce them (Ferreira and Swets, 2005). Because of these reasons, we forced subjects to intermittently and unexpectedly switch their planned syntactic structure mid-utterance. The logic is that switching structures increases demands on the neural resources involved in syntactic processing (as well for other mechanisms). We expected to capture this effect in the blood-oxygen level-dependent (BOLD) response during scanning.

In the present study, we controlled the syntactic structure of the subjects' utterances with a constrained sentence elicitation task, similar to Caplan and Hanna (1998). To induce syntactic restructuring, we borrowed the target perturbation paradigm from motor control research (Paulignan et al., 1991; Elliott et al., 1995; Izawa et al., 2008). In this paradigm subjects attempt to hit a target, e.g., reaching from one point to another on a screen. On most trials the subject's target and/or sensory feedback remain constant throughout the trial. On a smaller proportion of random trials, the subject's target or sensory feedback is altered mid-movement. For example, the target location changes, or a force is applied to the subject's arm. On such trials, the subject must adapt and correct the movement trajectory online to reach the goal. We adapted this approach to syntax, dubbing our paradigm "syntactic perturbation."

We trained subjects to produce either active sentences (e.g., *Susan is following Charlie*) or passive sentences (e.g., *Charlie is being followed by Susan*). On most trials (standard trials, 80%) subjects did not switch their planned structure. On a smaller proportion of random trials (switch trials, 20%) a cue prompted subjects to switch structures mid-utterance. In other words, on switch trials the cue prompted the subject to switch from active to passive or from passive to active. This task was artificial,

raising questions about the ecological validity of our experiment – such considerations should be kept in mind when evaluating the results. However, we assumed subjects would update their syntactic structure regardless of this artificial nature.

The key assumptions of our experiment are the following: (1) the planned syntactic structure of an utterance is built in advance of speech production (at least for mono-clausal active and passive sentences), and (2) this plan can be dynamically updated during speech production. The first assumption is supported by the fact that juxtaposition errors often occur for words or phrases of the same syntactic category and from the same syntactic position (Fromkin, 1971) – to account for this regularity, speakers must have built the syntactic structure in advance of articulation. The second assumption is supported by an experiment showing that speakers decrease their rate of speech predictively if the structure they ultimately utter contains a syntactic violation (Ferreira and Swets, 2005).

Our design consisted of two main contrasts: STRUCTURE (sentences > word lists) and PERTURBATION (switch trials > standard trials), and one secondary contrast: COMPLEXITY (passive sentences > active sentences). Using our novel paradigm in production rather than comprehension, we examined the response profile of these contrasts in areas traditionally associated with syntax. Our main goal was to further inform the debate on the role of these regions in syntactic processing. Also, we believe that the discussion of the neurobiology of syntax has focused overwhelmingly on the ATL and Broca's area because of the repeated use of similar experimental manipulations in comprehension. We sought to determine whether our experiment, differing in modality and task, found activation in areas outside of these regions, potentially indicating a role for them in syntactic processing. We discuss our predictions for each of these contrasts in turn.

The contrast of sentence > word lists in comprehension has frequently revealed activation in the ATL, often bilaterally but also left lateralized (e.g., Mazoyer et al., 1993; Humphries et al., 2005; Rogalsky and Hickok, 2009). Two recent MEG studies found increased activation for preparation to produce two-word phrases compared to production of single words (Pykkänen et al., 2014) and two-word lists (Del Prato and Pykkänen, 2014). Current research supports a semantic interpretation of ATL function that drives these effects. We expected that the contrast of STRUCTURE in our study would also activate the ATL because the production of sentences presumably requires semantic processing that the production of lists does not.

The sentence > list contrast in comprehension occasionally activates Broca's area (Bedny et al., 2011; Fedorenko et al., 2011; Pallier et al., 2011), but these effects are much less consistent than for the ATL (Rogalsky and Hickok, 2011). This suggests that this activation reflects the contribution of working memory resources or cognitive control mechanisms needed to parse difficult input rather than fundamental syntactic operations (Novick et al., 2005; Rogalsky et al., 2008). Our sentences are short, simple in structure, and guided by a strict template; we believe this minimizes demands on these mechanisms. Therefore we did not expect STRUCTURE to activate Broca's area.



For similar reasons, we also expected that Broca's area would not show a significant effect of PERTURBATION. During switch trials, subjects had an unambiguous selection of the alternative sentence construction, which should minimize selection demands (Miller and Cohen, 2001). With respect to working memory, these resources are taxed in conditions that place heavy demands on maintenance of information or retrieval across intervening material (Baddeley, 1992; Gibson, 2000; Lewis et al., 2006). During switch trials, subjects had to quickly restructure their utterance but did not have to maintain additional material or retrieve information across long distance.

We did not have equally strong predictions for the PERTURBATION contrast in the ATL. Although the existing evidence does not support a role for the ATL in syntax, changes in syntactic structure lead to changes in semantic interpretation (Chomsky, 2014). If the ATL plays a role in combinatory semantics, syntactic restructuring might induce activation in this region for semantic processes. Therefore we expected to potentially see increased activation for PERTURBATION in the ATL.

Significant effects of PERTURBATION outside these areas could reflect syntactic operations in areas not traditionally associated with syntax, such as subcortical areas (see Lieberman, 2001; Ullman, 2004; Boeckx and Benítez-Burraco, 2014 for these non-standard proposals). Activation for this contrast could also reflect non-syntactic mechanisms. These could be linguistic (e.g., reanalysis of thematic role assignment), or non-linguistic (e.g., error detection, attention). In the discussion section we discuss robust activation of the right IFG for PERTURBATION in the context of the literature on action inhibition and the role it may play in syntactic restructuring.

The secondary COMPLEXITY contrast (passive/complex > active/simple) is an extension of previous work that has shown increased activation in Broca's area for passive sentences (e.g., Ye and Zhou, 2009; Mack et al., 2013). The standard interpretation of this finding is that increased syntactic processing resources are used to process passive sentences. However, while historical approaches of generative grammar (Chomsky, 1982, 2002, 2014) posited a syntactic complexity difference between passives and actives (application of a movement operation in passives), modern syntactic theory does not (largely due to the VP-internal subject hypothesis – active sentences also involve movement of the subject, Kitagawa, 1994). Any complexity difference between passives and actives therefore likely lies in non-syntactic factors, such as the mapping of arguments to thematic roles (we thank an anonymous reviewer for pointing this out). The fact that Broca's area *does* show increased activation for passives during comprehension supports a non-syntactic interpretation of the function of this region. Whether this region shows increased activation for passive compared to active sentences during production is an open question and should inform hypotheses of this region's function. We did not have strong predictions for this contrast, but included it because it allowed comparison with previous research in comprehension.

## MATERIALS AND METHODS

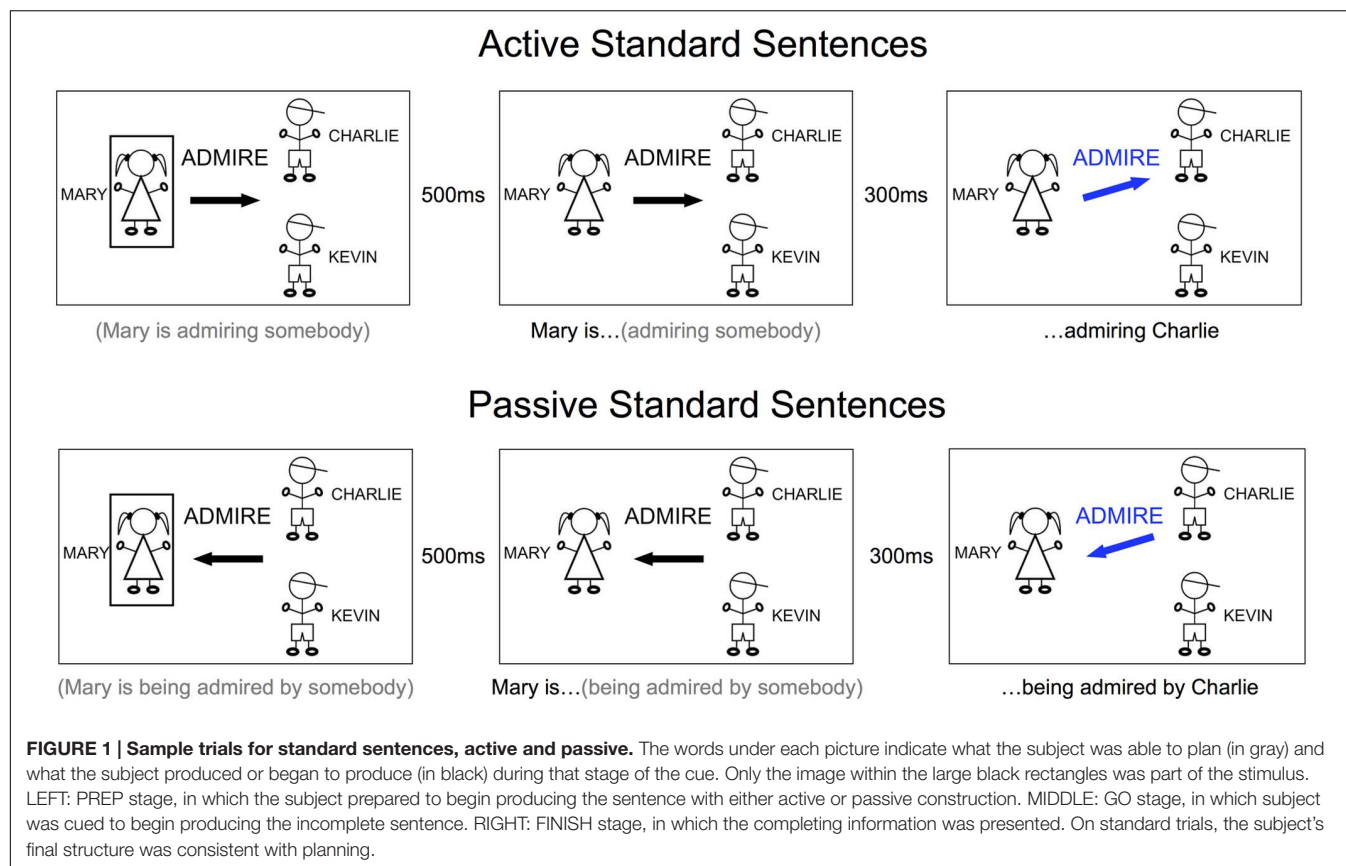
### Subjects

Twenty-one right-handed, native speakers of English (age 19–33, 10 female) volunteered for participation. Subjects had normal or corrected-to-normal vision, no hearing impairment, and reported no history of neurological disorder. Subjects were paid \$10 for participation in a 1-hour behavioral training session. One subject was excluded from the fMRI portion of the experiment due to difficulty with the task during the behavioral session resulting in 20 remaining subjects in the fMRI experiment. Subjects were paid \$30 an hour for participation in the fMRI session. Consent was acquired from each subject before participation and all procedures were approved by the Institutional Review Board of UC Irvine.

### Stimuli

The stimulus for each trial consisted of a cue that progressed through three stages: PREP, GO, and FINISH. Every stimulus had the same basic appearance: simple line drawings of the people engaged in the target sentence, the names of the people in large font next to the drawings, the verb to be used in the sentence in the middle of the screen, and an arrow underneath the verb pointing to the right or the left (**Figure 1**). Identical stimulus presentation was used for both sentences and lists – only the subject's task changed. Twenty different transitive verbs were used. Verb length varied from one to three syllables. Verbs were selected for a mix of articulatory complexity. Here is the complete list of verbs: admire, deceive, examine, follow, frighten, greet, harass, help, hug, kick, kiss, pinch, poke, protect, punch, push, rob, scare, tease, tickle.

Verbs were randomly distributed throughout the experimental runs. Four people were used with these names: Mary, Susan, Charlie, and Kevin. The first person was always a different gender than the final person, and people were randomly distributed in different positions throughout the experimental runs. Three people appeared on each cue: one person on the left (START) and two on the right (END). The END people were displayed vertically, one above the other. During the first stage, the PREP stage, a rectangular box surrounded the START person. The arrow, in black color, pointed from the START person horizontally toward the middle of the END people, and not directly toward either of them. During the second stage, the GO stage, the rectangular box disappeared, serving as a “go” signal for the subject to begin articulating. During the third stage, the FINISH stage, the arrow turned blue and tilted up or down to point to the target END individual for that trial. This design forced the subject to begin articulating without knowing which person to end the sentence with and to use the information provided at the FINISH stage to complete the sentence with the correct person. The PREP stage lasted for 500 ms. We chose this time to give subjects enough time to process the information and plan their utterances. The GO stage began immediately after the 500 ms, and subjects began articulating in synchrony with the disappearance of the box. The interval between the GO stage and the FINISH stage was 300 ms, and



the FINISH stage remained on the screen for 1000 ms, followed by fixation until the next trial. During the behavioral training session, the subject would initiate the next trial whenever ready. During the fMRI session, the inter-trial-interval was fixed at 4200 ms, for a total trial duration/inter-trial interval of 6 s.

## Task

The task was production of either sentences or lists and to restructure appropriately to the switch cue. This resulted in a  $2 \times 2$  design: **STRUCTURE** (sentence, list) and **PERTURBATION** (standard, switch). In the sentence condition, subjects produced sentences with either active or passive construction using the template detailed below. These two constructions comprised an additional sub-factor within the sentence condition, **COMPLEXITY** (active/simple, passive/complex). Active sentences were cued with an arrow pointing away from the first person, and passive sentences were cued with an arrow pointing toward the first person (**Figure 1**). Active sentences were produced with this template: (person 1) *is (verb)ing* (person 2). e.g., *Mary is following Charlie*. Passive sentences were produced with this template: (person 1) *is being (verb)ed by* (person 2). e.g., *Mary is being followed by Charlie*. We instructed subjects to use the progressive aspect on every trial and not to deviate from the template. In the list condition, subjects produced a list of words based on the information from the cue. Subjects ignored the identity of the particular verb on the cue and did not use

it in their lists. When the arrow pointed to the right (as in active sentences), subjects produced a list with this template: (person 1) “word **right** arrow” (person 2), e.g., “Mary word right arrow Charlie.” When the arrow pointed to the left (as in passive sentences), subjects produced a list with this template: (person 1) “word **left** arrow” (person 2). e.g., “Mary word left arrow Charlie.” We chose the word *word* to approximately control for the duration of planning and articulation that would take place for the word *is* in the sentence condition. This timing was relevant to when the subjects were cued to restructure during their utterance (we discuss this in more detail below). Subjects made their utterances at a natural speaking rate.

Subjects did not know how to complete the sentence/list at the beginning of each trial. The FINISH stage indicated which person (top or bottom) would be the second person in the sentence. Subjects were instructed to begin their utterances at the GO stage and use the information on the FINISH stage to determine which name to produce. We set the ISI between the GO stage and the FINISH stage to be 300 ms to allow subjects enough time while speaking naturally to update their utterance without making mistakes on the switch trials. As an example, if the target sentence were “Mary is following Charlie,” at the GO stage subjects started speaking “Mary is following . . .,” then 300 ms later at the FINISH stage they updated their plan to include “Charlie” and finished. Similarly for the list condition, if the target list were “Mary word left arrow Charlie,” they would start speaking “Mary word left

arrow...” at the GO stage, and update at the FINISH stage to include “Charlie.”

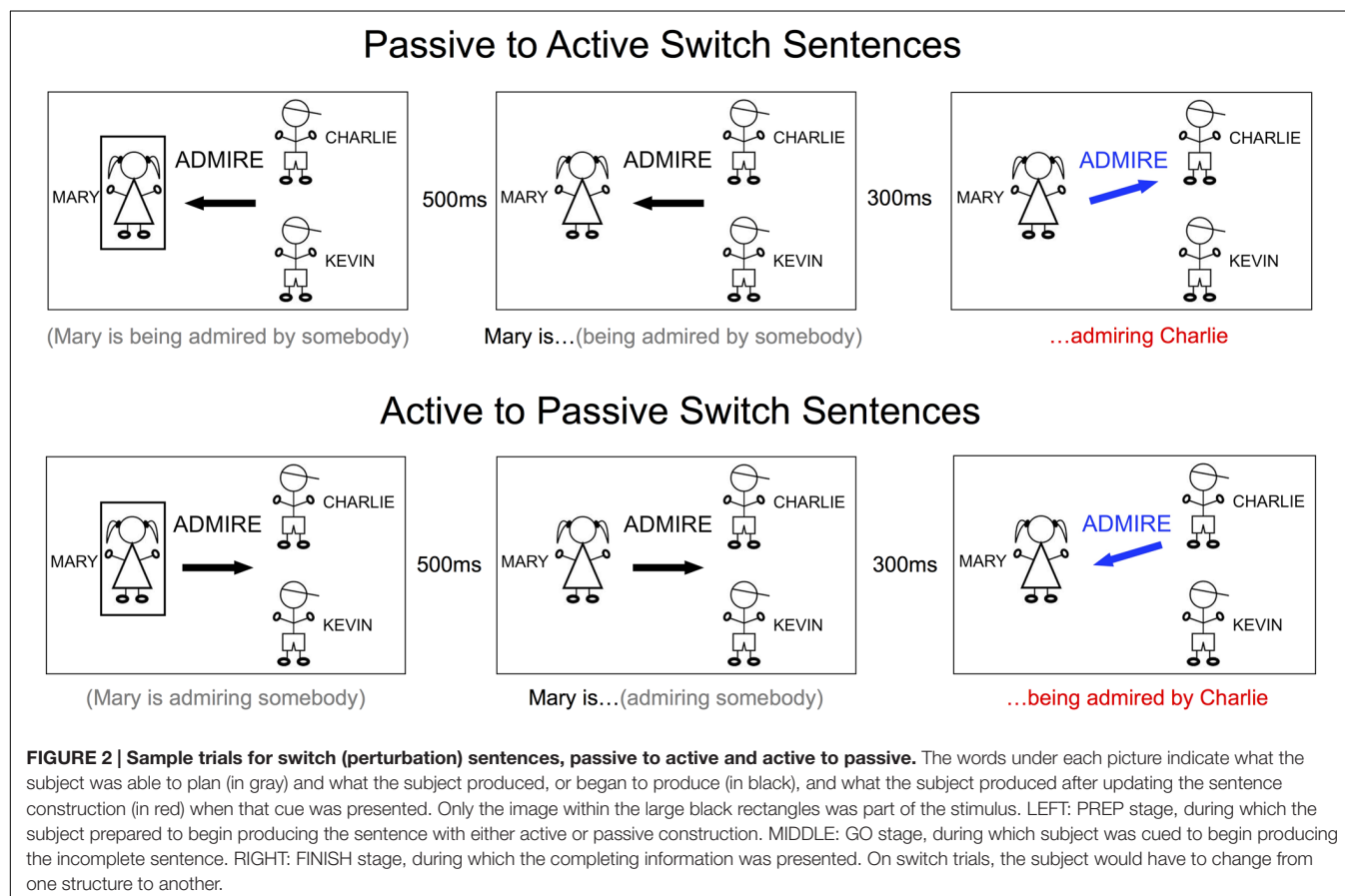
Standard trials occurred as described above; switch trials involved not only updating person 2, but also switching the orientation of the arrow mid-production (**Figure 2**). On sentence switch trials, subjects switched their target sentence from active to passive or vice versa, e.g., *Mary is following (person 2)* → *Mary is being followed by Charlie*. During list switch trials, subjects needed to switch whether they said *right arrow* or *left arrow*, e.g., *Mary word left arrow (person 2)* → *Mary word right arrow Charlie*. Standard and switch trials were presented at a 4/1 ratio and in random order within each run, such that subjects could not predict what the next trial would be. We used this ratio because this approximate ratio was used in previous studies of target perturbation and fMRI (3/1 ratio used by Tourville et al., 2008), and a smaller ratio of standard to switch trials might have resulted in anticipation of switch trials. We did not want subjects to use a strategy of not committing to a syntactic plan on every trial in order to avoid errors.

The sentence and list conditions were presented in separate runs to avoid confusion and task-switching effects. To balance the spatial orientation of the cues, we counterbalanced across sides by presenting subjects with cues that flowed from left to right (depicted in **Figures 1** and **2**) and cues that flowed from right to left (active sentences correspondingly began with a left arrow instead of a right arrow). Subjects always received two runs from

either the sentence or list condition in a row, one each of left and right cue orientation (order counterbalanced across subjects), and we collapsed all analyses across the two orientations.

## Behavioral Training Session

Before running the experiment in the fMRI scanner, we familiarized subjects on the task in a behavioral training session. We wanted subjects to be well prepared for the task in the fMRI scanner to limit variance in performance as well as minimize effects of exposure. In the training session we explained the task to the subjects, including a demonstration by the experimenter on several trials. Then, subjects were asked to perform the task themselves. In the first several trials, the experimenter remained in the testing room to give feedback and instruction. When the subject grasped the task, the experimenter left the room and the subject proceeded self-paced. Subjects performed both tasks with both orientations for a total of four experimental runs, consisting of 50 trials apiece, for a total of 100 trials in the sentence condition and 100 trials in the list condition. The subjects' utterances were recorded and their performance was analyzed. A subject's response was considered an error if they produced the incorrect sentence construction (e.g., active instead of passive), produced the word *right* instead of the word *left* (or vice versa), or if they made a speech error during the trial (e.g., produced the wrong speech sound, extensive delays, etc.). Substituting the names of people (e.g., *Mary* instead of *Susan*) or substituting one verb for



another (e.g., *push* instead of *punch*) were not counted as errors, unless the subject also made an additional error as described above. We were only able to collect and analyze behavioral data from 14 out of 20 subjects due to equipment issues. To assess the effect of perturbation on behavior, we averaged across constructions in the sentence conditions and direction in the list conditions. We then performed a  $2 \times 2$  ANOVA (STRUCTURE  $\times$  PERTURBATION). Subjects underwent the fMRI portion of the experiment after completing the behavioral session, either the same day or on a subsequent day, within a week after the behavioral session.

## fMRI Experiment

Before scanning, subjects were briefly re-familiarized with the task by performing a few trials in each condition outside the scanner. Subjects were instructed to produce their utterances out loud in the scanner, but quietly and with minimal articulation. Subjects received 12 total experimental runs during the experiment (six sentence, six list, counterbalanced by orientation). During the experiment, a fixation cross was displayed on a screen in-between trials. Stimuli were delivered with Matlab software (Mathworks, Inc, USA) utilizing Psychtoolbox (Brainard, 1997; Kleiner et al., 2007). Subjects were given ear covers and foam earplugs to attenuate scanner noise. Each run contained 40 standard trials and 10 switch trials in random order with no explicit rest trials. Presentation order of sentence and list runs was counterbalanced along with cue orientation across subjects. Active/passive constructions and left/right arrow lists were presented at equal frequency. The high-resolution anatomical image was collected following the experimental runs. The scanning session lasted about 1 h and 15 min in total.

## fMRI Data Collection and Analysis

MR images were obtained in a Philips Achieva 3T (Philips Medical Systems, Andover, MA, USA) fitted with an eight-channel RF receiver head coil at the high field scanning facility at UC Irvine. We first collected a total of 1896 T2\*-weighted EPI volumes over 12 runs using Fast Echo EPI in ascending order ( $TR = 2$  s,  $TE = 25$  ms, flip angle =  $90^\circ$ , in-plane resolution =  $1.95$  mm  $\times$   $1.95$  mm, slice thickness =  $3$  mm with  $0.5$  mm gap). The first four volumes of each run were collected before stimulus presentation and discarded to control for T1 saturation effects. The high-resolution T1-weighted anatomical image was acquired in the axial plane ( $TR = 8$  ms,  $TE = 3.7$  ms, flip angle =  $8^\circ$ , size =  $1$  mm isotropic).

Slice-timing correction, motion correction, and spatial smoothing were performed using AFNI software (<http://afni.nimh.nih.gov/afni>). Motion correction was achieved by using a 6-parameter rigid-body transformation, with each functional volume in a run first aligned to a single volume in that run. Functional volumes were aligned to the anatomical image, and subsequently aligned to Talairach space (Talairach and Tournoux, 1988). Functional images were resampled to  $2.5$  mm isotropic voxels and spatially smoothed using a Gaussian kernel of  $6$  mm FWHM. Finally, functional images were rescaled to reflect percent signal change from the mean signal during each run.

First-level analyses were performed on each individual subject's data using AFNI's 3dDeconvolve function. The regression analysis was performed to find parameter estimates that best explained variability in the data. Each predictor variable representing the time course of activity associated with the task was entered into a deconvolution analysis that estimated parameters best representing the timecourse of the hemodynamic response function in percent signal change values. Timecourse estimates were modeled beginning with the onset of the PREP stage, i.e., when the subject began planning the sentence. The following eight regressors of interest were used in the experimental analysis: sentence active, sentence passive, list left, list right, sentence switch: active to passive, sentence switch: passive to active, list switch: left to right, and list switch right to left. The six motion parameters were included as regressors of no interest. Second-level group analyses were then performed. The values from the experimental contrasts from each subject and condition were entered into a mixed-effects analysis with subjects as random variables using AFNI's 3dMEMA function. We tested the following contrasts: sentence vs. list (STRUCTURE), active vs. passive (COMPLEXITY), and switch vs. standard (PERTURBATION). Because we were particularly interested in switch effects for the sentence condition, we examined the effects of PERTURBATION for sentences and lists separately in addition to the interaction of STRUCTURE and PERTURBATION. We corrected for multiple comparisons though Monte Carlo simulation using AFNI's 3dClustSim function to hold the family-wise error (FWE) rate to less than  $0.05$ . We estimated smoothness in the data from the residual error time series for each subject's first-level analysis using AFNI's 3dFWHMx function. These estimates were averaged across participants for input to 3dClustSim (simulations were restricted to in-brain voxels). Activations were considered significant with a per-voxel threshold of  $p < 0.001$  (one-tailed) and a cluster size threshold of  $610$  mm<sup>3</sup> (39 voxels).

## ROI Analyses

Given the extensive literature documenting a relationship between Broca's area, the ATL, and sentence processing, we performed ROI analyses on these regions. We extracted percent signal change values within structural ROIs for the left and right ATL, Broca's area, and the right hemisphere homolog of Broca's area, the right inferior frontal gyrus (RIFG) and ran statistical analyses. For Broca's area and the RIFG, we used templates in Talairach space for BA44 and BA45 provided by AFNI based on the cytoarchitectonic probability maps of Amunts et al. (1999). We included every voxel in each map and combined both maps together to form a single mask for Broca's area and a single mask for the RIFG. The relevant functional regions of interest for the ATL do not align well to probability maps based on cytoarchitectonics; we constructed left and right ATL ROIs based on coordinates reported in the neuroimaging literature. We obtained the center of mass coordinates reported by Rogalsky and Hickok (2009) for the sentence > list contrast in the left and right ATL, and created spheres with radius  $10$  mm around the coordinates. We averaged across all voxels within each ROI and analyzed the average percent signal change values across

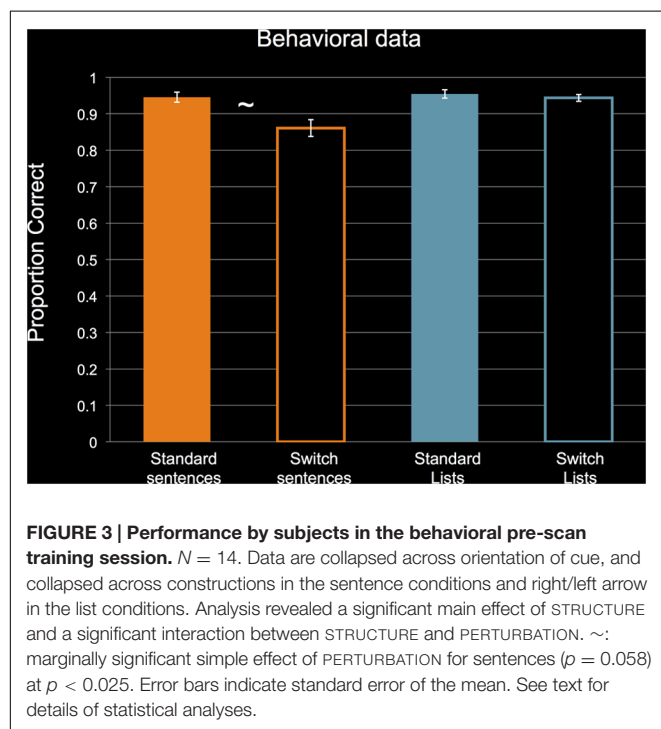


the entire estimated timecourse. We first analyzed the effect of COMPLEXITY (passive > active) within each ROI with paired *t*-tests. We then collapsed our analyses across constructions in the sentence conditions and direction in the list conditions, resulting in  $2 \times 2$  ANOVAs for each ROI (STRUCTURE  $\times$  PERTURBATION).

## RESULTS

### Behavioral Performance

To reiterate, we only collected behavioral data during the behavioral training session before the fMRI session. **Figure 3** shows the behavioral performance of the 14 subjects for whom we collected data. For non-switch standard trials, subjects performed near ceiling for the sentence and list conditions. The clear outlier is the sentence switch condition. Even though subjects' performance dropped during switch sentence trials, their performance was still above 80%, indicating that they could successfully perform the task. A  $2 \times 2$  ANOVA revealed a significant main effect of STRUCTURE,  $F(1,13) = 5.282$ ,  $p = 0.039$ ,  $\eta^2 = 0.289$ , no significant main effect of PERTURBATION,  $F(1,13) = 3.232$ ,  $p = 0.095$ ,  $\eta^2 = 0.199$ , and a significant interaction,  $F(1,13) = 5.353$ ,  $p = 0.038$ ,  $\eta^2 = 0.292$ . Follow-up two-tailed *t*-tests ( $\alpha = 0.025$ ) revealed a marginally significant effect of PERTURBATION for sentences,  $t(1,13) = 2.077$ ,  $p = 0.058$ , Cohen's  $d = 0.555$ , and no effect of PERTURBATION for lists,  $t(1,13) = 0.668$ ,  $p = 0.516$ , Cohen's  $d = 0.169$ . These results confirm that performance was only impaired during the sentence switch condition.



### Whole-Brain fMRI Analyses

The whole-brain contrasts of STRUCTURE and COMPLEXITY did not reveal activation in the ATL or Broca's area. The effect of STRUCTURE (sentences > lists) revealed increased activation for sentences in left visual cortex, right precentral gyrus, right postcentral gyrus, and bilateral middle frontal gyrus (**Figure 4**). The effect of COMPLEXITY (passive > active sentences) revealed one cluster in the left postcentral gyrus (**Figure 4**). See **Table 1** for Talairach coordinates for each significant cluster of activation for these contrasts.

The effect of PERTURBATION in the sentence condition (sentence switch > sentence control) revealed increased activation during the switch condition in a network including areas typically found for experiments of response selection/inhibition as in the Go/No-Go task (Simmonds et al., 2008; Swann et al., 2009). The GO/No-Go task requires subjects to inhibit a planned motor response when a "stop" signal appears, as well in areas found for perturbation in low-level motor control (Diedrichsen et al., 2005; Tourville et al., 2008; **Figure 5**). Particularly strong activation was observed in the right IFG and anterior insula that has been shown to be involved in "stopping" or the cancelation of a planned response (Aron et al., 2003, 2014). Activations for this contrast also included the supplementary motor area (SMA), pre-SMA, basal ganglia (right caudate nucleus), left inferior parietal cortex, right STS, and right IFG/MFG (**Figure 5**). The effect of PERTURBATION in the list condition (switch lists > standard lists) revealed one cluster in the left cerebellum (**Figure 5**, bottom). See **Table 2** for Talairach coordinates for each significant cluster of activation for these contrasts.

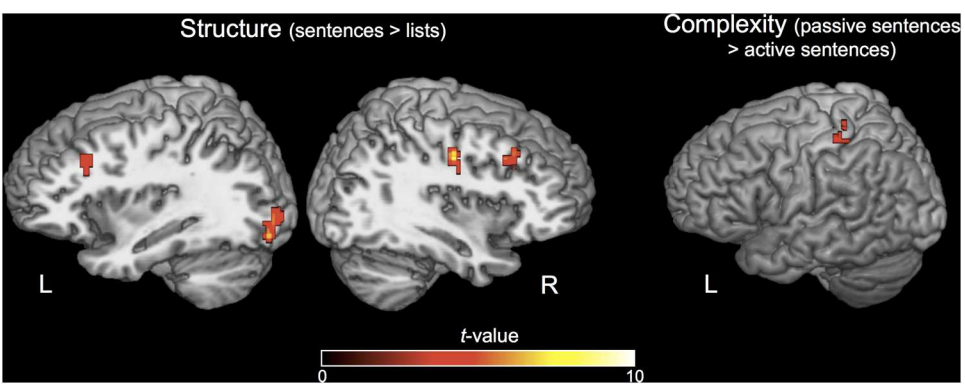
The interaction contrast of PERTURBATION with STRUCTURE did not reveal any significant clusters when cluster-corrected for multiple comparisons, suggesting that there was a similar switch effect across the sentence and list conditions in the brain, although the separate contrasts for these conditions activated different sets of areas.

### ROI Analyses

For the ROI analyses, based on our expectations from the literature, we separately examined the effect of COMPLEXITY (passive > active sentences) using a one-way *t*-test and performed a  $2 \times 2$  ANOVA of PERTURBATION (switch vs. control) and STRUCTURE (sentences vs. lists).

There was no effect of COMPLEXITY (passive > active) for any of the ROIs (all reported tests are one-tailed *t*-tests). Broca's area:  $t(1,19) = -0.059$ ,  $p = 0.477$ ; RIFG:  $t(1,19) = 0.069$ ,  $p = 0.473$ ; left ATL:  $t(1,19) = 1.746$ ,  $p = 0.952$ ; right ATL:  $t(1,19) = 0.799$ ,  $p = 0.783$ . The high *t*-value of the left ATL indicates that there was a possibility of higher activation for *active* – less complex – sentences.

In Broca's area, there was no significant effect of STRUCTURE,  $F(1,19) = 1.443$ ,  $p = 0.244$ , or PERTURBATION,  $F(1,19) = 0.714$ ,  $p = 0.408$ , and no significant interaction,  $F(1,19) = 0.164$ ,  $p = 0.408$ . In the RIFG, there was no significant effect of STRUCTURE,  $F(1,19) = 0.005$ ,  $p = 0.946$ , a significant effect of PERTURBATION,  $F(1,19) = 13.541$ ,  $p = 0.002$ , and no significant



**FIGURE 4 | Activations for the STRUCTURE and COMPLEXITY contrasts.** *N* = 20. Activations are displayed on a Talairach template brain and cluster-corrected for multiple comparisons at FWE *p* < 0.05, individual voxel threshold *p* < 0.001 (one-tailed), cluster size threshold 610 mm<sup>3</sup>.

**TABLE 1 | Effects of STRUCTURE and COMPLEXITY.**

Region	Hemisphere	x	y	z	Cluster size (mm <sup>3</sup> )
<b>Effect of structure</b>					
Middle frontal gyrus	Right	30	19	37	891
Inferior occipital gyrus	Left	−29	−86	−5	844
Middle frontal gyrus	Left	−27	24	33	813
Precentral gyrus	Right	38	−13	37	750
<b>Effect of complexity</b>					
Postcentral gyrus	Left	−32	−36	51	641

*N* = 20. FWE cluster-corrected *p* < 0.05; individual voxel threshold *p* < 0.001, cluster size threshold 610 mm<sup>3</sup>. Coordinates reflect the center of mass of each significant cluster. Coordinates are reported in Talairach space.

effect of the interaction,  $F(1,19) = 0.663$ ,  $p = 0.426$ . Activations for each of these conditions in Broca’s area and the RIFG are displayed in **Figure 6**. Both ATL regions showed a significant main effect of PERTURBATION (reduced activity for perturbation), no main effect of STRUCTURE, and no interaction. Left ATL: STRUCTURE,  $F(1,19) = 0.597$ ,  $p = 0.449$ ; PERTURBATION,  $F(1,19) = 6.963$ ,  $p = 0.016$ ; interaction,  $F(1,19) = 2.820$ ,  $p = 0.110$ . Right ATL: STRUCTURE,  $F(1,19) = 0.123$ ,  $p = 0.729$ ; PERTURBATION,  $F(1,19) = 13.161$ ,  $p = 0.002$ ; interaction,  $F(1,19) = 0.396$ ,  $p = 0.537$ . Activation for each of these conditions in left and right ATL ROIs are displayed in **Figure 7**. While the test of the interaction between STRUCTURE and PERTURBATION in the left ATL was not significant, it should be noted that this effect trended toward significance.

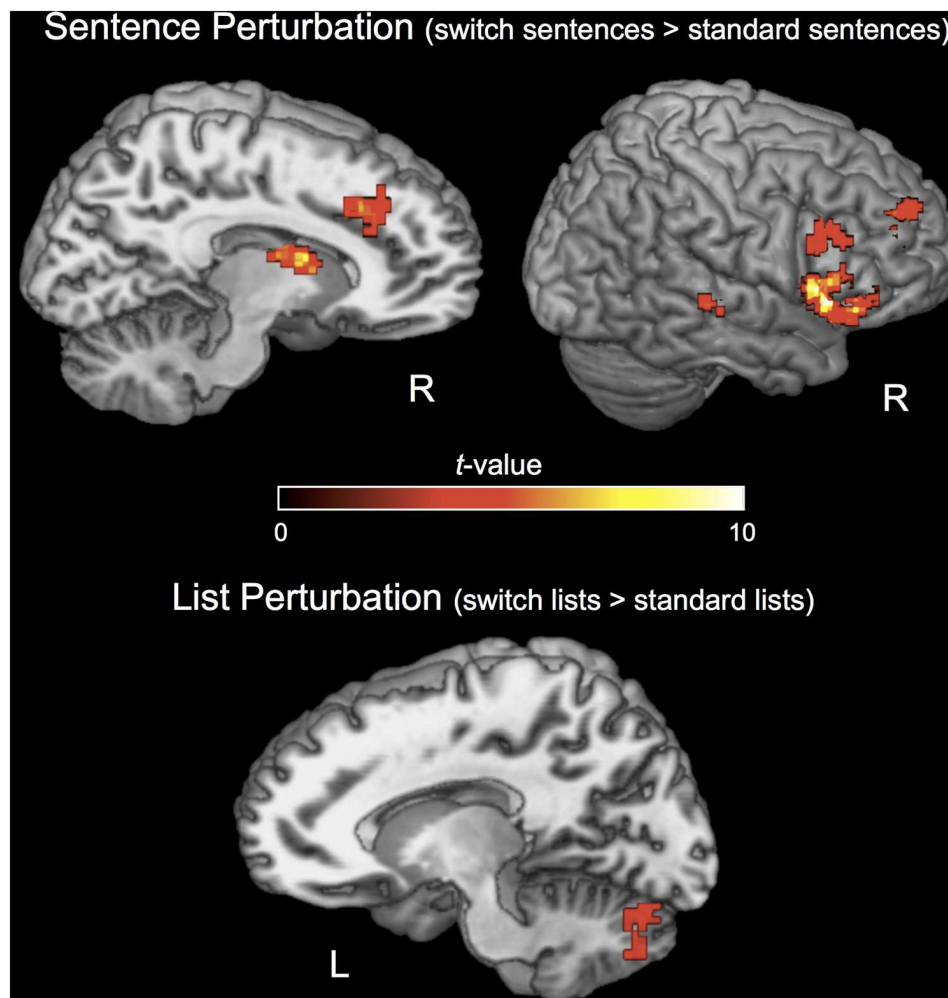
DISCUSSION

We performed a novel investigation in the effort to understand the neural bases of syntax: a constrained speech production task, including two different sentence constructions (active and passive), unstructured lists, and a “syntactic perturbation” paradigm. One goal was to probe the response profile of the traditional candidates for syntactic processing and their right

hemisphere homologs with this novel PERTURBATION paradigm and a contrast of STRUCTURE (sentences > word lists). We also included a secondary contrast of COMPLEXITY (passive > active sentences) to determine if this effect previously found for comprehension in Broca’s area extended to production. Finally, another goal of the experiment was exploratory – to determine whether networks outside of the traditional candidate regions for syntax would activate to syntactic perturbation. We will first focus our discussion on the activation profiles of Broca’s area and the ATL. Following this we discuss the effects we obtained for the whole-brain contrasts, particularly the activation we obtained for PERTURBATION in the right IFG and the potential role this region plays in sentence processing.

The Activation Profile of Broca’s Area

The domain-general hypotheses of Broca’s area suggest that this region underlies a non-syntactic mechanism during sentence processing, either resolving representational conflict through cognitive control (Novick et al., 2005) or providing working memory resources (Rogalsky and Hickok, 2011). The lack of effects for our PERTURBATION contrast in this region is consistent with these accounts, and contrary to the expectations of a region involved in syntax, as perturbation was expected to tax syntactic processing. We also did not observe a significant effect of STRUCTURE in Broca’s area. Previous work has shown that this contrast is observable in small subregions directly adjacent to regions that do not show this contrast (Hickok and Rogalsky, 2011; Fedorenko et al., 2012a). Our structural ROIs may have contained sentence-selective and non-selective subregions, thus weakening our power to detect effects of STRUCTURE. Regardless, the contrast was clearly not robust, and combined with the fact that the PERTURBATION contrast did not approach significance in this region speaks against a syntactic function. We did not replicate previous findings for passive > active sentences in Broca’s area in comprehension (Ye and Zhou, 2009; Mack et al., 2013). Hypotheses of Broca’s area function in sentence processing should take this disparity into account, while noting that the task constraints of our study may have substantially



**FIGURE 5 | Activations for the PERTURBATION contrasts.**  $N = 20$ . Activations are displayed on a Talairach template brain and cluster-corrected for multiple comparisons at FWE  $p < 0.05$ , individual voxel threshold  $p < 0.001$  (one-tailed), cluster size threshold  $610 \text{ mm}^3$ .

reduced our ability to detect activation differences between these constructions.

### The Activation Profile of the Anterior Temporal Lobe (ATL)

Our whole-brain analysis did not reveal any effects of COMPLEXITY and STRUCTURE in the left or right ATL. However, the ROI analysis did reveal a PERTURBATION effect for the ATL bilaterally – *decreased* activity for perturbation. We attribute the null effect of STRUCTURE and the decreased activity for PERTURBATION to a semantic rather than syntactic function of the ATL and decreased attention to semantic content in our study.

Our ROI plots showed less activity for switch sentences than for natural sentences, which reduced the sensitivity of our analyses to detect a main effect of STRUCTURE. This reduction can be explained by decreased attention to the semantic content of the stimulus for switch trials. Rogalsky and Hickok (2009)

showed that attention substantially affects activation to semantic content in the ATL. The demanding nature of our task may have distracted subjects away from the semantic content of the sentences, reducing the difference in semantic processing between lists and sentences. Our stimuli also had limited semantic content generally. We used proper nouns instead of common nouns (e.g., *Mary* instead of *the acrobat*), and simplistic line drawings devoid of detail rather than pictures of actual people engaging in action as used in other studies (e.g., Menenti et al., 2011; Segaert et al., 2012). Future studies seeking to obtain effects of structure in the ATL during speech production should enrich the semantic nature of the materials and choose a task that does not require heavy attentional demands.

The decreased activity for PERTURBATION is contrary to the expectations of a region involved in syntax, but compatible with a role for semantics. Any effect of PERTURBATION would presumably increase demands on syntactic structure building, rather than decrease them. The increased attentional demands of switching syntactic structures, however, likely reduced attention

**TABLE 2 | Effects of PERTURBATION.**

Region	Hemisphere	x	y	z	Cluster size (mm <sup>3</sup> )
<b>Effect of perturbation in sentences</b>					
Inferior frontal gyrus/anterior insula	Right	41	18	2	7,297
Superior frontal gyrus	Right	24	45	25	1,969
Medial frontal gyrus	Right	7	27	33	1,531
Superior temporal sulcus	Right	52	−28	1	844
Caudate nucleus	Right	11	1	14	813
Middle frontal gyrus	Right	52	17	28	766
<b>Effect of perturbation in lists</b>					
Cerebellum	Left	−8	−74	−26	1,078

*N* = 20. FWE cluster-corrected  $p < 0.05$ ; individual voxel threshold  $p < 0.001$ , cluster size threshold 610 mm<sup>3</sup>. Coordinates reflect the center of mass of each significant cluster. Coordinates are reported in Talairach space.

to the semantic content of the sentences, accounting for a reduction of activity in the ATL as discussed above.

The major piece of data in support of a basic syntactic function of the ATL is the observation that the structural effect in the ATL can be found for sentences with the content words replaced by non-words, retaining the structural “feel” but with greatly impoverished semantic content (i.e., jabberwocky stimuli; Mazoyer et al., 1993; Humphries et al., 2006; Rogalsky et al., 2011). However, this effect is much less robust than for full sentences, with some studies failing to observe it at all (Pallier et al., 2011; Fedorenko et al., 2012c). Future research could determine the source of these discrepancies, including testing the notion that there may be a functional-anatomical subdivision within the ATL between syntactic and semantic processing (Rogalsky and Hickok, 2009).

## Whole-Brain Contrasts of Complexity and Structure

We first discuss the whole brain contrasts of COMPLEXITY and STRUCTURE. The whole-brain contrast of COMPLEXITY revealed one significant cluster in the left post-central gyrus. Since passive sentences are longer than active sentences, requiring additional articulation, this cluster likely reflects the increased motor speech output and corresponding somatosensory input rather than any core linguistic function. The whole-brain contrast of STRUCTURE only revealed activity in visual cortex and bilateral superior frontal areas. These regions have been previously associated with visual attention (Kastner and Ungerleider, 2000; Corbetta and Shulman, 2002). This suggests that demands on visual attention were stronger during the sentence condition than during the list condition, which is supported by the behavioral data.

The lack of additional effects in language-related regions for these contrasts deserves explanation. We have already discussed Broca’s area and the ATL; other language-related areas that are

typically activated by this contrast include the left posterior temporal lobe and the angular gyrus (Bedny et al., 2011; Pallier et al., 2011; Fedorenko et al., 2012c). The difference between our results and previous studies cannot be attributed solely to the differences between production and comprehension; several production studies have revealed effects in these areas (Menenti et al., 2011, 2012; Segaert et al., 2012, 2013). As discussed in the introduction, the structurally simple and short sentences that we used minimized demands on working memory and cognitive control, and our stimuli did not encourage rich semantic processing. It may be the case that effects in these language-related regions are due to these processes. Previous research points to a role for the posterior temporal lobe in working memory and cognitive control (Hickok et al., 2003; Glaser et al., 2013) and the angular gyrus in semantic processing (Binder et al., 2009; Price et al., 2015), consistent with this speculation.

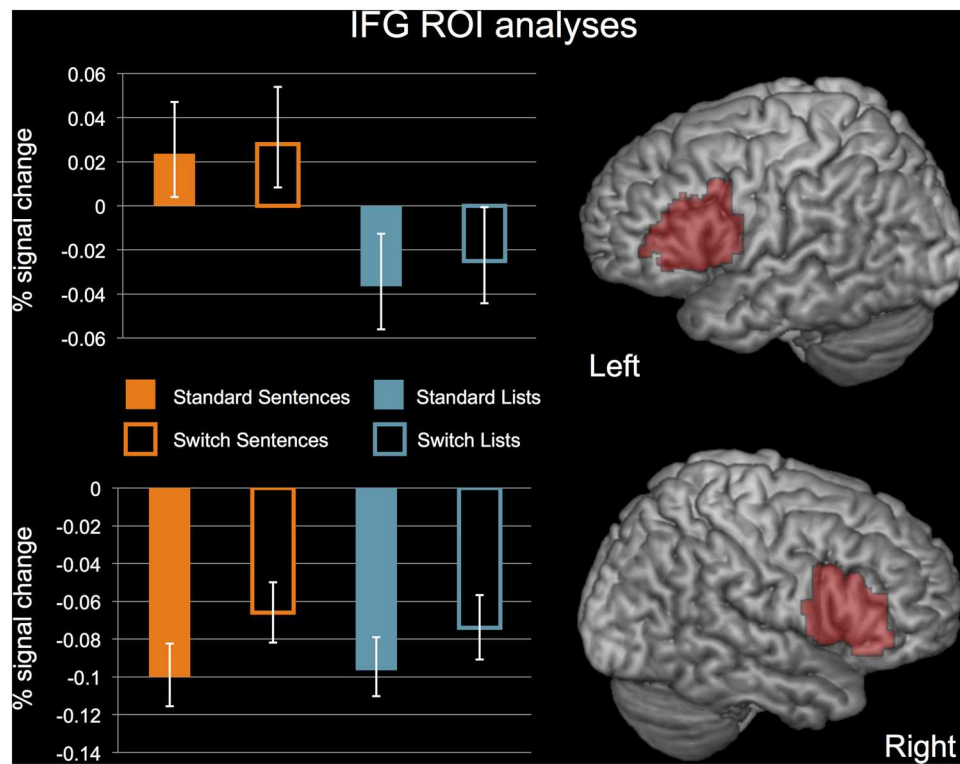
## Syntactic Perturbation Reveals a Network for Response Selection, Action Inhibition, and Motor Control

While syntactic PERTURBATION did not activate traditional language areas of the left hemisphere, it did activate other brain regions, including medial frontal areas (SMA, pre-SMA), the right caudate nucleus, the right posterior STS, the right IFG, and the right anterior insula. These are regions that have been reported in studies of perturbation and motor control in other domains (Diedrichsen et al., 2005; Suminski et al., 2007; Tourville et al., 2008) and studies of response selection/action inhibition implementing go/no-go designs (Simmonds et al., 2008). The list PERTURBATION contrast activated only the cerebellum. This disparity of results between the sentence and list conditions must be treated carefully, as the interaction contrast did not reveal a significant statistical interaction between STRUCTURE and PERTURBATION in any regions. This suggests that there were similar activation patterns for both conditions, but that the effect was somewhat stronger in the sentence condition.

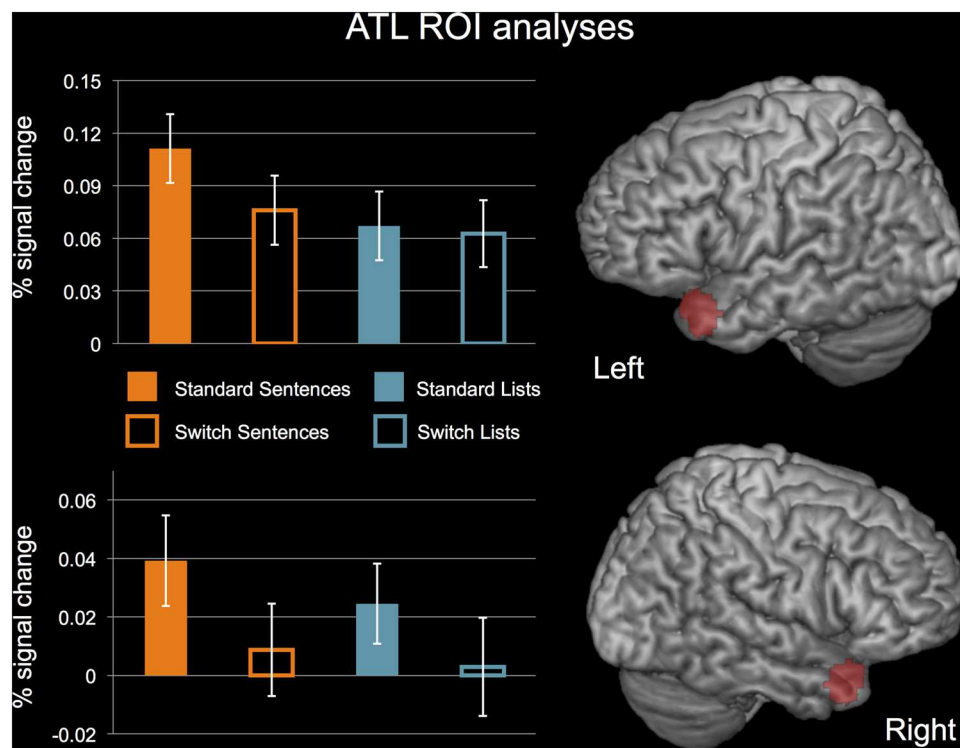
The activation of the right caudate nucleus is consistent with the suggestion that the basal ganglia are involved in syntactic operations (Lieberman, 2001; Ullman, 2004). However, we do not believe that this activation in our study reflects syntax. This is because the right basal ganglia are part of a larger network that is strongly implicated in stopping, discussed below.

While RIFG activation is sometimes reported for syntactic manipulations (Embick et al., 2000; Friederici et al., 2000; Meyer et al., 2000; Fiebach et al., 2005; Tyler et al., 2010), it is not common for experiments of basic sentence processing, and the aphasia literature does not support a strong association between deficits in sentence processing and the RIFG (Damasio, 1992; but see Caplan et al., 1996). The effect of PERTURBATION in this region therefore likely reflects non-syntactic mechanisms. The operative mechanism may be action inhibition, or “stopping,” which has been attributed specifically to the RIFG in conjunction with the other areas activated by the PERTURBATION contrast (Aron et al., 2003, 2014). Under this hypothesis, the RIFG operates as a “brake.” We can apply this braking hypothesis to the





**FIGURE 6 | ROI analyses for Broca's area and the RIFG.**  $N = 20$ . Error bars indicate standard error of the mean. See text for details of statistical analyses.



**FIGURE 7 | ROI analyses for the left and right ATL.**  $N = 20$ . Error bars indicate standard error of the mean. See text for details of statistical analyses.

current study through reverse inference. After subjects planned to produce a sentence with a given sentence construction, on switch trials they utilized the brake to inhibit this plan. When subjects planned to produce a list of words, they also relied on the brake, but less so.

Our study provides insight into a surprisingly large amount of previous studies of syntax and sentence comprehension that report activation of the RIFG. Such studies can be divided into two groups: studies of complex/non-canonical sentence constructions and garden-path sentences (Meyer et al., 2000; Fiebach et al., 2005; Grewe et al., 2006; Bornkessel-Schlesewsky et al., 2012; Chan et al., 2012), and studies involving syntactic violations (Embick et al., 2000; Moro et al., 2001; Ben-Shachar et al., 2003; Friederici et al., 2003, 2006; Bahlmann et al., 2008). The fact that our task explicitly involved stopping suggests that this mechanism may account for RIFG activations in these previous studies. When subjects process a sentence with non-canonical sentence structure or syntactic violations, they must revise their initial parse to arrive at the correct interpretation. This revision may rely on an inhibition function to quickly reject the current parse in favor of a new one. Supporting this hypothesis, Caplan et al. (1996) found that patients with right hemisphere lesions had significantly worse sentence comprehension than control subjects, particularly for complex sentence constructions (although these effects were not as strong as in patients with left hemisphere lesions). Future research could further investigate the hypothesis of a “braking” function during sentence comprehension.

## CONCLUSION

The present study sought to implement a novel paradigm in the study of syntax and the brain: a constrained sentence production

task with a perturbation paradigm applied to syntactic structure. While our activations point to a possibility of a stopping mechanism in the RIFG that facilitates structural revision, it is difficult to make any firm conclusions based on this study alone. The lack of effects for syntactic PERTURBATION and STRUCTURE in Broca’s area suggest that this region performs a non-syntactic function during sentence processing. This supports the previous body of evidence against a role for syntax in Broca’s area (Rogalsky and Hickok, 2011). Finally, we did not extend previous effects of sentences > word lists in the ATL to production, although the lack of an effect may have been due to reduced activity in this region during perturbation. This is consistent with a role for the ATL in combinatorial semantics.

## AUTHOR CONTRIBUTIONS

WM and GH conceptualized and designed the experiment. WM created the stimuli and collected and analyzed the data. WM and GH wrote and revised the manuscript.

## FUNDING

This research was supported by NIH Grant # DC03681 awarded to GH.

## ACKNOWLEDGMENTS

We would like to thank Jonathan Venezia for useful suggestions, Maral Aghvinian for assistance with data analysis, and two anonymous reviewers for their comments.

## REFERENCES

- Amunts, K., Schleicher, A., Bürgel, U., Mohlberg, H., Uylings, H., and Zilles, K. (1999). Broca’s region revisited: cytoarchitecture and intersubject variability. *J. Comp. Neurol.* 412, 319–341. doi: 10.1002/(SICI)1096-9861(19990920)412:2<319::AID-CNE10>3.0.CO;2-7
- Aron, A. R., Fletcher, P. C., Bullmore, T., Sahakian, B. J., and Robbins, T. W. (2003). Stop-signal inhibition disrupted by damage to right inferior frontal gyrus in humans. *Nat. Neurosci.* 6, 115–116. doi: 10.1038/nn1003
- Aron, A. R., Robbins, T. W., and Poldrack, R. A. (2014). Inhibition and the right inferior frontal cortex: one decade on. *Trends Cogn. Sci. (Regul. Ed.)* 18, 177–185. doi: 10.1016/j.tics.2013.12.003
- Baddeley, A. (1992). Working memory. *Science* 255, 556–559. doi: 10.1126/science.1736359
- Bahlmann, J., Schubotz, R. I., and Friederici, A. D. (2008). Hierarchical artificial grammar processing engages Broca’s area. *Neuroimage* 42, 525–534. doi: 10.1016/j.neuroimage.2008.04.249
- Bedny, M., Pascual-Leone, A., Dodell-Feder, D., Fedorenko, E., and Saxe, R. (2011). Language processing in the occipital cortex of congenitally blind adults. *Proc. Natl. Acad. Sci. U.S.A.* 108, 4429–4434. doi: 10.1073/pnas.1014818108
- Bemis, D. K., and Pykkänen, L. (2011). Simple composition: a magnetoencephalography investigation into the comprehension of minimal linguistic phrases. *J. Neurosci.* 31, 2801–2814. doi: 10.1523/JNEUROSCI.5003-10.2011
- Ben-Shachar, M., Hendler, T., Kahn, I., Ben-Bashat, D., and Grodzinsky, Y. (2003). The neural reality of syntactic transformations evidence from functional magnetic resonance imaging. *Psychol. Sci.* 14, 433–440. doi: 10.1111/1467-9280.01459
- Binder, J. R., Desai, R. H., Graves, W. W., and Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb. Cortex* 19, 2767–2796. doi: 10.1093/cercor/bhp055
- Blank, S. C., Scott, S. K., Murphy, K., Warburton, E., and Wise, R. J. (2002). Speech production: wernicke, Broca and beyond. *Brain* 125, 1829–1838. doi: 10.1093/brain/awf191
- Boeckx, C., and Benítez-Burraco, A. (2014). The shape of the human language-ready brain. *Front. Psychol.* 5:282. doi: 10.3389/fpsyg.2014.00282
- Bornkessel-Schlesewsky, I., Grewe, T., and Schlewsky, M. (2012). Prominence vs. aboutness in sequencing: a functional distinction within the left inferior frontal gyrus. *Brain Lang.* 120, 96–107. doi: 10.1016/j.bandl.2010.06.004
- Bornkessel-Schlesewsky, I., and Schlewsky, M. (2013). Reconciling time, space and function: a new dorsal-ventral stream model of sentence comprehension. *Brain Lang.* 125, 60–76. doi: 10.1016/j.bandl.2013.01.010
- Brainard, D. H. (1997). The psychophysics toolbox. *Spat. Vis.* 10, 433–436. doi: 10.1163/156856897X00357
- Braun, A. R., Guillemin, A., Hosey, L., and Varga, M. (2001). The neural organization of discourse. *Brain* 124, 2028–2044. doi: 10.1093/brain/124.10.2028
- Brennan, J., Nir, Y., Hasson, U., Malach, R., Heeger, D. J., and Pykkänen, L. (2012). Syntactic structure building in the anterior temporal lobe during natural story listening. *Brain Lang.* 120, 163–173. doi: 10.1016/j.bandl.2010.04.002

- Caplan, D., and Hanna, J. E. (1998). Sentence production by aphasic patients in a constrained task. *Brain Lang.* 63, 184–218. doi: 10.1006/brln.1998.1930
- Caplan, D., Hildebrandt, N., and Makris, N. (1996). Location of lesions in stroke patients with deficits in syntactic processing in sentence comprehension. *Brain* 119, 933–950. doi: 10.1093/brain/119.3.933
- Chan, Y. C., Chou, T. L., Chen, H. C., and Liang, K. C. (2012). Segregating the comprehension and elaboration processing of verbal jokes: an fMRI study. *Neuroimage* 61, 899–906. doi: 10.1016/j.neuroimage.2012.03.052
- Chomsky, N. (1982). *Some Concepts and Consequences of the Theory of Government and Binding*, Vol. 6. Cambridge MA: MIT Press.
- Chomsky, N. (1995). *The Minimalist Program*, Vol. 28. Cambridge, MA: MIT Press.
- Chomsky, N. (2002). *Syntactic Structures*. Berlin: Walter de Gruyter.
- Chomsky, N. (2014). *Aspects of the Theory of Syntax (No. 11)*. Cambridge MA: MIT Press.
- Corbetta, M., and Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215. doi: 10.1038/nrn755
- Damasio, A. R. (1992). Aphasia. *New Engl. J. Med.* 326, 531–539. doi: 10.1056/NEJM199202203260806
- Del Prado, P., and Pyllkanen, L. (2014). MEG evidence for conceptual combination but not numeral quantification in the left anterior temporal lobe during language production. *Front. Lang. Sci.* 5:524. doi: 10.3389/fpsyg.2014.00524
- Diedrichsen, J., Hashambhoy, Y., Rane, T., and Shadmehr, R. (2005). Neural correlates of reach errors. *J. Neurosci.* 25, 9919–9931. doi: 10.1523/JNEUROSCI.1874-05.2005
- Elliott, D., Lyons, J., Chua, R., Goodman, D., and Carson, R. G. (1995). The influence of target perturbation on manual aiming asymmetries in right-handers. *Cortex* 31, 685–697. doi: 10.1016/S0010-9452(13)80020-2
- Embick, D., Marantz, A., Miyashita, Y., O'Neil, W., and Sakai, K. L. (2000). A syntactic specialization for Broca's area. *Proc. Natl. Acad. Sci. U.S.A.* 97, 6150–6154. doi: 10.1073/pnas.100098897
- Fedorenko, E., Behr, M. K., and Kanwisher, N. (2011). Functional specificity for high-level linguistic processing in the human brain. *Proc. Natl. Acad. Sci. U.S.A.* 108, 16428–16433. doi: 10.1073/pnas.1112937108
- Fedorenko, E., Duncan, J., and Kanwisher, N. (2012a). Language-selective and domain-general regions lie side by side within Broca's area. *Curr. Biol.* 22, 2059–2062. doi: 10.1016/j.cub.2012.09.011
- Fedorenko, E., Nieto-Castanon, A., and Kanwisher, N. (2012b). Syntactic processing in the human brain: what we know, what we don't know, and a suggestion for how to proceed. *Brain Lang.* 120, 187–207. doi: 10.1016/j.bandl.2011.01.001
- Fedorenko, E., Nieto-Castanon, A., and Kanwisher, N. (2012c). Lexical and syntactic representations in the brain: an fMRI investigation with multi-voxel pattern analyses. *Neuropsychologia* 50, 499–513. doi: 10.1016/j.neuropsychologia.2011.09.014
- Ferreira, F., and Swets, B. (2005). "The production and comprehension of resumptive pronouns in relative clause "island" contexts," in *Twenty-First Century Psycholinguistics: Four Cornerstones*, ed. A. Cutler (Hillsdale, NJ: Lawrence Erlbaum Associates), 263–278.
- Fiebach, C. J., Schlesewsky, M., Lohmann, G., von Cramon, D. Y., and Friederici, A. D. (2005). Revisiting the role of Broca's area in sentence processing: syntactic integration versus syntactic working memory. *Hum. Brain Mapp.* 24, 79–91. doi: 10.1002/hbm.20070
- Friederici, A. D. (2011). The brain basis of language processing: from structure to function. *Physiol. Rev.* 91, 1357–1392. doi: 10.1152/physrev.00006.2011
- Friederici, A. D., Fiebach, C. J., Schlesewsky, M., Bornkessel, I. D., and von Cramon, D. Y. (2006). Processing linguistic complexity and grammaticality in the left frontal cortex. *Cereb. Cortex* 16, 1709–1717. doi: 10.1093/cercor/bhj106
- Friederici, A. D., Meyer, M., and von Cramon, D. Y. (2000). Auditory language comprehension: an event-related fMRI study on the processing of syntactic and lexical information. *Brain Lang.* 74, 289–300. doi: 10.1006/brln.2000.2313
- Friederici, A. D., Rüschemeyer, S. A., Hahne, A., and Fiebach, C. J. (2003). The role of left inferior frontal and superior temporal cortex in sentence comprehension: localizing syntactic and semantic processes. *Cereb. Cortex* 13, 170–177. doi: 10.1093/cercor/13.2.170
- Fromkin, V. A. (1971). The non-anomalous nature of anomalous utterances. *Language (Baltim)* 57, 27–52.
- Gibson, E. (2000). "The dependency locality theory: a distance-based theory of linguistic complexity," in *Image, Language, Brain*, eds Y. Miyashita, A. Marantz, and W. O'Neil (Cambridge, MA: MIT Press), 95–126.
- Glaser, Y. G., Martin, R. C., Van Dyke, J. A., Hamilton, A. C., and Tan, Y. (2013). Neural basis of semantic and syntactic interference in sentence comprehension. *Brain Lang.* 126, 314–326. doi: 10.1016/j.bandl.2013.06.006
- Golestani, N., Alario, F. X., Meriaux, S., Le Bihan, D., Dehaene, S., and Pallier, C. (2006). Syntax production in bilinguals. *Neuropsychologia* 44, 1029–1040. doi: 10.1016/j.neuropsychologia.2005.11.009
- Grew, T., Bornkessel, I., Zysset, S., Wiese, R., von Cramon, D. Y., and Schlesewsky, M. (2006). Linguistic prominence and Broca's area: the influence of animacy as a linearization principle. *Neuroimage* 32, 1395–1402. doi: 10.1016/j.neuroimage.2006.04.213
- Grodzinsky, Y., and Santi, A. (2008). The battle for Broca's region. *Trends Cogn. Sci. (Regul. Ed.)* 12, 474–480. doi: 10.1016/j.tics.2008.09.001
- Hagoort, P. (2005). On Broca, brain, and binding: a new framework. *Trends Cogn. Sci. (Regul. Ed.)* 9, 416–423. doi: 10.1016/j.tics.2005.07.004
- Haller, S., Radue, E. W., Erb, M., Grodd, W., and Kircher, T. (2005). Overt sentence production in event-related fMRI. *Neuropsychologia* 43, 807–814. doi: 10.1016/j.neuropsychologia.2004.09.007
- Hauser, M. D., Chomsky, N., and Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science* 298, 1569–1579. doi: 10.1126/science.298.5598.1569
- Hickok, G., Buchsbaum, B. R., Humphries, C., and Muftuler, T. (2003). Auditory-motor interaction revealed by fMRI: speech, music, and working memory in area Spt. *Cogn. Neurosci. J.* 15, 673–682. doi: 10.1162/089892903322307393
- Hickok, G., and Rogalsky, C. (2011). What does Broca's area activation to sentences reflect? *J. Cogn. Neurosci.* 23, 2629–2631. doi: 10.1162/jocn\_a.00044
- Humphries, C., Binder, J. R., Medler, D., and Liebenthal, E. (2006). Syntactic and semantic modulation of neural activity during auditory sentence comprehension. *Cogn. Neurosci. J.* 18, 665–679. doi: 10.1162/jocn.2006.18.4.665
- Humphries, C., Love, T., Swinney, D., and Hickok, G. (2005). Response of anterior temporal cortex to syntactic and prosodic manipulations during sentence processing. *Hum. Brain Mapp.* 26, 128–138. doi: 10.1002/hbm.20148
- Izawa, J., Rane, T., Donchin, O., and Shadmehr, R. (2008). Motor adaptation as a process of reoptimization. *J. Neurosci.* 28, 2883–2891. doi: 10.1523/JNEUROSCI.5359-07.2008
- Kaan, E., and Swaab, T. Y. (2002). The brain circuitry of syntactic comprehension. *Trends Cogn. Sci. (Regul. Ed.)* 6, 350–356. doi: 10.1016/S1364-6613(02)01947-2
- Kastner, S., and Ungerleider, L. G. (2000). Mechanisms of visual attention in the human cortex. *Annu. Rev. Neurosci.* 23, 315–341. doi: 10.1146/annurev.neuro.23.1.315
- Kitagawa, Y. (1994). *Subjects in Japanese and English*. Abingdon: Taylor & Francis.
- Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., and Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception* 36, 1.1–16.
- Lau, E., Stroud, C., Plesch, S., and Phillips, C. (2006). The role of structural prediction in rapid syntactic analysis. *Brain Lang.* 98, 74–88. doi: 10.1016/j.bandl.2006.02.003
- Lewis, R. L., Vasishth, S., and Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends Cogn. Sci. (Regul. Ed.)* 10, 447–454. doi: 10.1016/j.tics.2006.08.007
- Lieberman, P. (2001). Human language and our reptilian brain: the subcortical bases of speech, syntax, and thought. *Perspect. Biol. Med.* 44, 32–51. doi: 10.1353/pbm.2001.0011
- Mack, J. E., Meltzer-Asscher, A., Barbieri, E., and Thompson, C. K. (2013). Neural correlates of processing passive sentences. *Brain Sci.* 3, 1198–1214. doi: 10.3390/brainsci3031198
- Mazoyer, B. M., Tzourio, N., Frak, V., Syrota, A., Murayama, N., Levrier, O., et al. (1993). The cortical representation of speech. *J. Cogn. Neurosci.* 5, 467–479. doi: 10.1162/jocn.1993.5.4.467
- Menenti, L., Gierhan, S. M., Segaert, K., and Hagoort, P. (2011). Shared language overlap and segregation of the neuronal infrastructure for speaking and listening revealed by functional MRI. *Psychol. Sci.* 22, 1173–1182. doi: 10.1177/0956797611418347
- Menenti, L., Segaert, K., and Hagoort, P. (2012). The neuronal infrastructure of speaking. *Brain Lang.* 122, 71–80. doi: 10.1016/j.bandl.2012.04.012

- Meyer, M., Friederici, A. D., and von Cramon, D. Y. (2000). Neurocognition of auditory sentence comprehension: event related fMRI reveals sensitivity to syntactic violations and task demands. *Cogn. Brain Res.* 9, 19–33. doi: 10.1016/S0926-6410(99)00039-7
- Miller, E. K., and Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202. doi: 10.1146/annurev.neuro.24.1.167
- Moro, A., Tettamanti, M., Perani, D., Donati, C., Cappa, S. F., and Fazio, F. (2001). Syntax and the brain: disentangling grammar by selective anomalies. *Neuroimage* 13, 110–118. doi: 10.1006/nimg.2000.0668
- Novick, J. M., Trueswell, J. C., and Thompson-Schill, S. L. (2005). Cognitive control and parsing: reexamining the role of Broca's area in sentence comprehension. *Cogn. Affect. Behav. Neurosci.* 5, 263–281. doi: 10.3758/CABN.5.3.263
- Pallier, C., Devauchelle, A. D., and Dehaene, S. (2011). Cortical representation of the constituent structure of sentences. *Proc. Natl. Acad. Sci. U.S.A.* 108, 2522–2527. doi: 10.1073/pnas.1018711108
- Paulignan, Y., MacKenzie, C., Marteniuk, R., and Jeannerod, M. (1991). Selective perturbation of visual input during prehension movements. *Exp. Brain Res.* 83, 502–512. doi: 10.1007/BF00229827
- Price, A. R., Bonner, M. F., Peelle, J. E., and Grossman, M. (2015). Converging evidence for the neuroanatomic basis of combinatorial semantics in the angular gyrus. *J. Neurosci.* 35, 3276–3284. doi: 10.1523/JNEUROSCI.3446-14.2015
- Pykkänen, L., Bemis, D. K., and Elorrieta, E. B. (2014). Building phrases in language production: an MEG study of simple composition. *Cognition* 133, 371–384. doi: 10.1016/j.cognition.2014.07.001
- Rogalsky, C., and Hickok, G. (2009). Selective attention to semantic and syntactic features modulates sentence processing networks in anterior temporal cortex. *Cereb. Cortex* 19, 786–796. doi: 10.1093/cercor/bhn126
- Rogalsky, C., and Hickok, G. (2011). The role of Broca's area in sentence comprehension. *J. Cogn. Neurosci.* 23, 1664–1680. doi: 10.1162/jocn.2010.21530
- Rogalsky, C., Matchin, W., and Hickok, G. (2008). Broca's area, sentence comprehension, and working memory: an fMRI study. *Front. Hum. Neurosci.* 2:14. doi: 10.3389/neuro.09.014.2008
- Rogalsky, C., Rong, F., Saberi, K., and Hickok, G. (2011). Functional anatomy of language and music perception: temporal and structural factors investigated using functional magnetic resonance imaging. *J. Neurosci.* 31, 3843–3852. doi: 10.1523/JNEUROSCI.4515-10.2011
- Segaert, K., Kempen, G., Petersson, K. M., and Hagoort, P. (2013). Syntactic priming and the lexical boost effect during sentence production and sentence comprehension: an fMRI study. *Brain Lang.* 124, 174–183. doi: 10.1016/j.bandl.2012.12.003
- Segaert, K., Menenti, L., Weber, K., Petersson, K. M., and Hagoort, P. (2012). Shared syntax in language production and language comprehension—an fMRI study. *Cereb. Cortex* 22, 1662–1670. doi: 10.1093/cercor/bhr249
- Simmonds, D. J., Pekar, J. J., and Mostofsky, S. H. (2008). Meta-analysis of Go/No-go tasks demonstrating that fMRI activation associated with response inhibition is task-dependent. *Neuropsychologia* 46, 224–232. doi: 10.1016/j.neuropsychologia.2007.07.015
- Stromswold, K., Caplan, D., Alpert, N., and Rauch, S. (1996). Localization of syntactic comprehension by positron emission tomography. *Brain Lang.* 52, 452–473. doi: 10.1006/brln.1996.0024
- Suminski, A. J., Rao, S. M., Mosier, K. M., and Scheidt, R. A. (2007). Neural and electromyographic correlates of wrist posture control. *J. Neurophysiol.* 97, 1527–1545. doi: 10.1152/jn.01160.2006
- Swann, N., Tandon, N., Canolty, R., Ellmore, T. M., McEvoy, L. K., Dreyer, S., et al. (2009). Intracranial EEG reveals a time-and frequency-specific role for the right inferior frontal gyrus and primary motor cortex in stopping initiated responses. *J. Neurosci.* 29, 12675–12685. doi: 10.1523/JNEUROSCI.3359-09.2009
- Talairach, J., and Tournoux, P. (1988). *Co-planar Stereotaxic Atlas of the Human Brain*. New York, NY: Thieme Medical Publishers, 122.
- Tourville, J. A., Reilly, K. J., and Guenther, F. H. (2008). Neural mechanisms underlying auditory feedback control of speech. *Neuroimage* 39, 1429–1443. doi: 10.1016/j.neuroimage.2007.09.054
- Tyler, L. K., Wright, P., Randall, B., Marslen-Wilson, W. D., and Stamatakis, E. A. (2010). Reorganization of syntactic processing following left-hemisphere brain damage: does right-hemisphere activity preserve function? *Brain* 133, 3396–3408. doi: 10.1093/brain/awq262
- Ullman, M. T. (2004). Contributions of memory circuits to language: the declarative/procedural model. *Cognition* 92, 231–270. doi: 10.1016/j.cognition.2003.10.008
- Wilson, S. M., Demarco, A. T., Henry, M. L., Gesierich, B., Babiak, M., Mandelli, M. L., et al. (2014). What role does the anterior temporal lobe play in sentence-level processing? neural correlates of syntactic processing in semantic variant primary progressive aphasia. *J. Cogn. Neurosci.* 26, 970–985. doi: 10.1162/jocn\_a\_00550
- Ye, Z., and Zhou, X. (2009). Conflict control during sentence comprehension: fMRI evidence. *Neuroimage* 48, 280–290. doi: 10.1016/j.neuroimage.2009.06.032

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Matchin and Hickok. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# Brain asymmetry in the white matter making and globularity

Constantina Theofanopoulou \*

Department of General Linguistics, Universitat de Barcelona, Barcelona, Spain

## OPEN ACCESS

### Edited by:

Antonio Benítez-Burraco,  
University of Huelva, Spain

### Reviewed by:

Pedro Tiago Martins,  
Pompeu Fabra University, Spain  
Bridget Samuels,  
Pomona College, USA

### \*Correspondence:

Constantina Theofanopoulou,  
Department of General Linguistics,  
Universitat de Barcelona, Edifici  
Josep Carner, 5a Planta, Gran Via  
de les Corts Catalanes 585,  
08007 Barcelona, Spain  
constantinaki@hotmail.com

### Specialty section:

This article was submitted to  
Language Sciences,  
a section of the journal  
Frontiers in Psychology

**Received:** 27 July 2015

**Accepted:** 24 August 2015

**Published:** 10 September 2015

### Citation:

Theofanopoulou C (2015) Brain  
asymmetry in the white matter  
making and globularity.  
Front. Psychol. 6:1355.  
doi: 10.3389/fpsyg.2015.01355

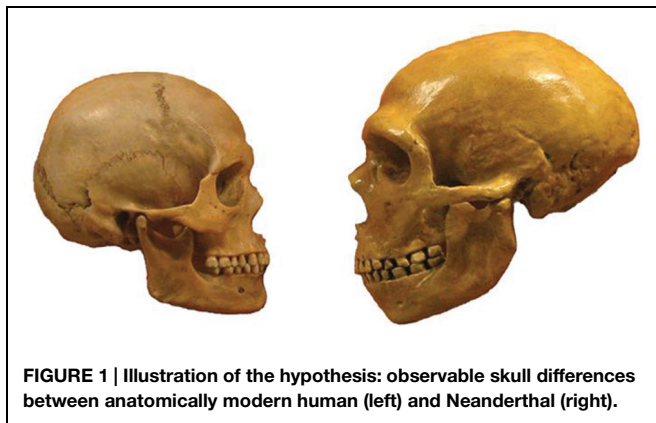
Recent studies from the field of language genetics and evolutionary anthropology have put forward the hypothesis that the emergence of our species-specific brain is to be understood not in terms of size, but in light of developmental changes that gave rise to a more globular braincase configuration after the split from Neanderthals-Denisovans. On the grounds that (i) white matter myelination is delayed relative to other brain structures and, in humans, is protracted compared with other primates and that (ii) neural connectivity is linked genetically to our brain/skull morphology and language-ready brain, I argue that one significant evolutionary change in *Homo sapiens*' lineage is the interhemispheric connectivity mediated by the Corpus Callosum. The size, myelination and fiber caliber of the Corpus Callosum present an anterior-to-posterior increase, in a way that inter-hemispheric connectivity is more prominent in the sensory motor areas, whereas "high- order" areas are more intra-hemispherically connected. Building on evidence from language-processing studies that account for this asymmetry ('lateralization') in terms of brain rhythms, I present an evo-devo hypothesis according to which the myelination of the Corpus Callosum, Brain Asymmetry, and Globularity are conjectured to make up the angles of a co-evolutionary triangle that gave rise to our language-ready brain.

**Keywords:** brain asymmetry, lateralization, skull, globularity, corpus callosum, white matter, brain rhythms, language

## Introduction

The general aim of this paper is to support the idea that the key underlying our human-specific cognitive profile is to be found in the changes that brought about a more globular brain shape. As far as I can tell, two scientific hypotheses have already been put forward claiming that it was essentially this globular shape that determined the brain profile of *Homo sapiens*: one, by Boeckx and Benítez-Burraco, 2014a,b), called it "globularity" and hypothesized that modifications in the fronto-parieto-thalamic network ought to be taken into account; the other, by Hublin et al. (2015), called it the "globularization phase," focusing, thus, on the developmental phase of the shaping, which, according to their findings, is due to bulging parietal and occipital bones (see Figure 1).

These two hypotheses can be understood as viewing the same changes from different perspectives. While Hublin et al. (2015) concentrate mostly on the evolutionary (occipito-parietal protrusion) and developmental (cerebellum) facet of the problem, Boeckx and Benítez-Burraco, 2014a,b; Benítez-Burraco and Boeckx, 2015) mainly focus on the link between globularity and language on a genetic level: their positing that the fronto-parieto-thalamic network might be of relevance seems to reflect the great majority of the findings that implicate the fronto



(B44/45)-parietal (BA22) lobes in linguistic processing, broadening this cortical network to its subcortical “afferent and efferent expansion,” namely the thalamus (Buzsáki, 2006; Theofanopoulou and Boeckx, 2015). The integration of the thalamus should not strike us as irrelevant at all, as it essentially relays the cerebellar input to the frontal lobe (BA44/45 included; for cerebello-thalamic connectivity: Leiner et al., 1989; Schmahmann, 1997; Engelborghs et al., 1998, for thalamico-BA44/45 connectivity: Ford et al., 2013; Bohsali et al., 2015) and its growth is correlated to associated expanded areas, like the parietal lobe (and more specifically the novel precuneus, Cavanna and Trimble, 2006; Bruner, 2014; Bruner et al., 2014). By the logic of co-evolution (namely that two tightly connected brain parts exert pressure to each other, affecting each other’s evolution), it can be deduced that both the cerebellum and the thalamus are crucially involved in the cognitive mechanisms that result in language.

In this paper, however, I will take the expansion of the cerebellum to be the guiding line for the following reasons: firstly, the cerebellum is directly connected to the evolutionarily-significantly splayed parieto-occipital bone; secondly, it is particularly in these posterior sensorimotor regions of the cortex where the Corpus Callosum permits interhemispheric connectivity thanks to its anterior-to-posterior increase in size, myelination and fiber caliber (Aboitiz et al., 1992; Doron and Gazzaniga, 2008); thirdly, this connectivity is taken to be crucial for permitting the rhythmic interhemispheric interplay observed in language processing sensorimotor networks (Morillon et al., 2010).

The frontal cortex will not figure much in my core hypothesis, as according to Barton and Venditti (2014), even though absolute and proportional frontal region size increased rapidly in humans, this change was tightly correlated with corresponding size increases in other areas and overall brain size; besides, research has demonstrated that the parieto-occipital fossa’s protrusion was the most decisive; lastly, as regards to evolutionary changes in frontal connectivity, Neubert et al. (2014) have shown that actually what differentiates humans’ and macaques’ frontal cortex is its coupling to posterior auditory areas (which is much stronger in humans), something that seems to justify my attention toward the posterior sensorimotor areas.

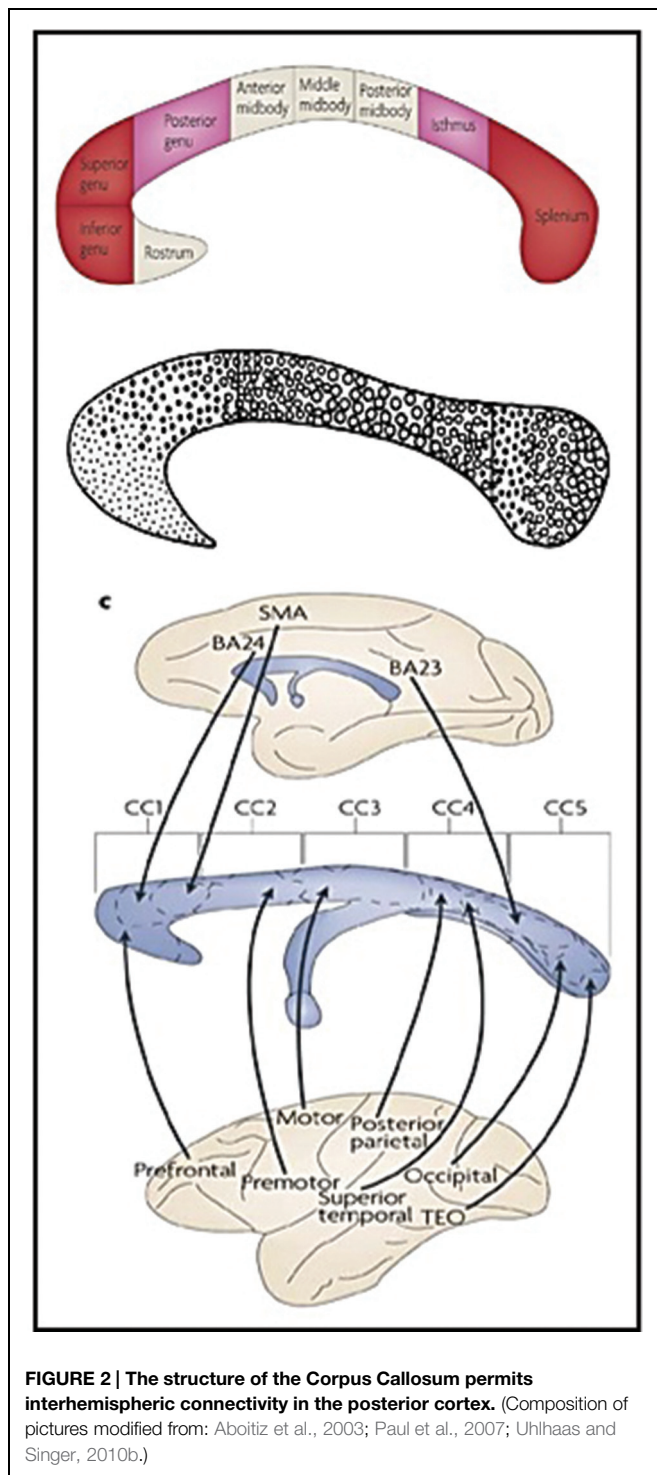
Ultimately, my objective is to bring out how -under this novel perspective- we could also make sense of the intricate idea of Brain Asymmetry (‘Lateralization’) and, thus, elucidate the hitherto unexplored relation of Brain Asymmetry-Corpus Callosum-Globularity. More specifically, I will argue that the architecture of the Corpus Callosum, allowing for interhemispheric connectivity in the posterior cortex (posterior temporal- posterior parietal- occipital cortex) and for intrahemispheric connectivity in the anterior- medial cortex (frontal- anterior parietal cortex-anterior/medial temporal cortex), suggests a novel way of capturing how brain asymmetry (a phylogenetically common trait) made it possible for our language- ready brain to arise (see Figure 2).

The structure of the paper will be the following: first, I will briefly introduce the idea of Brain Asymmetry- Lateralization (see Brain Asymmetry – Lateralization) and explain how it fits in the framework I wish to put forward; then I will look into how Asymmetry can be captured in terms of brain rhythms and analyze how the morphology of the Corpus Callosum renders this asymmetry possible (1.1). Next, I will draw attention to alpha and beta rhythms and propose a way in which they could constitute an overlooked window into both Ontogeny and Phylogeny (1.2), adducing supporting data from deficits that could be seen as speech-related oscillopathies (1.3). In Section “White Matter-Globular Brain Pattern,” I will delve into the idea of the White Matter begetting our Globular Brain Pattern during development. In the end, I will discuss how the posterior brain and skull co-evolved in *H. sapiens* (see Posterior Brain and Skull Enlargement in *Homo sapiens*).

## Brain Asymmetry – Lateralization

Brain Asymmetry, long thought to be human- specific, has been shown to lie along an evolutionary continuum (Fitch and Braccini, 2013), so that even the first biological pillar of the uniqueness of human language, namely its strong left-lateralization (Lenneberg, 1966) has fallen down. Comparative studies have suggested a left-hemispheric dominance for conspecific communication in a wide variety of species (Ocklenburg and Güntürkün, 2012), such as chimpanzees (Tagliabata et al., 2008), rhesus monkeys (Hauser and Andersson, 1994), dogs (Siniscalchi et al., 2008), mice (Ehret, 1987), sea lions (Böye et al., 2005), and frogs (Bauer, 1993). More tellingly, a left-dominance has been reported in canaries as regards to hypoglossal functions (Nottebohm, 1971), in zebra finches concerning vocal learning (Voss et al., 2007; Moorman et al., 2012) and in Bengalese finches for song discrimination (Okanoya et al., 2001). Another shared dominance worth mentioning is that of emotional processing in the right hemisphere (e.g., Önal-Hartmann et al., 2011). There is evidence that it obtains also in gelada babboons (Casperd and Dunbar, 1996), mangabeys (Baraud et al., 2009), rhesus macaques (Vermeire and Hamilton, 1998), chimpanzees (Parr and Hopkins, 2000), marmosets (Hook-Costigan and Rogers, 1998), and dogs (Siniscalchi et al., 2008).

Last but not least, asymmetry in motor behavior and more concretely, in handedness, has been erroneously thought to be human- specific and furthermore to imply -along with language-



a general left hemispheric dominance common to humans (Harris, 1991). On the one hand, left-motor-lateralization is not unique to humans: as Smaers et al. (2013) review, lateralization in motor behavior has been found in primates (Nudo et al., 1992; Bogart et al., 2012; Sun et al., 2012), non-primate mammals (Ehret, 1987; Rogers et al., 1994), birds (Vallortigara and Andrew, 1994), fish (Cantalupo et al., 1995; Bisazza et al., 1997), reptiles

(Engbretson et al., 1981; Hosoi et al., 2007) and amphibians (Bauer, 1993; e.g., pawedness in toads Bisazza et al., 1996), footedness in birds (Rogers and Workman, 1993) and finnedness in fish (Hori, 1993). Handedness, in particular, was recently shown to be present also in non-primate mammals (bipedal marsupials), something that challenges the notion that ‘true’ handedness is unique to primates (Giljov et al., 2015). On the other hand, the right-handedness rule would imply “that most left handed people display right hemispheric dominance for language, an assertion not validated by rigorous empirical studies (Knecht et al., 2000),” as Washington and Tillinghast (2015) observe. Ocklenburg and Güntürkün (2012) suggest there are both genetic and epigenetic factors we should take into account in the context of brain asymmetry. They provide the example of pigeons: their determined embryonic egg-position (genetic factor) permits only their right eye to be stimulated by light (epigenetic factor), resulting in left hemisphere superiority for visual object discrimination. (For a good experiment on how early navigational experience in pigeons affects lateralization, see Mehlhorn et al., 2010.) The authors finally suggest that similar “early spinal asymmetries could act as lateralized ‘precursors’ of asymmetrical cortical motor functions” in humans (such as prenatal bias on turning the head to the right); epigenetic factors though should not be overlooked, as their relevance in human handedness is much more important (unlike birds in the case of vision, humans are not genetically confined to using only one hand!). The lower incidence of left-handedness in countries where the left hand is associated with uncleanliness is a good example portraying how much epigenetic factors affect handedness (Zverev, 2006). Siding with Benítez-Burraco and Longa (2012), I conclude that “the relationships between right-handedness (structural and functional) brain lateralization, and language are perhaps not significant enough, or illuminating from an evolutionary perspective.”

With the aforementioned I wish to underline that human language lateralization is not due to a dominance of the left hemisphere for language as such: none of the hitherto known cognitive functions emerged during hominin evolution. Rather, they are phylogenetically shared, as one should expect given the conservation of brain rhythms across a wide range of species (Buzsáki et al., 2013; Boeckx and Theofanopoulou, 2015). This should lead us to consider the following: given that gray matter-subcortical parts of the brain are associated to sensorimotor and cognitive functions, and white matter modulates the distribution of action potentials among them and the neocortex (Fields, 2005, 2008), it is probably this modulatory function that gives to the core-cognitive functions (gray matter) the level of complexity detected only in our species and required for language.

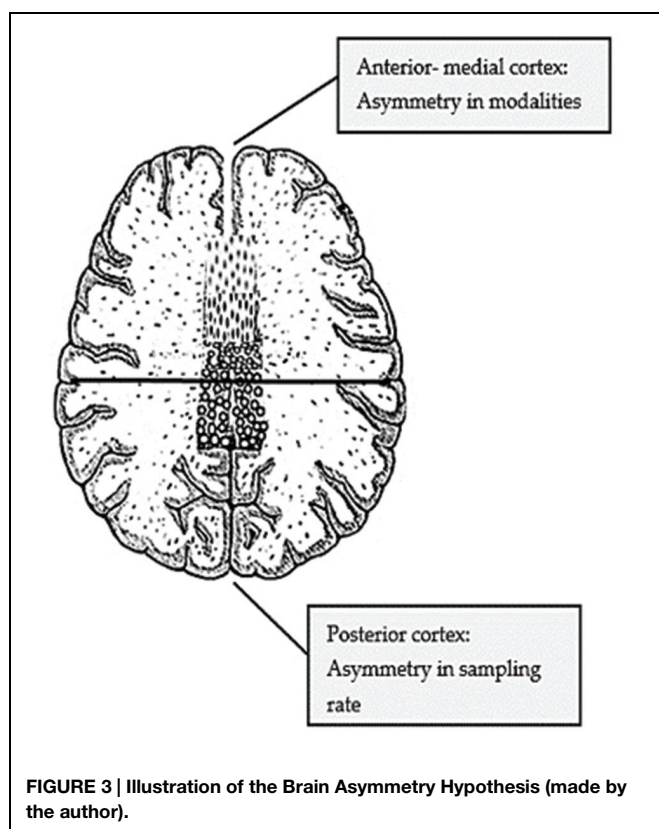
I agree with Ocklenburg and Güntürkün (2012) in that white matter might be an overlooked window to brain asymmetry-issues: “A common conception is that functional asymmetries are a consequence of structural asymmetries in the brain . . . . research . . . has focused on macroscopic gray matter asymmetries . . . evidence from recent studies in animal models suggests that structural asymmetries in connectivity patterns of homologous regions in the two hemispheres may be of greater functional relevance.” Another reason to believe so is the



developmental nature of white matter's myelination; occurring relatively slowly over the lifespan (Hynd et al., 1995), white matter constitutes a perfect mirror candidate of the developmental nature of human language acquisition. Besides, as I noted above, the chemical mechanisms of myelination are decisive for axons' generating action potentials, something that can be directly associated with the oscillatory basis of language I will shortly highlight.

In what follows, I will try to illustrate the relevance of the brain's largest white tissue structure, i.e., the Corpus Callosum, in Brain Asymmetry, and afterward provide a link to Globularity. Let me clarify that the Corpus Callosum was not chosen merely because of its size, but because of its decisive position in the brain and its human- specific structure. The size, myelination and fiber caliber of the Corpus Callosum presents an anterior-to-posterior increase (Doron and Gazzaniga, 2008), resulting in interhemispheric connectivity being more prominent in the sensory motor areas, whereas "high- order" areas are more intrahemispherically connected (Figure 3). Studies comparing humans' and monkeys' corpora callosa tellingly revealed that in humans the proportion of large diameter fibers in callosal regions that interconnect primary sensory areas is higher than in macaques (Aboitiz et al., 1992) and that the fiber organization has nothing in common with the callosal organization reported in monkeys (Jones et al., 1978; Killackey et al., 1983).

The hypothesis I wish to put forward is the following: it is the structure of the Corpus Callosum that makes humans' brains display a sophisticated, selective asymmetry:

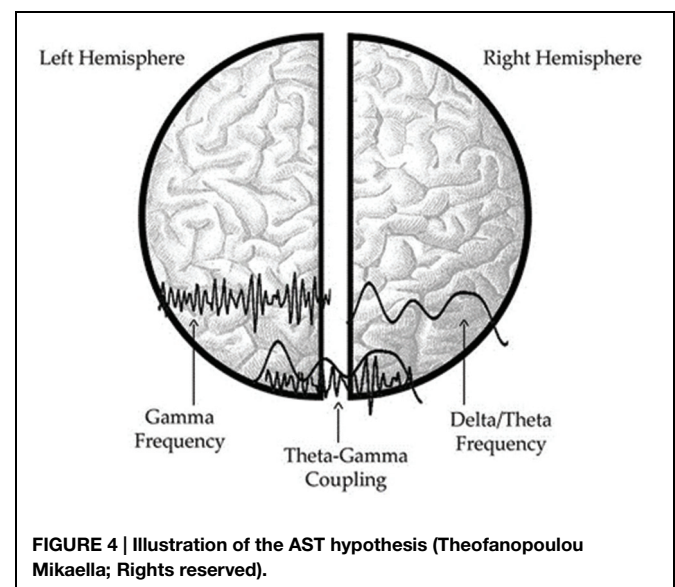


in the anterior/medial cortex, where callosal fibers are narrow and intrahemispheric connectivity is enhanced, asymmetry is expressed at the level of small- world networks, i.e., cognitive functions appear to be lateralized as modules (e.g., the default network in the left and the attentional in the right hemisphere: Wang et al., 2008; De Schotten et al., 2011); in the posterior cortex, there is no such asymmetry, as visual, auditory, and motor functions appear in both hemispheres. What makes the posterior cortex asymmetrical is to be found in the hemispheres' refinement toward processing input of specific 'sampling rate' (temporal-faster rate sampling executed in the left hemisphere and spectral-slower rate in the right). In the coming section, data fostering my hypothesis will be added (see Figure 3).

### Asymmetry in the Dynome and the Corpus Callosum

Contemporary neural models of auditory language processing proposed that the two hemispheres are differently specialized in either temporal (left hemisphere) or spectral (right hemisphere) resolution (Zatorre et al., 2002), or in other terms that they differ in terms of their preferred "sampling rate" with the left hemisphere being well suited for faster- rate sampling and the right for slower rate (Poeppel, 2003; Hickok and Poeppel, 2007, 2015; Morillon et al., 2010). According to Celesia and Hickok (2015), "These two proposals are not incompatible as there is a relation between sampling rate and spectral vs. temporal resolution: rapid sampling allows the system to detect changes that occur over short timescales, but sacrifices spectral resolution, and vice versa."

More concretely, Morillon et al. (2010) found that there are two auditory speech sampling mechanisms working in parallel: while syllabic parsing of the input (slow- rate delta-theta oscillations,  $\sim 4$  Hz) is predominantly assigned to the right hemisphere, the left hemisphere has been shown to have a primacy for the processing of phonemic input (fast-rate gamma oscillations,  $\sim 40$  Hz; see Figure 4). A dynamic interplay is assumed to allow for the timely coordination of both information types, namely for fast phonemic gamma being





modulated by syllabic theta oscillations. The hypothesis of functional asymmetry concerning hemispheres' preferential cues is known as the AST (asymmetric sampling in time) hypothesis (Poeppel, 2003; Ghazanfar and Poeppel, 2014).

Tellingly, even though this asymmetry was most pronounced in the auditory cortex, motor areas also express natural oscillatory activity that corresponds to the same rates: intrinsic jaw movements oscillate at delta/theta oscillations, thus presenting an overlapping parsing with the syllabic network, while the phonemic fast gamma oscillations underlie tongue and formant transition movements (e.g., trill at 35–40 Hz; Morillon et al., 2010).

Another significant finding of the experiment conducted by Morillon et al. (2010) was a strong intrinsic asymmetry (also manifest at rest) between the articulatory (left hemisphere) and the hand motor cortex (right hemisphere). This asymmetry is suggested to be phylogenetically “inherited,” probably because of the long shared sinistral pharyngeal muscle control on the one hand and the dextral hand gestures control on the other. (Let me parenthesize here to point out from another perspective that erroneously right-handedness has attracted all the attention: what is significant for language is hand gestures accompanying language and not handedness *per se*.)

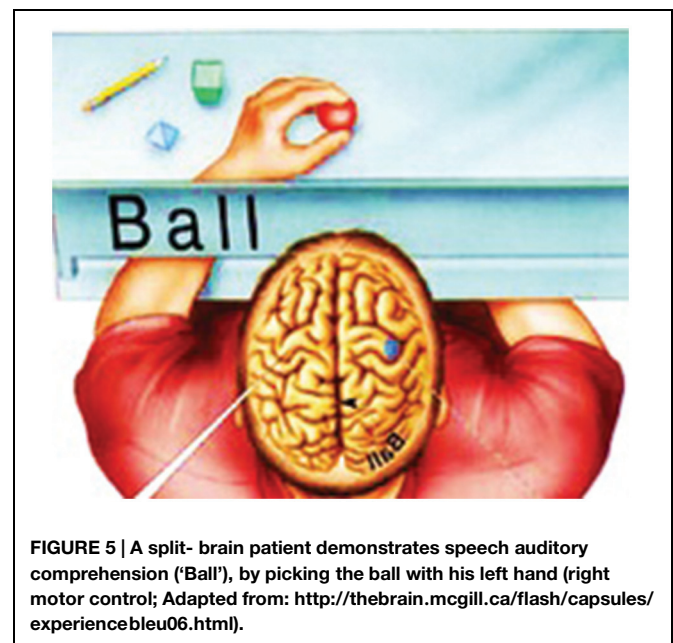
In addition, there are cases of other, non-human, even non-vocal-learning species, whose lip-smacking is tuned into the same slow oscillatory cycles, present in human syllabic sampling (such as the Gelada Baboons *Theropithecus gelada*, Bergman, 2013, see also Ghazanfar et al., 2012). It would then be critical to find out whether the Gelada Baboons display also a human-like right dominance of these lip movements and whether they are coupled with other faster oscillatory cycles, subserving communicatory processes. With the latter, I don't mean to imply that our linguistic profile is due to our capacity of housing spectral and temporal information within a narrow time-window, since this competence is again found to be present in other species: mustached bats exhibit the same oscillatory asymmetry in echolocation processing (Washington and Tillinghast, 2015). Rather, what all these phylogenetic observations are meant to highlight is that we indeed share both generic and elemental mechanisms with other species: the key to our questions is to be found in how *H. sapiens* ‘coupled’ the modalities inherited. Adopting a Darwinian thinking, it seems indeed plausible that the connectivity across and within modalities afforded by the peculiar structure of the Corpus Callosum is evolutionarily significant. In the case of auditory processing, it has been experimentally shown that it is the posterior Corpus Callosum that gives rise to this linguistically-crucial theta-gamma coupling (Rumsey et al., 1996; Pollmann et al., 2002; Nosarti et al., 2004). Even early dichotic listening experiments on patients with Corpus Callosum abnormalities had specified that agenesis in the splenium is pertinent to aberrant auditory interplay (Sugishita et al., 1995; Pollmann et al., 2002). Friederici et al. (2007) tellingly observe that “an intact posterior third of the *C[orpus] C[allosum]* connecting temporal regions is a necessary precondition for a prosody-induced N400 mismatch effect. Lesions in the anterior two-thirds of the CC that connect frontal regions, in contrast, can cause a modulation of the prosody-induced

mismatch effect but cannot eliminate the effect.” Sammler et al. (2010) conducted an experiment with two groups of patients with lesions either in the anterior or the posterior Corpus Callosum: the latter did not exhibit the expected mismatch between segmental (temporal) and suprasegmental (spectral) features of language.

Apart from the role of the Corpus Callosum in auditory processing, all these studies also hint at the involvement of the right hemisphere in language processing (for a review see Lindell, 2006). Fully consistent with my hypothesis, Overath et al. (2015), after finding that the effect of speech segment length was robust in both hemispheres, inferred the following: “It is therefore possible that laterality effects are driven more by higher order linguistic processing demands than by speech analysis *per se*.”

From this perspective, we can also explain the initial results from split-brain patients (Corballis, 1998), that apparently confirmed the linguistic incompetence of the right hemisphere: callosotomy patients could not verbally answer to language questions presented to isolated right hemispheres, because the articulatory (only) ability is left-dominant. However, when asked for non-verbal responses, the patients demonstrated speech auditory comprehension, by picking with their left hand (right motor control) the object (uttered by the experimenter) among an array of objects (see Figure 5).

Lastly, the idea of lateralization in terms of synchronic activity in the posterior cortex-mediated by large-diameter, fast-conducting callosal regions- was firstly formulated by Aboitiz et al. (2003). However, back then there was no evidence reporting a right-dominance for spectral and a left-dominance for temporal processing, so that the authors made the following conclusion: “There is not yet evidence for the existence of synchronic ensembles during performance in working memory tasks, but 40 Hz synchronic oscillations have been reported during linguistic performance. . . The role of synchronic activity in working



**FIGURE 5 | A split-brain patient demonstrates speech auditory comprehension ('Ball'), by picking the ball with his left hand (right motor control; Adapted from: <http://thebrain.mcgill.ca/flash/capsules/experiencebleu06.html>).**

*memory processes, be they linguistic, auditory or visual, urgently needs to be investigated.*" The proposal of dynamic asymmetry presented here should be really close to what Aboitiz et al. (2003) had in mind.

### Alpha and Beta Rhythms- a Window to Ontogeny and Phylogeny of Mirror Neurons

Although research has mostly focused on the rhythms that have a discernible effect in auditory processing (delta/theta and gamma oscillations), alpha and beta rhythms are suggested to be equally implicated but in a reverse mode: by being suppressed.

This shift of interest toward the significance of these two rhythms is apparent if one pays attention to the way alpha/beta suppression is treated by Poeppel and colleagues: Doelling et al. (2014) investigated to what extent the oscillation-based envelope tracking (discussed above) also mediates the relationship between sharpness and intelligibility; the results supported this correlation but the authors also noted that another rhythm (alpha) was suppressed in their experiment. They limited their observation to mentioning Obleser and Weisz (2012), who have shown that alpha power suppression is related to intelligibility, but 1 year after they sought to put this to test (Luc and Arnal, 2014). Interestingly, they found that the typical increase in theta band was always followed by a broadband suppression of alpha (9–14 Hz) and beta (15–25 Hz) bands and suggested that *"the brain exploits the level of post-stimulus alpha suppression as internal evidence to determine how well the stimulus matched the prediction."*

If we now turn back to the experiments made by Obleser and Weisz (2012), alpha rhythm seems to have far-reaching implications in auditory processing: on the one hand, it has been shown to be suppressed to reinforce acoustic intelligibility, but on the other hand, alpha rhythm is enhanced during auditory memory retention (Knecht et al., 2000; Jensen et al., 2002; Obleser et al., 2012; Theofanopoulou and Boeckx, 2015). These findings suggest that alpha rhythm moderates the interplay between working memory/attention and intelligibility. Tellingly, when the auditory memory is overloaded (hence alpha rhythm enhancement is further employed), acoustic degradation affects processing, because of the non-canonical ellipsis of alpha suppression (Obleser et al., 2012). This is also consistent with studies reporting an increased activation of the cerebellum in high-load tasks, suggesting a prominent role of the cerebellum in working memory processing (Kirschen et al., 2010; Stoodley et al., 2012; Luis et al., 2015).

On the grounds that Morillon et al. (2010) regard the synchronization of the motor (jaw movements/trill or formant transitions) and auditory (syllabic/phonemic) modalities as significant, I take it that there must be such a correlation also between the motor and auditory alpha and beta suppression. This conclusion can be reached thanks to studies focusing on the Mirror Neurons System: a system of neurons which were first thought to be activated only during action execution and action-observation (Rizzolatti and Craighero, 2004). However, recent experiments leave no doubt that mirror neurons (MN) are also implicated in

auditory processing (Cuellar et al., 2012) and sensorimotor learning (Catmur et al., 2007; Hickok and Hauser, 2010). At present, it is acknowledged that MN integrate cross-modal information and crucially all the information that has been said to be involved in language processing (Senkfor, 2002; Molnar-Szakacs and Overy, 2006).

What is even more relevant to the present paper is that EEG and MEG studies report a suppression of alpha/mu and beta-band activity in the sensorimotor area, among other areas (Rizzolatti and Luppino, 2001; Muthukumaraswamy and Johnson, 2004; Ulloa and Pineda, 2007; Pineda, 2008; Perry and Bentin, 2009; Perry et al., 2010; Frenkel-Toledo et al., 2013; Lange et al., 2015; from now on, I will focus on the mu-alpha rhythm suppression, as beta-band suppression has only very recently been shown to be involved in the context of MN; see Lange et al., 2015).

There is corroborating evidence that alpha/mu rhythm desynchronization in the sensorimotor system appears early in infancy and its functional properties are so strongly modulated by maturation, that the sensorimotor system evolves from a random (in infants) to a "small-world" organization (in children and adults; Ferrari et al., 2009; Marshall and Meltzoff, 2011; Cuevas et al., 2014; Berchicci et al., 2015). This "small-world" networking is achieved by a wiring pattern in the brain that is thickly intra-connected locally and sparsely interconnected globally (Changizi, 2001, p. 571; Karbowski, 2003; Sporns and Kötter, 2004; Sporns and Zwi, 2004). Pineda (2005) proposed that *"mu rhythms represent an important information processing function that links perception and action-specifically, the transformation of 'seeing' and 'hearing' into 'doing'."*

In addition, if we go back to Morillon et al. (2010), they suggest that *"inherent auditory-motor tuning at the syllabic rate and acquired tuning at the phonemic rate are also compatible with two recognized stages of language development in infants; an early stage with production of syllables that does not depend on hearing (also observed in deaf babies), followed by a later stage in which infants match their phonemic production to what they hear in caregiver speech."* It seems to me that alpha/mu rhythm's maturation in development could indeed be the key for the interplay of the two sampling rates discussed, given that it is directly connected to the maturation and myelination of the white matter: Jann et al. (2012) found positive correlations of Fractional Anisotropy with alpha frequency within the splenium of the corpus callosum. Let me reiterate that the splenium, being at the posterior-myelinated part of the Corpus Callosum, presents an overlap of myelin water fraction with Fractional Anisotropy values within their thick axons that permits fast signal conductance. It is, furthermore, noteworthy that Miller (1994) sees such a connection between myelination, alpha frequency and intelligence that he put forward a brain myelination hypothesis of intelligence.

More importantly, these ontogenetic observations can be linked also to phylogenetic issues, which could result in interesting future experiments. It is known that MN were originally detected in monkeys' area F5. What is not yet appreciated is that MN respond to the observation of lip-smacking and hand-actions (Ferrari et al., 2003) by inhibiting

mu rhythm (Vanderwert et al., 2013; Cook et al., 2014). Given that there are hypotheses -in the context of human ontogeny- proposing that MNS in humans pass from being purely visual to multimodal (Iacoboni and Dapretto, 2006; Ferrari et al., 2009), it could be conjectured that this developmental shift, mediated by the myelinated posterior Corpus Callosum, was crucial for our linguistic cognition. (For data fostering this idea, see Autism Spectrum Disorder.)

## Deficits as Speech- Related Oscillopathies

### *Autism spectrum disorder*

It shouldn't strike us as strange that a plethora of evidence in line with the above comes from autism, where actually the developmental process is most obviously affected. Murphy et al. (2014) found that alpha- band deployment was severely impaired, giving rise to increased distraction, and Jochaut et al. (2015) encountered that ASD patients, instead of down- regulating gamma activity by theta, presented an opposite dependency such that gamma and theta- coupling jointly increased out of physiological ranges.

In light of what has been said about the implication of these rhythms in speech processing, it is clear that dysfunctional theta/gamma coordination and alpha suppression would disrupt the alignment of neuronal excitability with syllabic onset, compromising speech decoding.

The data seem also in consonance with what has been conjectured before about an earlier visual system which later becomes multi- modal: Damarla et al. (2010) show that ASD patients display more activation in visuospatial (bilateral superior parietal extending to inferior parietal and right occipital) areas, something that possibly indicates a compensatory role of visual processing during speech perception. The latter is supported by experiments according to which ASD subjects extensively explore the mouth region in face- to-face situations (Klin et al., 2002), and use specific attention modes to enhanced local visual processing (Schwarzkopf et al., 2014).

Regarding the Corpus Callosum in ASD, there is a abundance of studies proving that its size is degenerated in the posterior areas (Alexander et al., 2007; Just et al., 2007, among others). Moreover, the fact that abnormalities in the parietal lobes (Courchesne et al., 1993) and the posterior fossa (Courchesne et al., 1994) have been detected in infantile Autism lends credence to the contention that the enlarged brain and skull areas co-evolved in *H. sapiens*.

### *Schizophrenia*

Also in schizophrenia abnormal neural oscillations and synchrony has been associated with less organization in subdivision of the corpus callosum than controls (Uhlhaas and Singer, 2010a and references therein). Most studies have focused on deficits in the generation and maintenance of coherent gamma- range oscillations (Light et al., 2006; Minzenberg et al., 2010; Kirihaara et al., 2012). However, Moran and Hong (2011) after reviewing EEG studies on Schizophrenia and describing some of the key functional roles exerted by gamma, low frequencies, and their cross-frequency coupling, conclude that

even isolated alterations in gamma or low frequency oscillations may impact the interactions of high and low frequency bands.

Turning now to the Corpus Callosum: Leroux et al. (2015) revealed that reduced leftward functional lateralization for language in patients with schizophrenia was correlated with altered callosal integrity, reflecting decreased, and/or slower interhemispheric communication. In addition, Peters and Karlsgodt (2015) concluded that aberrant interhemispheric communication in schizophrenia is due to disrupted maturation at adolescence, with later changes likely due to disease neurotoxicity or to abnormal or excessive aging effects. In agreement with my hypothesis, neuroimaging studies showed lower callosal integrity (through FA or RD) in either the whole Corpus Callosum (Miyata et al., 2010; Knöchel et al., 2012; Freitag et al., 2013) or, more specifically, in the splenium region (Kyriakopoulos et al., 2008; Holleran et al., 2014; Balevich et al., 2015).

## White Matter- Globular Brain Pattern

In this section I will try to make clear how the development of the white matter underlies, or rather, co-evolves with the growth of our brain. More specifically, I will try to show how the processes of myelination and energy allocation (thermodynamics) of the white matter can shed light on our pursuit of what determines the shape of our brain. [The reasons why I am using the terms "shape" or "pattern" instead of "size" are well- explained in Boeckx and Benítez-Burraco (2014a) and Hublin et al. (2015). Suffice it to mention here one of their arguments: over the course of the past 30 000 years brain size declined slightly in recent *H. sapiens*, hence it strikes me as quite biased to keep focusing on brain's size solely.]

The idea is based on a statement in Boeckx and Benítez-Burraco (2014a): "*if the brain grows differently, it wires differently.*" I'd prefer to think of this in terms of allometric evolution, and say that "differences in brain growth and wiring co-evolve." In line with Buckner and Krienen (2013), I deem that the most telling wiring "element" can be found in the context of myelination and synaptic plasticity. As they note: "*Myelination of the cerebrum is delayed relative to other brain structures and in humans is globally protracted compared with other primates, including chimpanzees ... these collective observations suggest that the expanded cortical mantle of the human brain comprises networks that widely span the cortex without consistent feedforward/feedback connectivity and, further, that these circuits mature late into development.*"

In a similar vein, Hublin et al. (2015) take the extended period of growth during ontogeny and the delayed maturation of brain structure to contribute to our brain shape and its cognitive complexity. Prolonged human development is consequently thought to be key for the globularization developmental phase, present only in *H. sapiens* (Gunz et al., 2010, 2012). Furthermore, the fact that in humans myelination of the cortical axons is slow during childhood and extends beyond late adolescence allows their brain to "wire" while interacting with an enriched physical and cultural environment, viz. while being exposed to a vast variety of stimuli.



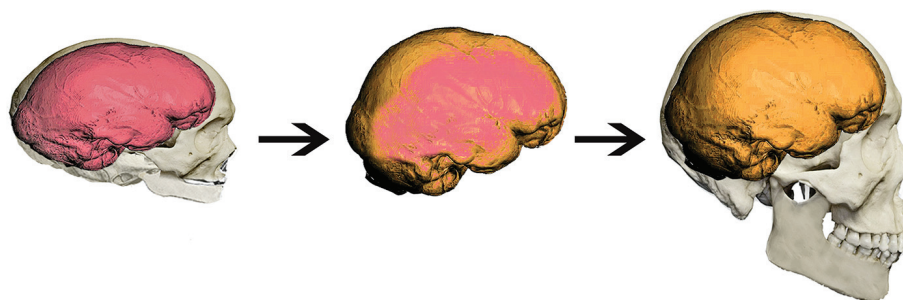
Even though studies showing that white matter volume increased in *H. sapiens* (Schoenemann et al., 2005; Sakai et al., 2011) are of relevance, Ventura-Antunes et al.'s (2013) observations should call our attention: building on Mota and Herculano-Houzel (2012), they drew the conclusion that cortical size is not only proportional to white matter volume, but to the average caliber and longitudinal tension along the axons of the white matter. This is how they explain that cortical size in rodents and primates scales differently: while rodents' brains wire with constant connectivity fraction, as a uniform network with the addition of isometrically longer fibers, primates' brains scale as a small-world network, growing through the addition of nodes that are densely intra-connected locally but only sparsely interconnected globally (Changizi, 2001; Karbowski, 2003; Sporns and Kötter, 2004; Sporns and Zwi, 2004).

Crucially, my hypothesis seems to fit very well in this picture. But the way I construe the role of the Corpus Callosum has little to do with the traditional view represented by Ringo et al. (1994). These authors set forth the idea that the strategy large brains use to compensate their conduction delays in transcallosal information transfer is not to increase the inter-hemispheric processing that depends solely on the Corpus Callosum, but to increase the intra-hemispheric amount of fibers that connect local lateralized networks. Ringo et al. (1994) and Hänggi et al. (2014) made a direct correlation between this observation and brain size; however, neither this small-world strategy nor large brain size is a specific trait of *H. sapiens*. In my opinion, humans' identifying features should be sought in the changes our neural wiring manifests. According to the hypothesis put forward in this paper, the Corpus Callosum is of great relevance, given that it displays a unique structure, which is related intrinsically to language processing coupling I described above. Let me just remind the reader that in humans the proportion of large diameter fibers in callosal regions that interconnect primary sensory areas is higher than in macaques (Aboitiz et al., 1992) and that the fiber organization of the Corpus Callosum has nothing in common with the callosal organization reported in monkeys, where the density of callosal connections varies according to body part within sensory representations (Jones et al., 1978; Killackey et al., 1983).

In order to conceive well of what connectivity means in brain terms, we should pay attention not only to the 'wiring' but also to the 're-wiring' of the brain. With the latter I am referring to the crucial phase of synaptic pruning in development, which proves to be very pertinent to my hypothesis. Moreover, viewed through the prism of energy allocation and thermodynamics in the brain, its relevance to brain asymmetry becomes conspicuous.

Indeed both Ventura-Antunes et al. (2013) and Hublin et al. (2015) hinted at the bearing of energy in the context of brain development. Human brains appear to use their prolonged development as a strategy to counterbalance large brains' energetic costs. The brain is extremely thermoregulated and vulnerable to energy shortages during development, as it requires circa 66% of the basal metabolic rate for functioning and maintenance by 4.2–4.4 years, when the brain approaches its adult size and synaptic densities are maximal (Holliday, 1986; Kuzawa, 1998). This exuberance of synapses is said to be needed to allow the synapse removal required for neural network refinement (Innocenti, 1995; Innocenti and Price, 2005; from an evolutionary standpoint, thermoregulation merits additional attention, considering that Neanderthals had a different endocranial heat dissipation pattern, when compared with modern humans, but a comparable amount of heat production, something that, according to Bruner (2014), could possibly be associated with the extinction of the first).

Myelination and synaptic pruning co-operate to adjust the energy consumption of the brain. Skoyles (2012) puts it boldly: *"This relative delay of myelination maturation is also consistent with small world connectivity refinement occurring particularly in the later stages of neuromaturation during adolescence, in which distant connections are pruned to create a more hub based connectivity. ... such network refinement depends upon the synaptic pruning that 'rewires' the local area neural networks formed between neighboring area neurons."* More importantly, Skoyles also provides a link between myelination and energy efficiency of axon transmission that can be directly linked to my hypothesis *"for the passage of each spike, a 0.5  $\mu$ m unmyelinated axon costs about 12-fold more in energy than when that spike is passed through a myelinated one."* Finally, if myelination reduces energy costs for interhemispheric communication, as it follows from the myelination and fiber structure of the



**FIGURE 6 | Modern human newborns have elongated brain-skull shape.** During the 'Globularization Phase' (in the middle) shape changes include a relative expansion of the cerebellum and parietal bulging. Modern human adults display finally a globular cranial and endocranial shape. (Mikaella Theofanopoulou; Rights reserved).



Corpus Callosum, synaptic pruning allows for neuron rewiring changes that refine anterior and medial cortex to establish and refine its intrahemispheric networks (Chklovskii et al., 2004; Skoyles, 2012).

These data can also be related to how the rhythms of the brain are tuned. Feinberg and Campbell (2010) found that the increase of myelination of long axonal fibers during adolescence results in long-range connectivity through reduced slow-wave activity (delta, theta) and decreased energy consumption. According to them and Uhlhaas and Singer (2010b), developmental changes correlate with the precision of rhythmic synchrony. More concretely, Zaehle and Herrmann (2011) observed a positive correlation between posterior callosal white matter density and inter-hemispheric frequency of visually evoked gamma oscillations, indicating a clear nexus between the connectome and the dynamo (Boeckx and Theofanopoulou, 2014). Finally, on the grounds that Barbato and Kinouchi (2000) find a great relationship between optimal pruning and the first learning experiences, it can be deduced that the linguistic input actually pilots the brain's development.

### Posterior Brain and Skull Enlargement in *Homo sapiens*

According to Hublin et al. (2015) “modern humans developed a more globular shape of the brain primarily resulting from a bulging in the parietal areas and a ventral flexion. In addition, modern humans display a proportionally larger cerebellum, larger olfactory bulbs and temporal lobe poles, and a wider orbitofrontal cortex” (see Figure 6).

For reasons I mentioned in the Introduction, I take the posterior cortex to be of more evolutionary significance than the frontal cortex; with this I don't mean to downgrade the acknowledged importance of the latter in higher-order language processing. It is noteworthy though that recent experiments show that the more eminent difference between humans' and monkeys' frontal cortex is its stronger connectivity with the sensorimotor cortex and not within frontal areas (Neubert et al., 2014). If we interpret these data in terms of co-evolution, we can say again that the anterior cortex co-evolved under the pressure of posterior cortex's enlargement.

Turning now to the posterior cortex, my hypothesis is fostered by the findings of Bruner (2010): “as brain size increases, the parietal lobes undergo relative flattening in non-modern humans. This pattern is stressed in Neanderthals, which show, however, a certain widening of the parietal volumes. Only *H. sapiens* shows a generalized enlargement of the entire parietal surface.” Furthermore, the bulging parietals of modern humans have been linked to evolutionary reorganization of deep parietal brain areas that gave rise to the novel precuneus (see Cavanna and Trimble, 2006 for the role of precuneus in cognition).

Buckner and Krienen (2013) review an array of studies concluding that the most telling change in the evolution of our lineage (that can be connected to cortical expansion) is the enlargement of the cerebellum (expressed mostly in the dentate nucleus) and its extensive projections to association cortex. In their final assessments they note: “The cerebellar

association zones are disproportionately expanded in humans, but the functional origins and importance of cerebellar expansion remain unresolved. Adaptionist ideas... seek explanations for cerebellar enlargement as a specific, selected feature of evolution.” Far from suggesting that the hypothesis presented here resolves the issue, I take the correlation between the myelination of the posterior Corpus Callosum and the enlargement of the cerebellum to be an insightful window into the evolution of our brain.

I side with Barton and Venditti (2014), when they propose the following: “cerebellar specialization was a far more important component of human brain evolution than hitherto recognized and that technical intelligence was likely to have been at least as important as social intelligence in human cognitive evolution. Given the role of the cerebellum in sensory-motor control and in learning complex action sequences, cerebellar specialization is likely to have underpinned the evolution of humans' advanced technological capacities, which in turn may have been a preadaptation for language.”

Essentially, as brains grow during ontogeny, the bones of the skull accommodate the expanding brain. The protrusion of the posterior cranial fossa in modern humans presents a good correlation with the cerebellar lobes, and the bulging parietal bones with the parietal lobe and specifically the precuneus.

It is also remarkable that some cranial changes have been associated with the Corpus Callosum: “In terms of evolution, shape and position of the corpus callosum are influenced by the general endocranial architecture, mainly by the flexion of the cranial base” (Bruner et al., 2012). Furthermore, considering that the tentorium cerebelli rotates inferoposteriorly in human fetuses (Jeffery, 2002) and that the antero-posterior stretching of the Corpus Callosum length was shown to vary in humans “due to the association between splenium and the anterior insertion of the tentorium cerebelli, caused by spatial proximity and consequent biomechanical relationships” (Bruner et al., 2012), there can be an overlooked relationship between some cranial fossa and the Corpus Callosum.

To conclude, I have argued in this paper that the special morphology of our Corpus Callosum provides an explanatory link between the (selective) Asymmetry long thought to be the key to understanding the evolution of our specific mode of cognition and the growth pattern that results in a globular brain(case), which sets us apart from other primates. On the basis of the evidence reviewed here, it can be said that accounts that ignore the critical role and anatomical position of the Corpus Callosum fail short of capturing what makes our brain's language-ready.

### Acknowledgments

I am indebted to Cedric Boeckx for his constant encouragement and help. Many thanks also to Antonio Benítez-Burraco for helpful discussion. Preparation of this work was supported by funds from the Spanish Ministry of Economy and Competitiveness (grants FFI2013-43823-P and FFI2014-61888-EXP).

## References

- Aboitiz, F., López, J., and Montiel, J. (2003). Long distance communication in the human brain: timing constraints for inter-hemispheric synchrony and the origin of brain lateralization. *Biol. Res.* 36, 89–99. doi: 10.4067/s0716-97602003000100007
- Aboitiz, F., Scheibel, A. B., Fisher, R. S., and Zaidel, E. (1992). Fiber composition of the human corpus callosum. *Brain Res.* 598, 143–153. doi: 10.1016/0006-8993(92)90178-c
- Alexander, A. L., Lee, J. E., Lazar, M., Boudos, R., DuBray, M. B., Oakes, T. R., et al. (2007). Diffusion tensor imaging of the corpus callosum in Autism. *Neuroimage* 34, 61–73. doi: 10.1016/j.neuroimage.2006.08.032
- Balevich, E. C., Haznedar, M. M., Wang, E., Newmark, R. E., Bloom, R., Schneiderman, J. S., et al. (2015). Corpus callosum size and diffusion tensor anisotropy in adolescents and adults with schizophrenia. *Psychiatry Res.* 231, 244–251. doi: 10.1016/j.psychres.2014.12.005
- Baraud, I., Buytet, B., Bec, P., and Blois-Heulin, C. (2009). Social laterality and 'transversality' in two species of mangabeys: influence of rank and implication for hemispheric specialization. *Behav. Brain Res.* 198, 449–458. doi: 10.1016/j.bbr.2008.11.032
- Barbato, D. M., and Kinouchi, O. (2000). Optimal pruning in neural networks. *Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics* 62, 8387–8394.
- Barton, R. A., and Venditti, C. (2014). Rapid evolution of the cerebellum in humans and other great apes. *Curr. Biol.* 24, 2440–2444. doi: 10.1016/j.cub.2014.08.056
- Bauer, R. H. (1993). Lateralization of neural control for vocalization by the frog (*Rana pipiens*). *Psychobiology* 21, 243–248. doi: 10.3758/BF03327141
- Benítez-Burraco, A., and Boeckx, C. (2015). Possible functional links among brain- and skull-related genes selected in modern humans. *Front. Psychol.* 6:794. doi: 10.3389/fpsyg.2015.00794
- Benítez-Burraco, A., and Longa, V. M. (2012). Right-handedness, lateralization and language in Neanderthals: a comment on Frayer et al. (2010). *J. Anthropol. Sci.* 90, 187–192. doi: 10.4436/jass.90002
- Berchicci, M., Tamburro, G., and Comani, S. (2015). The intrahemispheric functional properties of the developing sensorimotor cortex are influenced by maturation. *Front. Hum. Neurosci.* 9:39. doi: 10.3389/fnhum.2015.00039
- Bergman, T. J. (2013). Speech-like vocalized lip-smacking in geladas. *Curr. Biol.* 23, R268–R269. doi: 10.1016/j.cub.2013.02.038
- Bisazza, A., Cantalupo, C., Robins, A., Rogers, L. J., and Vallortigara, G. (1996). Right-pawedness in toads. *Nature* 379, 408. doi: 10.1038/379408a0
- Bisazza, A., Pignatti, R., and Vallortigara, G. (1997). Detour tests reveal task- and stimulus-specific behavioral lateralization in mosquitofish (*Gambusia holbrooki*). *Behav. Brain Res.* 89, 237–242. doi: 10.1016/s0166-4328(97)00061-2
- Boeckx, C., and Benítez-Burraco, A. (2014a). Globularity and language-readiness: generating new predictions by expanding the set of genes of interest. *Front. Psychol.* 5:1324. doi: 10.3389/fpsyg.2014.01324
- Boeckx, C. A., and Benítez-Burraco, A. (2014b). The shape of the human language-ready brain. *Front. Psychol.* 5:282. doi: 10.3389/fpsyg.2014.00282
- Boeckx, C., and Theofanopoulou, C. (2014). A multidimensional interdisciplinary framework for linguistics: the lexicon as a case study. *J. Cogn. Sci. (Seoul)* 15, 403–420.
- Boeckx, C., and Theofanopoulou, C. (2015). Commentary on: labels, cognomes, and cyclic computation: an ethological perspective. *Front. Psychol.* 6:784. doi: 10.3389/fpsyg.2015.00784
- Bogart, S. L., Mangin, J.-F., Schapiro, S. J., Reamer, L., Bennett, A. J., Pierre, P. J., et al. (2012). Cortical sulci asymmetries in chimpanzees and macaques: a new look at an old idea. *Neuroimage* 61, 533–541. doi: 10.1016/j.neuroimage.2012.03.082
- Bohsali, A. A., Triplett, W., Sudhyadhom, A., Gullett, J. M., McGregor, K., Fitzgerald, D. B., et al. (2015). Broca's area – thalamic connectivity. *Brain Lang.* 141, 80–88. doi: 10.1016/j.bandl.2014.12.001
- Böye, M., Güntürkün, O., and Vaclair, J. (2005). Right ear advantage for conspecific calls in adults and subadults, but not infants, California sea lions (*Zalophus californianus*): hemispheric specialization for communication? *Eur. J. Neurosci.* 21, 1727–1732. doi: 10.1111/j.1460-9568.2005.04005.x
- Bruner, E. (2010). Morphological differences in the parietal lobes within the human genus: a neurofunctional perspective. *Curr. Anthropol.* 51, S77–S88. doi: 10.1086/650729
- Bruner, E. (2014). "Functional craniology, human evolution, and anatomical constraints in the neanderthal braincase," in *Dynamics of Learning in Neanderthals and Modern Humans*, Vol. 2, eds T. Akazawa, N. Ogiwara, H. C. Tanabe, and H. Terashima (Dordrecht: Springer), 121–129.
- Bruner, E., de la Cuétara, J. M., Colom, R., and Martin-Loeches, M. (2012). Gender-based differences in the shape of the human corpus callosum are associated with allometric variations. *J. Anat.* 220, 417–421. doi: 10.1111/j.1469-7580.2012.01476.x
- Bruner, E., Rangel de Lázaro, G., de la Cuétara, J. M., Martín-Loeches, M., Colom, R., and Jacobs, H. I. L. (2014). Midsagittal brain variation and MRI shape analysis of the precuneus in adult individuals. *J. Anat.* 224, 367–376. doi: 10.1111/joa.12155
- Buckner, R. L., and Krienen, F. M. (2013). The evolution of distributed association networks in the human brain. *Trends Cogn. Sci. (Regul. Ed.)* 17, 648–665. doi: 10.1016/j.tics.2013.09.017
- Buzsáki, G. (2006). *Rhythms of the Brain*. Oxford: Oxford University Press.
- Buzsáki, G., Logothetis, N., and Singer, W. (2013). Scaling brain size, keeping timing: evolutionary preservation of brain rhythms. *Neuron* 80, 751–764. doi: 10.1016/j.neuron.2013.10.002
- Cantalupo, C., Bisazza, A., and Vallortigara, G. (1995). Lateralization of predator-evasion response in a teleost fish (*Girardinus falcatus*). *Neuropsychologia* 33, 1637–1646. doi: 10.1016/0028-3932(95)00043-7
- Caspari, J. M., and Dunbar, R. I. M. (1996). Asymmetries in the visual processing of emotional cues during agonistic interactions by gelada baboons. *Behav. Process.* 37, 57–65. doi: 10.1016/0376-6357(95)00075-5
- Catmur, C., Walsh, V., and Heyes, C. (2007). Sensorimotor learning configures the human mirror system. *Curr. Biol.* 17, 1527–1531. doi: 10.1016/j.cub.2007.08.006
- Cavanna, A. E., and Trimble, M. R. (2006). The precuneus: a review of its functional anatomy and behavioural correlates. *Brain* 129, 564–583. doi: 10.1093/brain/awl004
- Celesia, G. G., and Hickok, G. (2015). "Auditory neglect and related disorders," in *The Human Auditory System: Fundamental Organization and Clinical Disorders*, eds G. G. Celesia and G. S. Hickok (New York, NY: Elsevier).
- Changizi, M. A. (2001). Principles underlying mammalian cortical scaling. *Biol. Cybern.* 84, 207–215. doi: 10.1007/s004220000205
- Chklovskii, D. B., Mel, B. W., and Svoboda, K. (2004). Cortical rewiring and information storage. *Nature* 431, 782–788. doi: 10.1038/nature03012
- Cook, R., Bird, G., Catmur, C., Press, C., and Heyes, C. (2014). Mirror neurons: from origin to function. *Behav. Brain Sci.* 37, 177–192. doi: 10.1017/s0140525x13000903
- Corballis, M. (1998). Interhemispheric neural summation in the absence of the corpus callosum. *Brain* 121, 1795–1807. doi: 10.1093/brain/121.9.1795
- Courchesne, E., Press, G. A., and Yeung-Courchesne, R. (1993). Parietal lobe abnormalities detected with MR in patients with infantile autism. *Am. J. Roentgenol.* 160, 387–393. doi: 10.2214/ajr.160.2.8424359
- Courchesne, E., Townsend, J., and Saitoh, O. (1994). The brain in infantile autism: posterior fossa structures are abnormal. *Neurology* 44, 214–223. doi: 10.1212/wnl.44.2.214
- Cuellar, M., Bowers, A., Harkrider, A. W., Wilson, M., and Saltuklaroglu, T. (2012). Mu suppression as an index of sensorimotor contributions to speech processing: evidence from continuous EEG signals. *Int. J. Psychophysiol.* 85, 242–248. doi: 10.1016/j.ijpsycho.2012.04.003
- Cuevas, K., Cannon, E. N., Yoo, K., and Fox, N. A. (2014). The infant EEG mu rhythm: methodological considerations and best practices. *Dev. Rev.* 34, 26–43. doi: 10.1016/j.dr.2013.12.001
- Damarla, S. R., Keller, T. A., Kana, R. K., Cherkassky, V. L., Williams, D. L., Minshew, N. J., et al. (2010). Cortical underconnectivity coupled with preserved visuospatial cognition in autism: evidence from an fMRI study of an embedded figures task. *Autism Res.* 3, 273–279. doi: 10.1002/aur.153
- De Schotten, M. T., Dell'Acqua, F., Forkel, S. J., Simmons, A., Vergani, F., Murphy, D. G., et al. (2011). A lateralized brain network for visuospatial attention. *Nat. Neurosci.* 14, 1245–1246. doi: 10.1038/nn.2905
- Doelling, K. B., Arnal, L. H., Ghitza, O., and Poeppel, D. (2014). Acoustic landmarks drive delta-theta oscillations to enable speech comprehension by facilitating perceptual parsing. *Neuroimage* 85, 761–768. doi: 10.1016/j.neuroimage.2013.06.035

- Doron, K., and Gazzaniga, M. (2008). Neuroimaging techniques offer new perspectives on callosal transfer and interhemispheric communication. *Cortex* 44, 1023–1029. doi: 10.1016/j.cortex.2008.03.007
- Ehret, G. (1987). Left hemisphere advantage in the mouse brain for recognizing ultrasonic communication calls. *Nature* 325, 249–251. doi: 10.1038/325249a0
- Engbretson, G. A., Reiner, A., and Brecha, N. (1981). Habenular asymmetry and the central connections of the parietal eye of the lizard. *J. Comp. Neurol.* 198, 155–165. doi: 10.1002/cne.901980113
- Engelborghs, S., Mariën, P., Martin, J. J., and De Deyn, P. P. (1998). Functional anatomy, vascularisation and pathology of the human thalamus. *Acta Neurol. Belg.* 98, 252–265.
- Feinberg, I., and Campbell, I. G. (2010). Sleep EEG changes during adolescence: an index of a fundamental brain reorganization. *Brain Cogn.* 72, 56–65. doi: 10.1016/j.bandc.2009.09.008
- Ferrari, P. F., Gallese, V., Rizzolatti, G., and Fogassi, L. (2003). Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *Eur. J. Neurosci.* 17, 1703–1714. doi: 10.1046/j.1460-9568.2003.02601.x
- Ferrari, P. F., Paukner, A., Ruggiero, A., Darcey, L., Unbehagen, S., and Suomi, S. J. (2009). Interindividual differences in neonatal imitation and the development of action chains in rhesus macaques. *Child Dev.* 80, 1057–1068. doi: 10.1111/j.1467-8624.2009.01316.x
- Fields, R. D. (2005). Myelination: an overlooked mechanism of synaptic plasticity? *Neuroscientist* 11, 528–531. doi: 10.1177/1073858405282304
- Fields, R. D. (2008). White matter matters. *Sci. Am.* 298, 54–61. doi: 10.1038/scientificamerican0308-54
- Fitch, W. T., and Braccini, S. N. (2013). Primate laterality and the biology and evolution of human handedness: a review and synthesis. *Ann. N. Y. Acad. Sci.* 1288, 70–85. doi: 10.1111/nyas.12071
- Ford, A., Triplett, W., Sudhavadhom, A., Gullett, J., McGregor, K., FitzGerald, D. B., et al. (2013). Broca's area and its striatal and thalamic connections: a diffusion-MRI tractography study. *Front. Neuroanat.* 7:8. doi: 10.3389/fnana.2013.00008
- Freitag, M. T., van Bruggen, T., Fritzsche, K. H., Henze, R., Brunner, R., Parzer, P., et al. (2013). Reduced lateralization in early onset schizophrenia. *Neurosci. Lett.* 537, 23–28. doi: 10.1016/j.neulet.2013.01.024
- Frenkel-Toledo, S., Bentin, S., Perry, A., Lieberman, D. G., and Soroker, N. (2013). Dynamics of the EEG power in the frequency and spatial domains during observation and execution of manual movements. *Brain Res.* 1509, 43–57. doi: 10.1016/j.brainres.2013.03.004
- Friederici, A. D., von Cramon, D. Y., and Kotz, S. A. (2007). Role of the corpus callosum in speech comprehension: interfacing syntax and prosody. *Neuron* 53, 135–145. doi: 10.1016/j.neuron.2006.11.020
- Ghazanfar, A. A., and Poeppel, D. (2014). "The neurophysiology and evolution of the speech rhythm," in *The Cognitive Neurosciences V*, ed. M. S. Gazzaniga (Cambridge, MA: MIT Press), 629–638.
- Ghazanfar, A. A., Takahashi, D. Y., Mathur, N., and Fitch, W. T. (2012). Cineradiography of monkey lip-smacking reveals putative precursors of speech dynamics. *Curr. Biol.* 22, 1176–1182. doi: 10.1016/j.cub.2012.04.055
- Giljov, A., Karenina, K., Ingram, J., and Malashichev, Y. (2015). Parallel emergence of true handedness in the evolution of marsupials and placentals. *Curr. Biol.* 25, 1878–1884. doi: 10.1016/j.cub.2015.05.043
- Gunz, P., Neubauer, S., Golovanova, L., Doronichev, V., Maureille, B., and Hublin, J.-J. (2012). A uniquely modern human pattern of endocranial development. Insights from a new cranial reconstruction of the Neandertal newborn from Mezmaiskaya. *J. Hum. Evol.* 62, 300–313. doi: 10.1016/j.jhevol.2011.11.013
- Gunz, P., Neubauer, S., Maureille, B., and Hublin, J.-J. (2010). Brain development after birth differs between Neanderthals and modern humans. *Curr. Biol.* 20, R921–R922. doi: 10.1016/j.cub.2010.10.018
- Hänggi, J., Fövényi, L., Liem, F., Meyer, M., and Jäncke, L. (2014). The hypothesis of neuronal interconnectivity as a function of brain size—a general organization principle of the human connectome. *Front. Hum. Neurosci.* 8:915. doi: 10.3389/fnhum.2014.00915
- Harris, L. J. (1991). Cerebral control for speech in right-handers and left-handers: an analysis of the views of Paul Broca, his contemporaries, and his successors. *Brain Lang.* 40, 1–50. doi: 10.1016/0093-934X(91)90115-H
- Hauser, M. D., and Andersson, K. (1994). Left hemisphere dominance for processing vocalizations in adult, but not infant, rhesus monkeys: field experiments. *Proc. Natl. Acad. Sci. U.S.A.* 91, 3946–3948. doi: 10.1073/pnas.91.9.3946
- Hickok, G., and Hauser, M. (2010). (Mis)understanding mirror neurons. *Curr. Biol.* 20, R593–R594. doi: 10.1016/j.cub.2010.05.047
- Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8, 393–402. doi: 10.1038/nrn2113
- Hickok, G., and Poeppel, D. (2015). "Neural basis of speech perception," in *The Human Auditory System: Fundamental Organization and Clinical Disorders*, eds G. G. Celesia and G. S. Hickok (New York, NY: Elsevier), 149–160.
- Holleran, L., Ahmed, M., Anderson-Schmidt, H., McFarland, J., Emsell, L., Leemans, A., et al. (2014). Altered interhemispheric and temporal lobe white matter microstructural organization in severe chronic schizophrenia. *Neuropsychopharmacology* 39, 944–954. doi: 10.1038/npp.2013.294
- Holliday, M. A. (1986). "Body composition and energy needs during growth," in *Human Growth: A Comprehensive Treatise*, eds F. Falkner and J. M. Tanner (New York, NY: Plenum), 101–107.
- Hook-Costigan, M. A., and Rogers, L. J. (1998). Lateralized use of the mouth in production of vocalizations by marmosets. *Neuropsychologia* 36, 1265–1273. doi: 10.1016/S0028-3932(98)00037-2
- Hori, M. (1993). Frequency-dependent natural selection in the handedness of scale-eating cichlid fish. *Science* 260, 216–219. doi: 10.1126/science.260.5105.216
- Hoso, M., Asami, T., and Hori, M. (2007). Right-handed snakes: convergent evolution of asymmetry for functional specialization. *Biol. Lett.* 3, 169–172. doi: 10.1098/rsbl.2006.0600
- Hublin, J. J., Neubauer, S., and Gunz, P. (2015). Brain ontogeny and life history in Pleistocene hominins. *Philos. Trans. R. Soc. B Biol. Sci.* 370: 20140062. doi: 10.1098/rstb.2014.0062
- Hynd, G. W., Hall, J., Novey, E. S., Eliopoulos, D., Black, K., Gonzalez, J. J., et al. (1995). Dyslexia and corpus callosum morphology. *Arch. Neurol.* 52, 32–38. doi: 10.1001/archneur.1995.00540250036010
- Iacoboni, M., and Dapretto, M. (2006). The mirror neuron system and the consequences of its dysfunction. *Nat. Rev. Neurosci.* 7, 942–951. doi: 10.1038/nrn2024
- Innocenti, G. M. (1995). Exuberant development of connections, and its possible permissive role in cortical evolution. *Trends Neurosci.* 18, 397–402. doi: 10.1016/0166-2236(95)93936-r
- Innocenti, G. M., and Price, D. J. (2005). Exuberance in the development of cortical networks. *Nat. Rev. Neurosci.* 6, 955–965. doi: 10.1038/nrn1790
- Jann, K., Federspiel, A., Giezendanner, S., Andreotti, J., Kottlow, M., Dierks, T., et al. (2012). Linking brain connectivity across different time scales with electroencephalogram, functional magnetic resonance imaging, and diffusion tensor imaging. *Brain Connect.* 2, 11–20. doi: 10.1089/brain.2011.0063
- Jeffery, N. (2002). Differential regional brain growth and rotation of the prenatal human tentorium cerebelli. *J. Anat.* 200, 135–144. doi: 10.1046/j.0021-8782.2001.00017.x
- Jensen, O., Gelfand, J., Kounios, J., and Lisman, J. E. (2002). Oscillations in the alpha band (9–12 Hz) increase with memory load during retention in a short-term memory task. *Cereb. Cortex* 12, 877–882. doi: 10.1093/cercor/12.8.877
- Jochaut, D., Lehongre, K., Saitovitch, A., Devauchelle, A.-D., Olsagasti, I., Chabane, N., et al. (2015). Atypical coordination of cortical oscillations in response to speech in autism. *Front. Hum. Neurosci.* 9:171. doi: 10.3389/fnhum.2015.00171
- Jones, E. G., Coulter, J. D., and Hendry, S. H. (1978). Intracortical connectivity of architectonic fields in the somatic sensory, motor and parietal cortex of monkeys. *J. Comp. Neurol.* 181, 291–347. doi: 10.1002/cne.901810206
- Just, M. A., Cherkassky, V. L., Keller, T. A., Kana, R. K., and Minshew, N. J. (2007). Functional and anatomical cortical underconnectivity in autism: evidence from an fMRI study of an executive function task and corpus callosum morphometry. *Cereb. Cortex* 17, 951–961. doi: 10.1093/cercor/bhl006
- Karbowsky, J. (2003). How does connectivity between cortical areas depend on brain size? Implications for efficient computation. *J. Comput. Neurosci.* 15, 347–356. doi: 10.1023/a:1027467911225
- Killackey, H. P., Gould, H. J., Cusick, C. G., Pons, T. P., and Kaas, J. H. (1983). The relation of corpus callosum connections to architectonic fields and body surface maps in sensorimotor cortex of new and old world monkeys. *J. Comp. Neurol.* 219, 384–419. doi: 10.1002/cne.902190403



- Kirihara, K., Rissling, A. J., Swerdlow, N. R., Braff, D. L., and Light, G. A. (2012). Hierarchical organization of gamma and theta oscillatory dynamics in schizophrenia. *Biol. Psychiatry* 71, 873–880. doi: 10.1016/j.biopsych.2012.01.016
- Kirschner, M. P., Chen, S. H. A., and Desmond, J. E. (2010). Modality specific cerebro-cerebellar activations in verbal working memory: an fMRI study. *Behav. Neurosci.* 23, 51–63. doi: 10.1155/2010/587450
- Klimesch, W. (2012).  $\alpha$ -band oscillations, attention, and controlled access to stored information. *Trends Cogn. Sci.* 16, 606–617. doi: 10.1016/j.tics.2012.10.007
- Klin, A., Jones, W., Schultz, R., Volkmar, F., and Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Arch. Gen. Psychiatry* 59, 809–816. doi: 10.1001/archpsyc.59.9.809
- Knecht, S., Dräger, B., Deppe, M., Bobe, L., Lohmann, H., Flöel, A., et al. (2000). Handedness and hemispheric language dominance in healthy humans. *Brain* 123, 2512–2518. doi: 10.1093/brain/123.12.2512
- Knöchel, C., Oertel-Knöchel, V., Schönmeier, R., Rotarska-Jagiela, A., van de Ven, V., Prvulovic, D., et al. (2012). Interhemispheric hypoconnectivity in schizophrenia: fiber integrity and volume differences of the corpus callosum in patients and unaffected relatives. *Neuroimage* 59, 926–934. doi: 10.1016/j.neuroimage.2011.07.088
- Kuzawa, C. (1998). Adipose tissue in human infancy and childhood: an evolutionary perspective. *Am. J. Phys. Anthropol.* 107, 177–209. doi: 10.1002/(SICI)1096-8644(1998)107:27<177::AID-AJPA7>3.0.CO;2-B
- Kyriakopoulos, M., Bargiotas, T., Barker, G. J., and Frangou, S. (2008). Diffusion tensor imaging in schizophrenia. *Eur. Psychiatry* 23, 255–273. doi: 10.1016/j.eurpsy.2007.12.004
- Lange, J., Pavlidou, A., and Schnitzler, A. (2015). Lateralized modulation of beta-band power in sensorimotor areas during action observation. *Front. Integr. Neurosci.* 9:43. doi: 10.3389/fnint.2015.00043
- Leiner, H., Leiner, A., and Dow, R. (1989). Reappraising the cerebellum: what does the hindbrain contribute to the forebrain? *Behav. Neurosci.* 103, 998–1008. doi: 10.1037/0735-7044.103.5.998
- Lenneberg, E. H. (1966). “Speech development: its anatomical and physiological concomitants,” in *Brain Function: Speech, Language and Communication*, Vol. 3, ed. E. C. Carterette (Berkeley, CA: University of California Press), 37–66.
- Leroux, E., Delcroix, N., and Dollfus, S. (2015). Left-hemisphere lateralization for language and interhemispheric fiber tracking in patients with schizophrenia. *Schizophr. Res.* 165, 30–37. doi: 10.1016/j.schres.2015.03.028
- Light, G. A., Hsu, J. L., Hsieh, M. H., Meyer-Gomes, K., Sprock, J., Swerdlow, N. R., et al. (2006). Gamma band oscillations reveal neural network cortical coherence dysfunction in schizophrenia patients. *Biol. Psychiatry* 60, 1231–1240. doi: 10.1016/j.biopsych.2006.03.055
- Lindell, A. K. (2006). In your right mind: right hemisphere contributions to language processing and production. *Neuropsychol. Rev.* 16, 131–148. doi: 10.1007/s11065-006-9011-9
- Luc, H., and Arnal, K. B. D. (2014). Delta-beta coupled oscillations underlie temporal prediction accuracy. *Cereb. Cortex* 25, 3077–3085. doi: 10.1093/cercor/bhr325
- Luis, E. O., Arrondo, G., Vidorreta, M., Martínez, M., Loayza, F., Fernández-Seara, M. A., et al. (2015). Successful working memory processes and cerebellum in an elderly sample: a neuropsychological and fMRI study. *PLoS ONE* 10:e0131536. doi: 10.1371/journal.pone.0131536
- Marshall, P. J., and Meltzoff, A. N. (2011). Neural mirroring systems: exploring the EEG Mu rhythm in human infancy. *Dev. Cogn. Neurosci.* 1, 110–123. doi: 10.1016/j.dcn.2010.09.001
- Mehlhorn, J., Haastert, B., and Rehkämper, G. (2010). Asymmetry of different brain structures in homing pigeons with and without navigational experience. *J. Exp. Biol.* 213, 2219–2224. doi: 10.1242/jeb.043208
- Miller, E. M. (1994). Intelligence and brain myelination: a hypothesis. *Pers. Individ. Dif.* 17, 803–832. doi: 10.1016/0191-8869(94)90049-3
- Minzenberg, M. J., Firl, A. J., Yoon, J. H., Gomes, G. C., Reinking, C., and Carter, C. S. (2010). Gamma oscillatory power is impaired during cognitive control independent of medication status in first-episode schizophrenia. *Neuropsychopharmacology* 35, 2590–2599. doi: 10.1038/npp.2010.150
- Miyata, J., Yamada, M., Namiki, C., Hirao, K., Saze, T., Fujiwara, H., et al. (2010). Reduced white matter integrity as a neural correlate of social cognition deficits in schizophrenia. *Schizophr. Res.* 119, 232–239. doi: 10.1016/j.schres.2009.12.038
- Molnar-Szakacs, I., and Overy, K. (2006). Music and mirror neurons: from motion to ‘emotion’. *Soc. Cogn. Affect. Neurosci.* 1, 235–241. doi: 10.1093/scan/nsl029
- Moorman, S., Gobes, S. M. H., Kuijpers, M., Kerkhofs, A., Zandbergen, M. A., and Bolhuis, J. J. (2012). Human-like brain hemispheric dominance in birdsong learning. *Proc. Natl. Acad. Sci. U.S.A.* 109, 12782–12787. doi: 10.1073/pnas.1207207109
- Moran, L. V., and Hong, L. E. (2011). High vs low frequency neural oscillations in schizophrenia. *Schizophr. Bull.* 37, 659–663. doi: 10.1093/schbul/sbr056
- Morillon, B., Lehongre, K., Frackowiak, R. S. J., Ducors, A., Kleinschmidt, A., Poeppel, D., et al. (2010). Neurophysiological origin of human brain asymmetry for speech and language. *Proc. Natl. Acad. Sci. U.S.A.* 107, 18688–18693. doi: 10.1073/pnas.1007189107
- Mota, B., and Herculano-Houzel, S. (2012). How the cortex gets its folds: an inside-out, connectivity-driven model for the scaling of mammalian cortical folding. *Front. Neuroanat.* 6:3. doi: 10.3389/fnana.2012.00003
- Murphy, J. W., Foxe, J. J., Peters, J. B., and Molholm, S. (2014). Susceptibility to distraction in autism spectrum disorder: probing the integrity of oscillatory alpha-band suppression mechanisms. *Autism Res.* 7, 442–458. doi: 10.1002/aur.1374
- Muthukumaraswamy, S. D., and Johnson, B. W. (2004). Changes in rolandic mu rhythm during observation of a precision grip. *Psychophysiology* 41, 152–156. doi: 10.1046/j.1469-8986.2003.00129.x
- Neubert, F.-X., Mars, R. B., Thomas, A. G., Sallet, J., and Rushworth, M. F. S. (2014). Comparison of human ventral frontal cortex areas for cognitive control and language with areas in monkey frontal cortex. *Neuron* 81, 700–713. doi: 10.1016/j.neuron.2013.11.012
- Nosarti, C., Rushe, T. M., Woodruff, P. W. R., Stewart, A. L., Rifkin, L., and Murray, R. M. (2004). Corpus callosum size and very preterm birth: relationship to neuropsychological outcome. *Brain* 127, 2080–2089. doi: 10.1093/brain/awh230
- Nottebohm, F. (1971). Neural lateralization of vocal control in a passerine bird. I. Song. *J. Exp. Zool.* 177, 229–261. doi: 10.1002/jez.1401770210
- Nudo, R. J., Jenkins, W. M., Merzenich, M. M., Prejean, T., and Grenda, R. (1992). Neurophysiological correlates of hand preference in primary motor cortex of adult squirrel monkeys. *J. Neurosci.* 12, 2918–2947.
- Obleser, J., and Weisz, N. (2012). Suppressed alpha oscillations predict intelligibility of speech and its acoustic details. *Cereb. Cortex* 22, 2466–2477. doi: 10.1093/cercor/bhr325
- Obleser, J., Wöstmann, M., Hellbernd, N., Wilsch, A., and Maess, B. (2012). Adverse listening conditions and memory load drive a common alpha oscillatory network. *J. Neurosci.* 32, 12376–12383. doi: 10.1523/jneurosci.4908-11.2012
- Ocklenburg, S., and Güntürkün, O. (2012). Hemispheric asymmetries: the comparative view. *Front. Psychol.* 3:5. doi: 10.3389/fpsyg.2012.00005
- Okanoya, K., Ikebuchi, M., Uno, H., and Watanabe, S. (2001). Left-side dominance for song discrimination in Bengalese finches (*Lonchura striata* var. *domestica*). *Anim. Cogn.* 4, 241–245. doi: 10.1007/s10071-001-0120-9
- Önal-Hartmann, C., Pauli, P., Ocklenburg, S., and Güntürkün, O. (2011). The motor side of emotions: investigating the relationship between hemispheres, motor reactions and emotional stimuli. *Psychol. Res.* 76, 311–316. doi: 10.1007/s00426-011-0337-4
- Overath, T., McDermott, J. H., Zarate, J. M., and Poeppel, D. (2015). The cortical analysis of speech-specific temporal structure revealed by responses to sound quilts. *Nat. Neurosci.* 18, 903–911. doi: 10.1038/nn.4021
- Parr, L. A., and Hopkins, W. D. (2000). Brain temperature asymmetries and emotional perception in chimpanzees, Pan troglodytes. *Physiol. Behav.* 71, 363–371. doi: 10.1016/S0031-9384(00)00349-8
- Paul, L. K., Brown, W. S., Adolphs, R., Tyszka, J. M., Richards, L. J., Mukherjee, P., et al. (2007). Agensis of the corpus callosum: genetic, developmental and functional aspects of connectivity. *Nat. Rev. Neurosci.* 8, 287–299. doi: 10.1038/nrn2107
- Perry, A., and Bentin, S. (2009). Mirror activity in the human brain while observing hand movements: a comparison between EEG desynchronization in the mu-range and previous fMRI results. *Brain Res.* 1282, 126–132. doi: 10.1016/j.brainres.2009.05.059
- Perry, A., Bentin, S., Shalev, I., Israel, S., Uzevovsky, F., Bar-On, D., et al. (2010). Intranasal oxytocin modulates EEG mu/alpha and beta rhythms during perception of biological motion. *Psychoneuroendocrinology* 35, 1446–1453. doi: 10.1016/j.psyneuen.2010.04.011



- Peters, B. D., and Karlsgodt, K. H. (2015). White matter development in the early stages of psychosis. *Schizophr. Res.* 161, 61–69. doi: 10.1016/j.schres.2014.05.021
- Pineda, J. A. (2005). The functional significance of mu rhythms: translating “seeing” and “hearing” into “doing.” *Brain Res. Brain Res. Rev.* 50, 57–68. doi: 10.1016/j.brainresrev.2005.04.005
- Pineda, J. A. (2008). Sensorimotor cortex as a critical component of an ‘extended’ mirror neuron system: does it solve the development, correspondence, and control problems in mirroring? *Behav. Brain Funct.* 4, 47. doi: 10.1186/1744-9081-4-47
- Poeppel, D. (2003). The analysis of speech in different temporal integration windows: cerebral lateralization as ‘asymmetric sampling in time’. *Speech Commun.* 41, 245–255. doi: 10.1016/s0167-6393(02)00107-3
- Pollmann, S., Maertens, M., von Cramon, D. Y., Lepsien, J., and Hughdahl, K. (2002). Dichotic listening in patients with posterior CC and nonspatial callosal lesions. *Neuropsychologia* 16, 56–64. doi: 10.1037/0894-4105.16.1.56
- Ringo, J. L., Doty, R. W., Demeter, S., and Simard, P. Y. (1994). Time is of the essence: a conjecture that hemispheric specialization arises from interhemispheric conduction delay. *Cereb. Cortex* 4, 331–343. doi: 10.1093/cercor/4.4.331
- Rizzolatti, G., and Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.* 27, 169–192. doi: 10.1146/annurev.neuro.27.070203.144230
- Rizzolatti, G., and Luppino, G. (2001). The cortical motor system. *Neuron* 31, 889–901. doi: 10.1016/s0896-6273(01)00423-8
- Rogers, L. J., Ward, J. P., and Stafford, D. (1994). Eye dominance in the small-eared bushbaby, *Otolemur garnettii*. *Neuropsychologia* 32, 257–264. doi: 10.1016/0028-3932(94)90011-6
- Rogers, L. J., and Workman, L. (1993). Footedness in birds. *Anim. Behav.* 45, 409–411.
- Rumsey, J. M., Casanova, M., Mannheim, G. B., Patronas, N., DeVaughn, N., Hamburger, S. D., et al. (1996). Corpus callosum morphology, as measured with MRI, in dyslexic men. *Biol. Psychiatry* 39, 769–775. doi: 10.1016/0006-3223(95)00225-1
- Sakai, T., Mikami, A., Tomonaga, M., Matsui, M., Suzuki, J., Hamada, Y., et al. (2011). Differential prefrontal white matter development in chimpanzees and humans. *Curr. Biol.* 21, 1397–1402. doi: 10.1016/j.cub.2011.07.019
- Sammler, D., Kotz, S. A., Eckstein, K., Ott, D. V. M., and Friederici, A. D. (2010). Prosody meets syntax: the role of the corpus callosum. *Brain* 133, 2643–2655. doi: 10.1093/brain/awq231
- Schmahmann, J. (1997). *The Cerebellum and Cognition*. New York, NY: Academic Press.
- Schoenemann, P. T., Sheehan, M. J., and Grotzer, L. D. (2005). Prefrontal white matter volume is disproportionately larger in humans than in other primates. *Nat. Neurosci.* 8, 242–252. doi: 10.1038/nn1394
- Schwarzkopf, S., Schilbach, L., Vogeley, K., and Timmermans, B. (2014). ‘Making it explicit’ makes a difference: evidence for a dissociation of spontaneous and intentional level 1 perspective taking in high-functioning autism. *Cognition* 131, 345–354. doi: 10.1016/j.cognition.2014.02.003
- Senkfor, A. J. (2002). “Episodic action memory: characterization of time course and neural circuitry,” in *Mirror Neurons and the Evolution of Brain and Language*, eds M. I. Stamenov and V. Gallese (Pennsylvania, PA: John Benjamins), 42–87.
- Siniscalchi, M., Quaranta, A., and Rogers, L. J. (2008). Hemispheric specialization in dogs for processing different acoustic stimuli. *PLoS ONE* 3:e3349. doi: 10.1371/journal.pone.0003349
- Skoyles, J. (2012). Human neuromaturation, juvenile extreme energy liability, and adult cognition/cooperation. *Nature Precedings*. Available at: <http://hdl.handle.net/10101/npre.2012.7096.1>
- Smaers, J. B., Steele, J., Case, C. R., and Amunts, K. (2013). The evolution of human handedness. *Ann. N. Y. Acad. Sci.* 1288, 59–69. doi: 10.1159/000323671
- Sporns, O., and Kötter, R. (2004). Motifs in brain networks. *PLoS Biol.* 2:e369. doi: 10.1371/journal.pbio.0020369
- Sporns, O., and Zwi, J. D. (2004). The small world of the cerebral cortex. *Neuroinformatics* 2, 145–162. doi: 10.1385/ni.2:2:145
- Stoodley, C. J., Valera, E. M., and Schmahmann, J. D. (2012). Functional topography of the cerebellum for motor and cognitive tasks: An fMRI study. *Neuroimage* 59, 1560–1570. doi: 10.1016/j.neuroimage.2011.08.065
- Sugishita, M., Otomo, K., Yamazaki, K., Shimizu, H., Yoshioka, M., and Shinohara, A. (1995). Dichotic listening in patients with partial section of the corpus callosum. *Brain* 118, 417–427. doi: 10.1093/brain/118.2.417
- Sun, Z. Y., Klöppel, S., Rivière, D., Perrot, M., Frackowiak, R., Siebner, H., et al. (2012). The effect of handedness on the shape of the central sulcus. *Neuroimage* 60, 332–339. doi: 10.1016/j.neuroimage.2011.12.050
- Taglialetta, J. P., Russell, J. L., Schaeffer, J. A., and Hopkins, W. D. (2008). Communicative signaling activates ‘Broca’s’ homolog in chimpanzees. *Curr. Biol.* 18, 343–348. doi: 10.1016/j.cub.2008.01.049
- Theofanopoulou, C., and Boeckx, C. (2015). “The role of the thalamus in linguistic cognition,” in *Advances in Biolinguistics*, eds C. Boeckx and K. Fujita (New York, NY: Routledge).
- Uhlhaas, P. J., and Singer, W. (2010a). Abnormal neural oscillations and synchrony in schizophrenia. *Nat. Rev. Neurosci.* 11, 100–113. doi: 10.1038/nrn2774
- Uhlhaas, P. J., and Singer, W. J. (2010b). “Developmental changes in neuronal oscillations and synchrony: evidence for a late critical period,” in *Proceedings of the Working Group on Human Neuroplasticity and Education*, Vol. 117, eds A. M. Battro, S. Dehaene, and W. J. Singer (Vatican City: The Pontifical Academy of Sciences), 218–229.
- Ulloa, E., and Pineda, J. (2007). Recognition of point-light biological motion: mu rhythms and mirror neuron activity. *Behav. Brain Res.* 183, 188–194. doi: 10.1016/j.bbr.2007.06.007
- Vallortigara, G., and Andrew, R. J. (1994). Differential involvement of right and left hemisphere in individual recognition in the domestic chick. *Behav. Process.* 33, 41–57. doi: 10.1016/0376-6357(94)90059-0
- Vanderwert, R. E., Fox, N. A., and Ferrari, P. F. (2013). The mirror mechanism and Mu rhythm in social development. *Neurosci. Lett.* 540, 15–20. doi: 10.1016/j.neulet.2012.10.006
- Ventura-Antunes, L., Mota, B., and Herculano-Houzel, S. (2013). Different scaling of white matter volume, cortical connectivity, and gyrification across rodent and primate brains. *Front. Neuroanat.* 7:3. doi: 10.3389/fnana.2013.00003
- Vermeire, B. A., and Hamilton, C. R. (1998). Effects of facial identity, facial expression, and subject’s sex on laterality in monkeys. *Laterality* 3, 1–19. doi: 10.1080/713754290
- Voss, H. U., Tabelow, K., Polzehl, J., Tchernichovski, O., Maul, K. K., Salgado-Commissariat, D., et al. (2007). Functional MRI of the zebra finch brain during song stimulation suggests a lateralized response topography. *Proc. Natl. Acad. Sci. U.S.A.* 104, 10667–10672. doi: 10.1073/pnas.0611515104
- Wang, C. Z., Herbst, J. A., Keller, G. B., and Hahnloser, R. H. (2008). Rapid interhemispheric switching during vocal production in a songbird. *PLoS Biol.* 6:e250. doi: 10.1371/journal.pbio.0060250
- Washington, S. D., and Tillinghast, J. S. (2015). Conjugating time and frequency: hemispheric specialization, acoustic uncertainty, and the mustached bat. *Front. Neurosci.* 9:143. doi: 10.3389/fnins.2015.00143
- Zaehle, T., and Herrmann, C. S. (2011). Neural synchrony and white matter variations in the human brain — Relation between evoked gamma frequency and corpus callosum morphology. *Int. J. Psychophysiol.* 79, 49–54. doi: 10.1016/j.ijpsycho.2010.06.029
- Zatorre, R. J., Belin, P., and Penhune, V. B. (2002). Structure and function of auditory cortex: music and speech. *Trends Cogn. Sci. (Regul. Ed.)* 6, 37–46. doi: 10.1016/s1364-6613(00)01816-7
- Zverev, Y. P. (2006). Cultural and environmental pressure against left-hand preference in urban and semi-urban Malawi. *Brain Cogn.* 60, 295–303. doi: 10.1016/j.bandc.2005.07.008

**Conflict of Interest Statement:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Theofanopoulou. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Corrigendum: Brain asymmetry in the white matter making and globularity

Constantina Theofanopoulou \*

Department of General Linguistics, Universitat de Barcelona, Barcelona, Spain

**Keywords:** brain asymmetry, white matter, lateralization, brain rhythms, language

## OPEN ACCESS

**Edited and reviewed by:**

Antonio Benítez-Burraco,  
University of Huelva, Spain

**\*Correspondence:**

Constantina Theofanopoulou  
constantinaki@hotmail.com

**Specialty section:**

This article was submitted to  
Language Sciences,  
a section of the journal  
Frontiers in Psychology

**Received:** 13 November 2015

**Accepted:** 16 November 2015

**Published:** 14 December 2015

**Citation:**

Theofanopoulou C (2015)  
Corrigendum: Brain asymmetry in the  
white matter making and globularity.  
Front. Psychol. 6:1857.  
doi: 10.3389/fpsyg.2015.01857

## A corrigendum on

### Brain asymmetry in the white matter making and globularity

by Theofanopoulou, C. (2015). *Front. Psychol.* 6:1355. doi: 10.3389/fpsyg.2015.01355

In the Acknowledgments Section of the original Hypothesis & Theory article, the last sentence erroneously reported grant number FFI2014-61888-EXP as part of funds supporting the work. The only grant was #FFI2013-43823-P, provided by Spanish Ministry of Economy and Competitiveness.

This error does not change the scientific conclusions of the article in any way.

## ACKNOWLEDGMENTS

I am indebted to Cedric Boeckx for his constant encouragement and help. Many thanks also to Antonio Benítez-Burraco for helpful discussion.

**Conflict of Interest Statement:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Theofanopoulou. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# The brain dynamics of linguistic computation

Elliot Murphy\*

Division of Psychology and Language Sciences, University College London, London, UK

## OPEN ACCESS

### Edited by:

Antonio Benítez-Burraco,  
University of Huelva, Spain

### Reviewed by:

Aritz Irurtzun,  
Centre de recherche sur la langue et  
les textes basques, France  
Timothy Michael Ellmore,  
The City College of New York, USA  
Constantina Theofanopoulou,  
Universitat de Barcelona, Spain

### \*Correspondence:

Elliot Murphy,  
Division of Psychology and Language  
Sciences, University College London,  
Chandler House, 2 Wakefield Street,  
London WC1N 1PF, UK  
elliottmurphy91@gmail.com

### Specialty section:

This article was submitted to  
Language Sciences,  
a section of the journal  
Frontiers in Psychology

**Received:** 11 July 2015

**Accepted:** 18 September 2015

**Published:** 13 October 2015

### Citation:

Murphy E (2015) The brain dynamics  
of linguistic computation.  
Front. Psychol. 6:1515.  
doi: 10.3389/fpsyg.2015.01515

Neural oscillations at distinct frequencies are increasingly being related to a number of basic and higher cognitive faculties. Oscillations enable the construction of coherently organized neuronal assemblies through establishing transitory temporal correlations. By exploring the elementary operations of the language faculty—labeling, concatenation, cyclic transfer—alongside neural dynamics, a new model of linguistic computation is proposed. It is argued that the universality of language, and the true biological source of Universal Grammar, is not to be found purely in the genome as has long been suggested, but more specifically within the extraordinarily preserved nature of mammalian brain rhythms employed in the computation of linguistic structures. Computational-representational theories are used as a guide in investigating the neurobiological foundations of the human “cognome”—the set of computations performed by the nervous system—and new directions are suggested for how the dynamics of the brain (the “dynome”) operate and execute linguistic operations. The extent to which brain rhythms are the suitable neuronal processes which can capture the computational properties of the human language faculty is considered against a backdrop of existing cartographic research into the localization of linguistic interpretation. Particular focus is placed on labeling, the operation elsewhere argued to be species-specific. A Basic Label model of the human cognome-dynome is proposed, leading to clear, causally-addressable empirical predictions, to be investigated by a suggested research program, Dynamic Cognomics. In addition, a distinction between minimal and maximal degrees of explanation is introduced to differentiate between the depth of analysis provided by cartographic, rhythmic, neurochemical, and other approaches to computation.

**Keywords:** neural oscillations, biolinguistics, syntax, dynome, theta, alpha, beta, gamma

The argument for placing language at the center of investigations into human cognition has by now been pushed on a number of fronts, from palaeoanthropology to philosophy (McGilvray, 2013; Hauser et al., 2014). In contrast, attempts to place the brain at the center of the language sciences have been met with suspicion and even ridicule, typically due to the observation that higher cognitive constructs like *verb* and *phrase* cannot presently be made commensurable with lower-level neurophysiological structures like *dendrite* and *cortical column*. Substantial engagement with the biology literature is a feature still lacking in departments of linguistics, despite the Minimalist Program’s narrowing of the boundaries between the computational and conceptual capacities of humans and non-humans (Chomsky, 1995, 2012, 2015b).

One of the core motivations linguists have for leaving aside biology and keeping to computational investigations arises from Poeppel (2012) and Chomsky's (2000) insightful discussions concerning philosophy of science, theoretical reduction, and unification. These authors point out that, as with the reduction of physics to an unaltered chemistry in the early years of the twentieth century, it may well be that a new neurobiology yielded by a "Galilean" revolution is required for commensurability with the computational theories of syntacticians to be achieved, rather than a revolutionized theory of language. But the common claim that linguistics is biology at a suitable level of abstraction (Berwick, 2011) is also used to effectively get linguists "off the hook" of directly exploring the biology of language, satisfied as many are with concluding that this is purely the job of neuroscience. Yet if neuroscientists are not guided by the concerns of computationalists across the cognitive sciences, and not just linguistics, then there is little reason to believe that this goal will ever be achieved. As Lenneberg (1964, p. 76) noted, "[n]othing is gained by labeling the propensity for language as biological unless we can use this insight for new research directions—unless more specific correlates can be uncovered."

## Dynamic Cognomics: Preliminary Remarks

The central argument of this paper will be that recent developments in brain dynamics and neurochemistry can provide the type of framework needed to meet Poeppel and Embick's (2005) challenge of "granularity" mismatch, or the problem of reconciling the primitives of neuroscience with the primitives of linguistics (see also Fitch, 2009; Poeppel, 2011). The brain simply does not know what *syntax* or *phonology* are, and these concepts are much too coarse to be implemented neurally. In 1996, Poeppel noted of cell assemblies and oscillations that "it is unclear whether these are the right biological categories to account for cognition" (1996, p. 643), but by now the oscillation literature has sufficiently expanded to incorporate numerous cognitive processes.

Linguistics can direct the brain sciences insofar as its insights into the universality of operations like concatenation (set-formation) inform the goals of neurobiology, while the brain sciences can direct linguistics insofar as they place constraints on what possible operations neuronal assemblies and their oscillations can perform. While linguists should focus on making their claims about language biologically feasible, neuroscientists should conversely ensure they do not sideline the notion of computation, as stressed by Gallistel and King (2009).

In order to explore these manifold agendas, I will adopt the multidisciplinary approach promoted by Boeckx and Theofanopoulou (2014), which endorses an interweaving of the sciences concerned with the following topics: the computations performed by the human nervous system (the "cognome"; Poeppel, 2012), brain dynamics (the "dynamome"; Kopell et al., 2014), neural wiring (the "connectome"; Seung, 2012) and genomics. This framework exposes the misleading nature of common questions surrounding whether the brain's wiring "makes us who we are," which have been given an impetus by calls from Seung (2012) and others for a map of the connectome.

The connectome constrains the *kinds* of operations performed by the nervous system, but it cannot reveal *what* operations in particular are performed. What is needed, as Seung himself has explained, is not just a comprehensive model of neural wiring, but also neural computation, which is what a theory of the cognome can contribute (see Reimann et al., 2015 for a proposed algorithm to predict the connectome of neural microcircuits).

Bridging the two domains, I will argue, is the dynamome; or what physicists would term the mesoscale, and not the microscale. The dynamome is the level of brain dynamics, encompassing electrophysiology, and neural oscillations. It explores "not only *what* is connected, but *how* and in what directions regions of the brain are connected" (Kopell et al., 2014, p. 1319). The cartographic literature (e.g., fMRI and DTI studies) typically displays theoretical and empirical satisfaction with discussions of neural "activation," "firing," and "pathways," keeping at a connectomic level of spatiotemporal brain nodes and edges (Bressler and Menon, 2010). The dynamome adds to such a "functional connectome" an understanding of the regions involved in producing and processing brain signals. Although I will focus on brain rhythms, it should be noted that the dynamome extends beyond neural oscillations and includes other temporal structures (Larson-Prior et al., 2013).

I would also like to propose that the universality of language, and the true biological source of Universal Grammar, is not to be found purely in the genome as has long been suggested (where there are surprising layers of variation; Benítez-Burraco and Boeckx, 2014a,b), but more specifically within the extraordinarily preserved nature of mammalian brain rhythms (the oscillations of mice and rats have the same pharmacological profiles as humans) likely arising from the deployment of long-diameter axons of long-range neurons (Buzsáki et al., 2013, see also Calabrese and Woolley, 2015). Such cortical and sub-cortical structures are "among the most sophisticated scalable architectures in nature" (Buzsáki et al., 2013, p. 751), with scalability referring to the ability to perform the same operations with increasing efficiency despite escalating organizational complexity. Brain rhythms, yielded in part by such structures, would therefore be expected to be capable of complex forms of information-transmission and integration.

A central question posed by this paper, then, is "Why claim that neuroscience requires a Galilean revolution in order for it to be made commensurable with linguistics when the properties of syntax may be able to be translated into rhythmic brain processes?" The current paper will suggest a new research program, Dynamic Cognomics, to explore the neurobiology of language in a deeper and more electrophysiologically explicit fashion than many existing cartographic neuroimaging studies, but some important background is needed before any concrete research goals can be drawn up.

## Cartographic Directions

In Murphy (2015a) it was claimed that the ability to label linguistic structures with a categorical identity (e.g., determiner, verb, and adjective), having concatenated two elements into an unordered set, and transfer them in a cyclic fashion to the



conceptual-intentional (CI) interface is the defining property of the human computational system. This perspective will be maintained here. It will be argued that modifications in oscillatory couplings and the cell assemblies targeted by such dynamic operations are a viable candidate for what brought about what could be regarded as a phase transition from single-instance set-formation (of the kind seen in birdsong) to unbounded set-formation. For instance, the phase/non-phase rhythm of syntactic computation ( $[C/T[v/V[D/N]]]$ ), emphasized by Richards (2011), Uriagereka (2012) and Boeckx (2013), may translate well into the rhythmic processes of neural oscillations.

Since the origins of modern cognitive neuroscience, linguistic processes have been claimed to elicit numerous event-related potentials (ERPs) by psycholinguists using magnetoencephalography (MEG) and electroencephalography (EEG) (see Swaab et al., 2012 for a review). As time-frequency analysis and its Fourier transforms developed into a mainstay of “ERPology” (Luck, 2014) in the 1990s and 2000s, it became possible to test the involvement of distinct brain regions and the concomitant electrical activity for various linguistic processes, given the standard assumption that language is a cognitive system. The ERP community has spent a great deal of time decomposing the major components, such as the P600 and N400. It is taken for granted that the level of analysis provided by these “large” components does not suffice at the electrophysiological level to describe generic linguistic sub-operations. The urge to seek a finer level of granularity, then, is clearly manifested in the ERP community through EEG and MEG investigations (Lau et al., 2008), but this objective is not found in the vast majority of cartographic neuroimaging research.

In recent decades, neuroanatomical inquiry into the structures responsible for syntactic processing has led to a number of revelations concerning the biology of language. Petersson et al. (2012) reveal the inadequacy of the classical Broca-Wernicke-Lichtheim language model of the brain by noting how the language network extends to substantial parts of superior and middle temporal cortex, inferior parietal cortex, along with subcortical areas such as the basal ganglia (Balari and Lorenzo, 2013), the hippocampus and the thalamus (Theofanopoulou and Boeckx, Forthcoming a). The network is also implicated in more general cognitive systems like the default-mode network and the multiple demand system.

Brodman area 44 and the posterior superior temporal cortex appear to be involved in a pathway which supports core syntactic computations (Friederici et al., 2006, see also Tettamanti and Weniger, 2006; Santi and Grodzinsky, 2010), with the combinatorial network being identified by Poeppel (2014) as the anterior medial temporal gyrus and anterior inferior temporal sulcus. Lieberman’s (2006) “Basal Ganglia Grammar” model proposes the existence of a pattern generator whose excitation/inhibition mechanism is located in the basal ganglia. This interfaces with working memory space located in Broca’s area (Santi et al., 2015). Lieberman estimates that the dorsolateral prefrontal circuit is involved in sentence comprehension, projecting from the prefrontal cortex toward the lateral dorso-medial region of the globus pallidus, and the

thalamus, which projects back to the prefrontal cortex. Balari and Lorenzo (2013, pp. 100–102) have suggested that this may be the circuit used as language’s computational system operating within a structure of working memory networks (Balari et al., 2012).

## Evo-devo Directions

As the theory of evolution expands beyond the Modern Synthesis and into areas such as evolutionary-developmental (evo-devo) biology (Carroll, 2006; Bolker, 2008) there is in turn more potential for space for linguists to find their place within biology. In the evo-devo program, following the lead of traditional formalists such as Vicq-D’Azyr, Goethe and Owen (Amundson, 1998, 2006), natural selection is “a constantly operating background condition, but the specificity of its phenotypic outcome is provided by the developmental systems” (Pigliucci and Müller, 2010, p. 13). Evo-devo departs from Neo-Darwinian adaptationism (NDA), or “phylogenetic empiricism” (Chomsky, 1968), in that it takes the saltationist view that species are the result of punctuated genetic changes. The functionalism of NDA should also be rejected, since functions do not typically pre-exist organic form (Müller, 2008), which is determined by morphogenetic parameters such as the viscoelastic properties of cellular matrices and the kinetic activity of cellular diffusion (what Alberch termed “morphological evolution”), and which at best have what Balari and Lorenzo call a “functional potential” (2013, p. 37). Contrary to ideas in Dawkins (2006, p. 202) and Lieberman (2015), laws governing the conservation of developmental pathways should be “acknowledged with a creative character similar—if not superior—to that of natural selection” (Balari and Lorenzo, 2013, p. 115). Form often precedes function, then, and natural selection acts as a “filtering condition on pre-existent variants”; thus “arrival of the fittest, instead of survival of the fittest, is the core issue in any evolutionary study” (Narita and Fujita, 2010, P. 364, see also Bertossa, 2011).

In this connection, Rakic and Kornack (2001) observe that the phase of asymmetric cell division yielding neuronal cells differs in timing between humans and monkeys to the extent that human neuronal populations are thought to be between 8 and 16 times larger than those of monkeys. Human-specific neuronal traits include the protein ApoE4, providing stronger synaptic connections (Bufill and Carbonell, 2004). Parker and McKinney (1999) detail how the myelination of the neocortex occurs in humans until the age of 12, but lasts only 3.5 years in rhesus monkeys. Zhang et al. (2011) also propose the existence of 1241 primate-specific genes, 280 of which are human-specific. 54% of these human-specific genes are upregulated in a brain area implicated in higher cognition, the prefrontal cortex. These new genes are “much more likely to be involved in gene regulation” (Diller and Cann, 2013, p. 256), a major topic in evo-devo.

Recent research in avian genomics suggests that the evolution of externalization may also not be as difficult as typically considered by generative grammarians. Pfenning et al. (2014, p. 1333) demonstrated that the profiles of transcription genes in vocal learners can be aligned, with 50 genes being shared between humans and birds which are “enriched in motor control

and neural connectivity functions.” Both humans and birds appear to have converged on identical solutions to vocal learning; a remarkable finding considering the 310 million year gap separating birds from humans. In summary, a slight epigenetic change, termed the “Small Bang” in Murphy (2015a), could have produced an alteration in the human computational system. The next section will consider how these operations could be implemented in the brain.

## Rhythmic Directions

How much physiological detail is required to capture the operations of the language faculty? Theofanopoulou and Boeckx (Forthcoming b) claim that studying neural dynamics only at the level of brain waves is sufficient, but as demonstrated below, a more refined biophysical picture is not only possible but in fact necessary to adequately explain the origins of linguistic computations like concatenation, cyclic transfer and labeling. What is needed is not just a neuroscience of language, but a *neurophysiology* of language. For instance, at the most general mesoscopic physiological level of local neuronal groups, synchronized firing patterns result in coordinated input into other cortical areas, which gives rise to the large-amplitude oscillations of the local field potential. Inhibitory interneurons play an important role in producing neural ensemble synchrony by generating a narrow window for effective excitation and rhythmically modulating the firing rate of excitatory neurons. Interneurons place constraints on the oscillations responsible, as argued here, for computation. Subthreshold membrane potential resonance may also contribute to oscillatory activity by facilitating synchronous activity of neighboring neurons. As Cannon et al. (2014, p. 705) note, “the physiology underlying brain rhythms plays an essential role in how these rhythms facilitate some cognitive operations.”

Shifting focus from neuroimaging to more recent investigations of brain oscillations may provide a welcome (but as yet tenuous) way of reconstructing in neural terms the operations of theoretical linguistics. Brain rhythms “have come of age,” as Buzsáki and Freeman (2015, p. v) put it. They reflect synchronized fluctuations in neuronal excitability and are grouped by frequency, with the most common rhythms being delta ( $\delta$ :  $\sim 0.5$ – $4$  Hz), theta ( $\theta$ :  $\sim 4$ – $10$  Hz), alpha ( $\alpha$ :  $\sim 8$ – $12$  Hz), beta ( $\beta$ :  $\sim 10$ – $30$  Hz), and gamma ( $\gamma$ :  $\sim 30$ – $100$  Hz). These are generated by various cortical and subcortical structures, and form a hierarchical structure since slow rhythms phase-modulate the power of faster rhythms.

It is by now well established that neural oscillations are related to a number of basic and higher cognitive functions, for example speech perception (Giraud and Poeppel, 2012; Kayser et al., 2014). According to Giraud and Poeppel’s temporal linking hypothesis, oscillation-based decoding segments information into “units of the appropriate temporal granularity” (2012, p. 511). Oscillations may consequently explain how the brain decodes continuous speech, however Giraud and Poeppel’s form of dynamic research crucially centers on the segregation of phonological, and not semantic or syntactic units, which may implicate different brain areas and rhythms. The  $\gamma$ ,  $\theta$ , and

$\delta$  rhythms respectively correspond closely to (sub)phonemic, syllabic and phrasal processing, as Giraud and Poeppel note, restricting their experimental inquiry to the  $\gamma$  and  $\theta$  bands. In addition, the neural dynamics responsible for syntactic operations may be obscured by the processing of external sensory events like speech, and so different experimental designs may be required to control for this.

Oscillations have also been linked to the timing of cortical information processing (Klimesch et al., 2007). As Vaas notes, “Intrinsic oscillatory electrical activities, resonance and coherence are at the root of cognition” (2001, p. 86), with the condensing and dissolving of oscillatory bursts possibly explaining the “cinematic” nature of subjective experience (Freeman, 2015). As Poeppel has put it, the brain essentially “breathes” through oscillations. If such generic neural operations are also shown to be responsible for syntactic computations, and not just linguistic perception, this would lend weight to Hagoort’s (2014) interpretation of the cartographic literature, which holds that the establishment of an axis of language production and comprehension is not justifiable. Expanding on Giraud and Poeppel’s (2012, p. 511) goal of establishing a “principled relation between the time scales present in speech and the time constants underlying neuronal cortical oscillations,” one of the central challenges will be to draw up relations between oscillatory time constants and the time scales of syntactic computation. This latter topic has yet to be explored in any serious detail, possibly due to a widespread prejudice that neurolinguistic investigations of syntax must analyse phrasal units, such as noun and verb phrases, rather than the underlying operations which construct them, such as set-formation and labeling (although see Ohta et al., 2013 for an innovative approach to localizing Merge and Search operations).

## Oscillations as Functional Units

Recent debates about the origins of ERP component generation have led some (Tass, 2000; Makeig et al., 2002) to propose that components do not arise purely from latency-fixed polarity responses which are additive to continuing EEG responses, but rather arise through a superposition of oscillations which reset their phases in reaction to sensory input (although see Sauseng et al., 2007 for the methodological limitations of particular phase resetting claims). For our purposes, it is worth noting that this phase reset model was the first to propose a strong dependency between components and oscillations, introducing to brain dynamics a *functional* and not purely *electrophysiological* role. This immediately granted researchers the ability to transfer understanding of components (which are in turn linked to cognitive faculties) to brain rhythms whilst correspondingly inferring the nature of components from an emerging understanding of oscillations. While cognitive electrophysiologists have embraced this integrally reciprocal perspective (Klimesch et al., 2004), linguists generally remain hostile to the claim that the nature of mental computations—like components—could be explored explicitly through biophysics.

While the cognome resides at the Marrian computational level (Marr, 1982), I would like to suggest that there is in fact no algorithmic level at syntax. At most there are algorithms

at the interfaces. Psycholinguistic theories can algorithmically model language processing, as Neeleman (2013) discusses, but syntax itself (being composed of operations like Concatenate, Label, and Transfer) has no need for this. Nevertheless, the dynamome, with its operations of information segregation and spike timing organization, can in some sense be seen as an algorithmic level, implemented by the cellular structures of the connectome. These Marrian concerns become more vivid when we consider with Martins and Boeckx (2014) that syllables, which are unique to humans, evolved from primate lip-smacking. In terms of brain rhythms, they are both identical, yet one is human-specific and another is not. The implications for the study of labeling, not acknowledged in Murphy (2015a), are clear: only comparative investigations of domain-general neurophysiological mechanisms, and the context in which they operate, will lead to enhanced understanding of human-specific computations. There are two central approaches to the cognome-dynamome one could adopt: re-construct the cognome from the bottom-up, or import linguistic constructs into a model of the dynamome. I will be primarily concerned with the latter methodology, though the material reviewed and the model outlined open up the possibilities of using neurophysiology to guide linguistic investigations.

### The Basic Label Model of the Cognome-dynamome

At the most general level of analysis, neural oscillations emerge from the tension between the brain's two most central principles: segregation of function and dynamic integration (de Pasquale et al., 2012). Human brains are highly complex dynamical systems with principles of cellular and electrochemical organization which range across a hierarchy of scales. The brain cannot function purely through anatomical connections—the *locus classicus* of standard neuroimaging studies—but additionally requires dynamic functional connectivity, achieved through oscillatory synchronization. Frequency bands alone are not sufficient for computation; rather, it is their interactions which are significant. Intuitive prejudices against studying complex systems in these dynamical terms abound: for instance, chemical dynamics are typically thought about in terms of reaction kinetics, being stipulated as pre-formed stable variables, ignoring the molecular composition/decomposition process.

A core feature of the brain's functional complexity is created by rhythms generated in different cortical and subcortical tissue. Oscillations denote distinct states of brain activity, while oscillatory activity reflects a dynamic interplay between the dissimilar cell types of discrete circuits (Buzsáki, 2006). Brain rhythms, with their inter-wave hierarchies, provide “a syntactical structure for the spike traffic within and across circuits at multiple time scales” (Buzsáki and Freeman, 2015, p. viii). “Phase synchronization” will additionally be a central notion to the present discussion, referring to a consistent phase coupling between two neuronal signals oscillating at a given frequency.  $\gamma$  band synchronization (GBS) in particular has been intensively studied due to its apparent role in phase coding and perceptual integration (Fries, 2009), and is thought to be a major process subserving a fundamental operation of cortical computation

implicated in various cognitive functions. Which functions are involved depends ultimately on what neural circuits GBS operates on. The following sub-sections will present a way of exploring the operations of the cognome in terms of these dynamic operations, leading to a form of what I will call Dynamic Cognomics.

### Concatenation

The central proposal of the model pursued here is that the interaction of brain rhythms yields linguistic computation. Lower frequencies such as the  $\alpha$  range are known to synchronize distant cortical regions; procedures which may represent the substrates of linguistic cross-modular transactions (Kinzler and Spelke, 2007). More precisely, I will assume that the  $\alpha$  band embeds  $\gamma$  rhythms generated cross-cortically, yielding a form of inter-modular conceptual combination, the electrophysiological equivalent of concatenation. The assemblies implicated by the  $\gamma$  range may have been influenced by the extended neocortical myelination discussed above, with direct effects on the network of information stored across such regions. This is consistent with recent claims that  $\alpha$  is responsible for the binding of visuo-spatial features (Roux and Uhlhaas, 2014) and is deployed in the service of determining successful lexical decisions (Strauss et al., 2015). I will further assume that the items concatenated are also initially “lexicalized” by  $\alpha$ -embedded cell assemblies oscillating at the  $\gamma$  range within supragranular layers of the default-mode network (Raichle et al., 2001).

### Transfer

Linguists take concatenation to occur cyclically (Chomsky, 2008), and so I will additionally assume that this Spell-Out/Transfer process is realized through embedding the above  $\gamma$  rhythms inside the  $\theta$  band, which finds its source in the hippocampus. I will adopt the claim of Theofanopoulou and Boeckx (Forthcoming b) that  $\gamma$  must be decoupled from the  $\alpha$  band through the activity of the thalamic reticular nucleus for  $\gamma$ - $\theta$  embedding to take place. Both types of Transfer operations—Spell-Out to the sensorimotor interface, SM, Interpret to the conceptual-intentional interface, CI—will be subsumed under this approach, which at a minimum involves this desynchronization of  $\alpha$ -generated structures and consequent  $\theta$ -synchronization. Though the thalamic reticular nucleus is here identified as a core component of desynchronization, other regions may also be involved. Due to its role in  $\gamma$ - $\theta$  embedding in auditory processing (Nosarti et al., 2004), the posterior corpus callosum is also likely to be heavily involved in Transfer operations.

### Labeling

Along with concatenation and transfer, there is also labeling. Two major observations have been made about this operation: (i) It is unique to humans (Murphy, 2015a); (ii) It is based on principles of minimal computation (Chomsky, 2015a). Labeling is also monotonic in that once a set has been labeled (as a verb or determiner phrase, for instance) its identity is sustained when embedded inside another set. Since labeling must take place at the point of transfer to the interfaces (to prevent a structure being a Verb Phrase at CI but a different

phrase at SM), labeling must be seen as a core syntactic operation (Murphy, 2015b; Piattelli-Palmarini and Vitiello, 2015), and not emerging epiphenomenally at the interfaces, despite it having a less central role than unconstrained “Merge” (concatenation) which operates independently from either CI or SM.

As Boeckx and Theofanopoulou (2015) note, labeling was not formulated at a fine enough level in Murphy (2015a) to avoid the granularity mismatch problem. In order to correct for this, I will define labeling as the attribution to a concatenated set some categorical specification created from the Labeling Assembly, which is composed of aspects of (i) general cognitive constraints, (ii) the CI system, (iii) the cognome and (iv) the precursor lexicon ( $\rho$ LEX). The final of these four constituents is taken to be the set of flat and atomic “root” structures (Boeckx, 2014a), from which morphology constructs internally hierarchical words (Nóbrega and Miyagawa, 2015). When *John* is concatenated with *ran*, the labeling algorithm produces a Verb Phrase, not a Noun Phrase (see Adger, 2013; Narita, 2014a and Murphy, 2015a for further algorithmic details). This covers the basic outline of labeling, but in order to achieve a finer level of granularity it will be necessary to descend to the dynamic level, and ultimately (in the final section) the cellular level.

In dynamic terms, I will take labeling to be the slowing down of  $\gamma$  to  $\beta$  followed by  $\beta$ - $\alpha$  coupling, involving a basal ganglia-thalamic-cortical loop (see Cannon et al., 2014 for the rhythmogenesis of  $\beta$  in the basal ganglia). This would disinhibit the thalamic medio-dorsal nucleus via the  $\beta$  band. This frequency coupling arises from a relationship between oscillations which form a hierarchy such that the speed of the slower rhythm controls the power of the faster rhythm. Due to its involvement in phrasal processing, I will assume that the  $\delta$  band may be involved in the later stages of this process. The role of the thalamo-cortical network as a slow rhythm generator, and hence a single dynamic and functional unit of brain oscillations, has been recently supported by Crunelli et al.’s (2015) review of the EEG literature. Accumulating evidence suggests that  $\beta$  holds objects, whereas  $\gamma$  merely generates them (Martin and Ravel, 2014). Dean et al. (2012) also show how  $\beta$  is an excellent candidate for comparing old and new information from distinct modalities due to its wider temporal windows; that is, it would compare phase heads (old information) with late-merged non-phasal elements like complements (new information), likely drawing on different conceptual representations and hence different “core knowledge systems” and brain regions (Spelke, 2010). Related both to Balari and Lorenzo’s (2013) claim that the basal ganglia is the center of their “Central Computational Complex” and Jouen et al.’s (2013) findings that this structure is implicated in acquiring the serial response order of a sequence, Theofanopoulou and Boeckx (Forthcoming b) propose that this region holds one of the  $\gamma$ -supported items before slowing it down to the  $\beta$  frequency as a consequence of the conduction delays resulting from the surrounding neural regions. Thus the  $\beta$  band accomplishes the role of labels, a claim supported by findings that  $\beta$  activity maintains existing cognitive states (Engel and Fries, 2010). More broadly, the basal ganglia and the striatum are implicated in sequencing and chunking, with striatal structures operating at

the  $\beta$  range (Leventhal et al., 2012). The core position occupied by the basal ganglia in this labeling model also fits well with imaging studies which have revealed the region’s involvement in “syntactic complexity,” specifically the processing of type-identity intervention of matching labels, being activated in a recent fMRI study when a noun phrase similar to the dependency head in a long-distance dependency intervenes in the dependency (Santi et al., 2015). Basal ganglia nuclei in humans are also around twice as large as would be predicted for a primate of our size (Schoenemann, 2012), and since humans do not appear to have substantially more sophisticated movements than apes, this increase may well have supported higher cognitive capacities like labeling.

## Formal Considerations

Introducing new formalisms will permit a clearer explication of dynamic cognomics. Although they appear similar, what follows will have no direct bearing on, and should not be considered an extension of, standard set-theoretic notational conventions relating to such things as functional application.

First, we can notate  $\gamma$ - $\theta$  embedding as  $\{\theta(\gamma)\}$ , with  $\gamma$  being embedded inside  $\theta$  rhythms. If it is known how many  $\gamma$  cycles are to be embedded (for instance, 7), this can be notated as  $\{\theta(\gamma^7)\}$ . We can notate the decoupling process required to transfer concatenated structures as  $\gamma(\bullet)\alpha$ , where  $\gamma$  is decoupled from the  $\alpha$  band. Frequency coupling can correspondingly be notated as  $\gamma \bullet \alpha$ . The decreasing of  $\gamma$  to  $\beta$  can be represented as  $\gamma \leftrightarrow \beta$ , where “ $\rightarrow$ ” refers to a state change. Post-phrasal syntactic reanalysis and wrap-up effects can be represented with  $\psi$ . Finally, the (hypothetical) cell assemblies responsible for particular lexical features, such as the [+singular] feature of *man*, can be represented as  $\zeta[\text{man}(+\text{singular})]$ . If it is known in which regions (cytoarchitectonic or otherwise) such assemblies are located, this can be represented as, for instance,  $\zeta:\text{BA44}[\text{man}(+\text{singular})]$ , while the rhythm band can be additionally represented as  $\zeta[\text{man}]:\gamma$ .

We are now in a position to write a simple derivation. Take the sentence *The man is called John*. This can be represented in familiar syntactic terms as a Tense Phrase, ignoring superfluous details (e.g., morphological operations):  $[\text{TP}_{[\text{DP } \textit{The man}]}[\text{T}[\textit{is}]][\text{VP } \textit{called John}]]$ . In the interests of clarity, I will put aside precise categorical concerns and denote labeled phrases with “L,” with multi-phrasal labels being italicized. Even though sentences are parsed in a left-right fashion, generative linguistics holds that syntactic derivations proceed right-left. In order to deal with this perennial psycholinguistic problem, I suggest that structures are concatenated, labeled and transferred as and when they are heard, read or otherwise perceived, and after every lexical unit a “look back” procedure is triggered to reanalyse the labels and features of each structure, denoted here by  $\psi$  (see Chesi, 2015 for a comprehensive left-right derivational proposal). In psycholinguistic terms, this may account for certain wrap-up effects which occur when subjects reach the final word of a sentence during online processing (Field, 2004). This approach is also consistent with the “one-system” contention of Lewis and Phillips (2015) that grammatical theories and language processing models describe the same cognitive system,



as evidenced by the fact that grammar-parser misalignments only seem to occur as a consequence of limitations in domain-general systems such as memory access and control mechanisms. It follows that “online and offline representations are the product of a single structure-building system (the grammar) that is embedded in a general cognitive architecture, and misalignments between online (“fast”) and offline (“slow”) responses reflect the ways in which linguistic computations can fail to reflect the ideal performance of that system” (Lewis and Phillips, 2015, p. 39). This one-system hypothesis also proposes that the grammar goes through a series of structure destruction and rebuilding operations as new words are encountered; a process which aligns well with the rhythmicity of the present model and the effects of  $\psi$ .

The derivation will proceed as follows. *The* is generated by distributed  $\gamma$  activity in the supragranular cell assemblies responsible for its long-term storage,  $\zeta[the]$ . This rhythm would be embedded within  $\alpha$  activity before being transferred to the interfaces through being decoupled from  $\alpha$  and newly embedded within hippocampal  $\theta$  activity.  $\zeta[man]$ , operating at the  $\gamma$  range, would then be embedded within  $\alpha$  before being transferred. The two representations would then be labeled a Determiner Phrase at the Labeling Assembly, which I will identify as the circuits connecting the thalamus, basal ganglia, prefrontal cortex, and anterior temporal regions. To achieve labeling, the embedded cycles would be slowed to the  $\beta$  range ( $\gamma \leftrightarrow \beta$ ) before being coupled to  $\beta$  ( $\beta \bullet \alpha$ ). The labeled phrase [*the man*] would be maintained in memory via the  $\beta$  rhythm. The subsequent material [*is called John*] would then be added in a similar fashion:

$$\begin{aligned} &\zeta[the]:\gamma \rightarrow \{\alpha(\zeta[the]:\gamma)\} \rightarrow \alpha(\bullet)\zeta[the]:\gamma \rightarrow \{\theta(\zeta[the]:\gamma)\} \\ &\zeta[man]:\gamma \rightarrow \{\alpha(\zeta[man]:\gamma)\} \rightarrow \alpha(\bullet)\zeta[man]:\gamma \\ &\rightarrow \{\theta(\zeta[man]:\gamma)\} \\ &\{\theta(\zeta[the]:\gamma)(\zeta[man]:\gamma)\} \\ &\psi \\ &\gamma \leftrightarrow \beta \\ &\alpha(\bullet)((\zeta[the]:\beta)(\zeta[man]:\beta)) \rightarrow \alpha(\bullet)(\zeta[the\ man]:\beta) \\ &\zeta[is]:\gamma \rightarrow \{\alpha(\zeta[is]:\gamma)\} \rightarrow \alpha(\bullet)\zeta[is]:\gamma \rightarrow \{\theta(\zeta[is]:\gamma)\} \\ &\gamma \leftrightarrow \beta \\ &\alpha(\bullet)((\zeta[the\ man]:\beta)(\zeta[is]:\beta)) \rightarrow \alpha(\bullet)(\zeta[the\ man\ is]:\beta) \\ &\psi \\ &\zeta[called]:\gamma \rightarrow \{\alpha(\zeta[called]:\gamma)\} \rightarrow \alpha(\bullet)\zeta[called]:\gamma \\ &\rightarrow \{\theta(\zeta[called]:\gamma)\} \\ &\gamma \leftrightarrow \beta \\ &\alpha(\bullet)((\zeta[the\ man\ is]:\beta)(\zeta[called]:\gamma)) \\ &\rightarrow \alpha(\bullet)(\zeta[the\ man\ is\ called]:\beta) \\ &\psi \\ &\zeta[john]:\gamma \rightarrow \{\alpha(\zeta[john]:\gamma)\} \rightarrow \alpha(\bullet)\zeta[john]:\gamma \\ &\rightarrow \{\theta(\zeta[john]:\gamma)\} \\ &\gamma \leftrightarrow \beta \\ &\alpha(\bullet)((\zeta[the\ man\ is\ called]:\beta)(\zeta[john]:\gamma)) \\ &\rightarrow \alpha(\bullet)(\zeta[the\ man\ is\ called\ john]:\beta) \\ &\psi \end{aligned}$$

Notice that, as with Computational Ethology (Murphy, 2015a) and recent syntactic proposals (Hornstein, 2009; Adger, 2013), labeling is here placed at the center of the dynamo’s linguistic operations. As a result, call the above cognome-dynome hypothesis the Basic Label model. What remains to be added to the derivation by empirical investigation are the factors of time-frequency domain and the anatomical regions of cellular assemblies (e.g., “embed  $\gamma$  of region  $r$  within  $\alpha$  of region  $s$  for time  $t$ ”). All elements in the derivation, then, are created as simple  $\gamma$  assemblies, and only some (namely, labeled phase heads) become more complex  $\beta$  assemblies; consider the difference between adverbs like *nearly* and verbs like *ran*. The rhythmic division of complexity which follows from this is supported by Honkanen et al. (2014), who demonstrated that simple objects represented in visual working memory employ the  $\gamma$  band, while more complex objects are represented by the  $\beta$  band. This oscillatory procedure also matches the generative view that phase heads have a longer derivational life than non-phase heads (Boeckx, 2014a; Narita, 2014a). The role attributed here to  $\gamma$  assemblies additionally finds some support in Bastiaansen and Hagoort’s (2015) EEG study of semantic unification, which detected larger  $\gamma$ -band power for semantically coherent than semantically incongruent sentences. Larger  $\beta$ -band power was also found for syntactically correct sentences relative to ungrammatical sentences, lending support to the hypothesized labeling power assigned to  $\beta$  in the present model.

## Some Empirical Consequences of Dynamic Cognomics

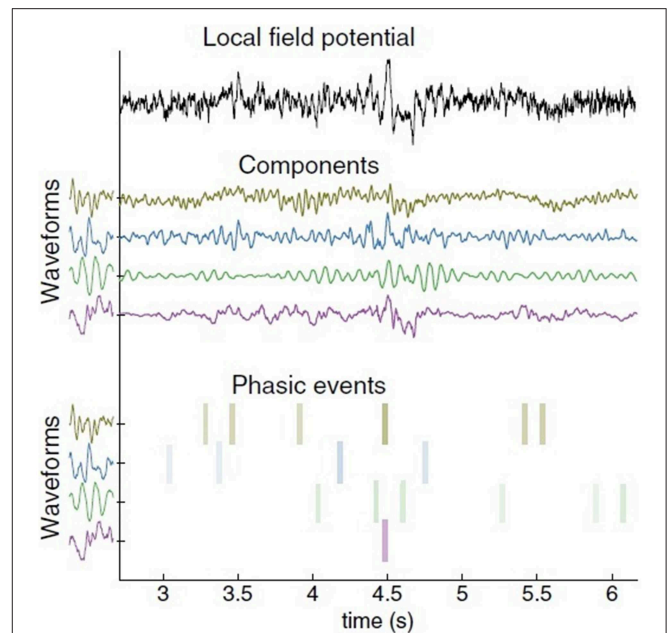
Among many other forms of imaging and behavioral data, neuroimaging studies should be used as a guide for dynamic cognomic investigations. With respect to linguistic computation, the left anterior temporal lobe has been implicated in basic combinatorics (concatenation) and phrasal construction (labeling) (Bemis and Pykkänen, 2013; Westerlund and Pykkänen, 2014), while the posterior middle temporal gyrus is involved in lexical access (lexicalization) and ambiguity resolution (Turken and Dronkers, 2011). Given the present rhythmic perspective on linguistic computation, the much-discussed fronto-temporal language network would consequently be purely an *output* system of the above operations, not a core syntax region. Friederici (2012) holds that distinct regions of the left inferior frontal gyrus are responsible for “different” types of syntax, arguing, for instance, that the dorsal stream is only implicated in embedded structures or structures deviating from normal ordering. Yet, as the above model makes clear, the basic combinatorics are universal across syntactic structures, whether simple or complex; set-formation is still set-formation whether it is found in a small clause or a Shakespearean sonnet.

While relatively little is known about how oscillations relate to cognitive operations, significant advances could come from direct empirical investigations teasing apart  $\gamma$  and  $\beta$  from other rhythms, demonstrating a correlation with a syntactic manipulation (and perhaps a dissociation with another operation which could be linked to slower rhythms and working memory

or attention processes; see Lakatos et al., 2008 for the role of oscillations in attention). Due to its high temporal and spatial resolution and signal-to-noise ratio, electrocortical topography is also highly applicable to testing the Basic Label model, having been used to investigate speech production (Bouchard and Chang, 2014), language comprehension (Cervenka et al., 2011), and having been flexibly deployed both in humans and animals. In addition to the cartographic studies above, paradigms such as that in Ohta et al. (2013), which differentiate the neural correlates of concatenation and search/agreement operations, could be employed. Despite having noted the limitations of cartographic studies, an area of ongoing neurolinguistic research is the spatial scales of brain rhythms. It could be explored, for instance, whether ongoing oscillations and generic computations share the same neuronal generators. Emerging technologies to experimentally test and refine the Basic Label model include high-density electrode recordings and optogenetic tools (Chow et al., 2010; Viventi et al., 2011), along with the more traditional EEG and MEG devices. Bemis and Pytkäinen (2013) showed that between 200 and 300 ms after the presentation of a word which can be combined with a previous item, the left anterior temporal lobe is activated, implicating this region in semantic composition. This would consequently be a good estimate of when oscillation studies might detect labeling effects to arise, given the role of labels in semantic composition (Hornstein and Pietroski, 2009; Murphy, 2015b). At the most general level of lexical comprehension, EEG and MEG studies would also predictably find coherent oscillatory activation of large neuronal assemblies when processing words relative to processing pseudowords, as Pulvermüller et al. (1994) found. A level of cortical entrainment would also be predicted for non-syllabic, phrasal, and sentential structures during the auditory presentation of simple stimuli; structures which are not part of any speech stream but are rather internally constructed by the comprehender, and whose rhythmic generators would likely align closely with the regions implicated in the Basic Label model.

Neural potentials have typically been analyzed through frequency, time-frequency, and wavelet representations (Kaiser, 2010). Independent component analysis (ICA) has also been used successfully in estimating the sources of neural systems given multiple recording locations, since these systems generate independent and continuous activity and combine linearly and instantaneously (Hyvärinen and Oja, 2000). However, spatial ICA does not allow the interpretation of time-varying patterns, and in the case of EEG it also does not produce a model of “phasic events” of rhythmic activity.

Given these shortcomings, I would like to introduce the possibility of analysing a continuous signal as a linear combination of reoccurring waveforms. This is achieved by combining overcomplete representations with adaptive signal models. If the goal is to extract waveforms from a single continuous channel, then it follows that we should adopt a generative model which summates impulse responses, being a multiple input, single output (MISO) model. Principe and Brockmeier (2015, p. 15) term this a *phasic event model*. This proceeds in two steps: learning a set of waveforms occurring repeatedly throughout a signal, and estimating an atomic



**FIGURE 1 | Decomposition of a single local field potential (LFP)**

**channel using the phasic event model.** Data collected by Brandi Marsh in Joseph Francis's laboratory at SUNY-Downstate. The original LFP signal is on top. In the middle are the component decompositions. The learned impulse response of the waveform is shown to the left of each component. The most significant amplitude atoms (timing, amplitude, and waveform index) appear at the bottom as colored bars. Color intensity corresponds to amplitude (from Principe and Brockmeier, 2015, p. 15).

decomposition of a signal in terms of timing, amplitude, and waveform index (see **Figure 1** for an example). The major advantages of this over other models is that the phasic event analysis learns the reoccurring waveform shape and allows the pinpointing of the amplitude and timing of phasic events. The model consequently captures the transitory nature of neural events.

Given the structure of the Basic Label model and the division of EEG patterns and local field potentials into rhythms ( $\alpha$ ,  $\beta$  etc.) and phasic events (sharp waves,  $\beta$  and  $\gamma$  ripples etc.), I think an approximate correlational (in the sense of Embick and Poeppel, 2015) division between computations and representations can be established between, respectively, phasic events (carried out in and between the cell assemblies of particular regions) and rhythms (necessarily localized at such regions).

Although oscillations are likely not all that is needed to provide a solution to the problem of linguistic computation, they nevertheless appear to be a vital part of the answer. Aside from language-centered obstacles, comparative dynamic cognomics will also face the notable challenge of the variation in oscillation presence across species, with the reasons for much rhythmic variation still unknown. For now, the Basic Label model satisfies the cognome-dynome operational level, but we would ultimately want to satisfy the connectome and other lower levels. As a result, the next section will expand on the bare electrophysiological details outlined above leading to the broadening of multi-disciplinary concerns and perspectives.

## Biophysical Directions

To adequately explore the neurochemical and biophysical details of the Basic Label model, it is useful to introduce a distinction between minimal and maximal degrees of explanation:

- (1) a. Minimal degree of explanation (MinDE): The use of brain dynamics to explain why the cognome performs the operations it does, and not some other imaginable operations.
- b. Maximal degree of explanation (MaxDE): The use of brain dynamics in addition to causally relatable accounts of neurochemistry and its underlying biophysics to explain why the cognome performs the operations it does, and not some other imaginable operations.

Note that MinDE has minimal requirements, whereas MaxDE has no stipulated limits, embracing the full range and plurality of the natural sciences. Neuroimaging studies, for instance, do not even reach the level of MinDE, whereas a purely rhythmic approach to the dynome of the kind found in Theofanopoulou and Boeckx (Forthcoming b) satisfies MinDE without reaching the neurochemical and biophysical precision of MaxDE. Kopell et al. (2014, p. 1319) stress that connectome-dynome linking hypotheses need to be supplemented with “the biological details that relate this connectivity more directly to function.” This is where I will attempt to depart from analyses which remain at the levels of the dynome and cognome (e.g., Sporns, 2013). For instance, Theofanopoulou and Boeckx (Forthcoming b) only refer in passing to basic interneuron classes, and their model lacks any serious neurobiological details. As Allen and Monyer (2015, p. 85) comment, “when considering interneurons, it would be important to investigate the role they play in the reactivation of cell ensembles occurring during sharp wave/ripples.”

Mechanistic ventures beyond the dynome are, I think, in the proper spirit of Turing’s (other) thesis regarding morphogenesis, which was concerned not just with a description of an organism’s forms (similar to the computational level of modern linguistics) but also with a proto-evo-devo theory of the cellular mechanisms which give rise to such forms (Turing, 1952, see also Maini, 2004). As Kopell et al. (2014, p. 1324) note, “an immersion in the physiology supporting temporal dynamics suggests mechanisms that would not be obvious if one were thinking abstractly about computation and rhythms”; a statement which carries urgent lessons for theoretical linguistics and neuroimaging.

Contrary to much of Koch’s (1999) ambitious work, the following section will argue that the divide between biophysics and computation is in fact incommensurable, and that a different biolinguistic strategy will be required to resolve the granularity mismatch problem. This approach will use the Basic Label model alongside neurochemistry as tools to construct a neurobiologically feasible cognome, free of the technical baggage—though not the methodological naturalism (Chomsky, 2000; Collins, 2015)—of minimalist syntax and its lexico-centrism and “featuritis” (Boeckx, 2014a).

## Feeble Currents and Cognomic Substrates

Though much interdisciplinary work remains to be carried out, dynamic cognomics has the potential to progress neurolinguistics beyond the situation described by Szathmáry in 1996: “Linguistics is at the stage at which genetics found itself immediately after Mendel. There are rules (of sentence production), but we do not yet know what mechanisms neural networks are responsible for each rule” (1996, p. 764). So far, I have only presented a model of how to embed the cognome within the dynome, but it is also vital to ground the dynome within the connectome and microlevel analyses, in turn addressing Szathmáry’s concern.

It has been shown that neuronal populations can synchronously discharge due to an internal or external event, and additionally as a result of dynamic interactions between reciprocally coupled networks, which serve to “tag the responses of neurones that need to be related to one another,” as König (1994, p. 31) put it in his seminal assessment of neural oscillations. This synchronous activity further tends to be oscillatory in nature (Liu et al., 2010). Oscillations have also been linked to neurochemistry (Muthukumaraswamy et al., 2009), as discussed below. While oscillatory electrical activity in cell assemblies has been observed since the 1920s beginning with Berger’s (1929) ground-breaking work, inspired by the Liverpool surgeon Caton’s (1875) studies of the “feeble currents” generated by rabbit and monkey brains, its role in cognitive capacities has been intensively explored only since the new millennium (Jensen et al., 2002; Ossandón et al., 2011), largely down to theoretical, technological, and optogenetic advances. Updating Caton’s imagery, McCormick et al. (2015, p. 133) summarize that brain rhythms are generated through “the interaction of stereotyped patterns of connectivity together with intrinsic membrane and synaptic properties.”

At the most common level of investigation, time-locked frequency analysis can decompose an EEG signal and identify changes in oscillations. But the widespread use of non-invasive and high-temporal resolution MEG, and recent advances in its source localization power (Wipf et al., 2010), have led to enhanced understanding of the spatiotemporal dynamics of oscillations and how they operate within neural networks. Recent work has begun to deliver an increasingly precise account of how, for instance, different classes of GABAergic interneurons in the hippocampus coordinate activity giving rise to network oscillations (Allen and Monyer, 2015), strengthening dynome-connectome correspondences. GABA<sub>B</sub> receptors also perform time integration of cell assemblies (classically defined as a set of neurons exhibiting stronger within-group connectivity than with other connected neurons; Hebb, 1949) from the subsecond to second scale (Deisz and Prince, 1989), a vital function in computing conceptual and linguistic information representations.

Going beyond this level of analysis will require mapping rhythms to the numerous interneuron classes, which are defined based on cell body location, expression of marker proteins, axonal arborization, and other properties (Whittington and Traub, 2003; Klausberger et al., 2005; Somogyi and Klausberger, 2005). Korotkova et al. (2010) attempted to reach such a



goal by showing how the removal of NMDA receptors in parvalbumin-expressing (PV) interneurons reduced the power of  $\theta$  oscillations in the CA1 hippocampal region, while also reducing the  $\gamma$ -power modulation by  $\theta$  oscillations. PV interneurons and somatostatin-expressing (SOM) interneurons preferentially synapse, respectively, onto the cell bodies and proximal dendrites of pyramidal cells and the distal dendrites of pyramidal cells (Royer et al., 2012). The silencing of PV interneurons, but not SOM interneurons, altered the  $\theta$  phase precession in the brains of mice running on a treadmill belt in the experiments conducted by Royer and colleagues, suggesting that PV interneurons are highly fit to control the firing phase of principal neurons during  $\theta$  oscillations, permitting the extension of a causal chain from cognome to dynome to a specific part of the connectome.

It should be noted, however, that PV and SOM expression is common to numerous hippocampal interneuron classes, and so further optogenetic work is needed in order to establish the role of individual interneuron classes in oscillation generation. Fruitful prospects for such work can be found in recent advances in juxtacellular recordings, permitting the monitoring of a single interneuron *in vivo*. To take a relevant case, Lapray et al. (2012) discovered that PV basket cells—providing inhibition to the pyramidal cell body and proximal dendrite—fire preferentially at the descending  $\theta$  phase (findings reproduced by Varga et al., 2012), while ivy cells—providing inhibitory currents onto pyramidal cell dendrites—fire preferentially during ascension and at the trough. These studies reveal that during a single  $\theta$  cycle the inhibitory power onto distinct pyramidal cell sectors varies systematically (see also Brandon et al., 2014).

Viewing cell assemblies as the fundamental unit of computation rather than single neurons can by now be justified in that assemblies can tolerate noise by not being redirected in their trajectory, unlike single or small clusters of neurons (which would also be effected by spike transmission failures), intensifying the justification for placing such assemblies at the center of the Basic Label model. Given the information chunking and feature merging roles attributed to  $\gamma$  cycles, Buzsáki suggests that episodes of  $\gamma$  oscillations, which contain strings of cell assemblies, “may be regarded as a neural word” (2010, p. 365); that is, a discrete unit of information. If induced  $\gamma$  is also responsible for constructing coherent conceptual objects by synchronizing neural discharges binding together distant brain regions, as proposed by Tallon-Baudry and Bertrand (1999), then oscillations may also be responsible for complex semantic phenomena like copredication, through which a single object or event can be conceptualized via simultaneously concatenated yet contradictory properties, e.g., *The newspaper I held this morning has gone bust* or *Lunch was delicious but took forever* (see Murphy, Forthcoming). Brain rhythms would consequently play a crucial role in constructing what Aristotle termed the “place of forms.”

Topics in electrophysiology should also direct the concerns of those investigating the brain dynamics of linguistic computation. Certain areas of recent research appear to be more commensurable with elementary computational operations than others. For instance, transfer of charges across

membranes of all brain structures leads to a current giving rise to an extracellular field, which in turn influences the membranes. The transmembrane voltage ( $V_m$ ) is defined as the difference between the intracellular ( $V_i$ ) and extracellular voltage ( $V_e$ ) at a time  $t$  and location  $x$ :  $V_m(x, t) = V_i(x, t) - V_e(x, t)$ . A topic of contemporary debate is whether this endogenous field with its spatiotemporal  $V_e$ -fluctuations changes neuronal functions through ephaptic coupling (see Jefferys, 1995 for an overview). This process amounts to a feedback mechanism through which the neural structures producing a given field are in turn affected by them, yielding a self-generated cyclic loop. In terms of range, ephaptic coupling influences structures ranging from synapses to discrete neurons to neural networks.

At the microscale, a linear relationship is seen between a chemical synaptic current  $I_{syn}$  and  $V_m$ , with such current being able to be described as  $I_{syn}(t) = g_{syn}(t)(V_m(t) - E_{rev})$ , where  $g_{syn}$  is the synaptic conductance and  $E_{rev}$  is the reverse current. Following the above self-generated model,  $V_e$  changes alter synaptic currents. In addition, ephaptic coupling of  $V_m$  to electric fields influences spiking due to its effect on active cell conductances (Anastassiou et al., 2011). The explanatory force of ephaptic coupling becomes clearer with parallel plate whole-slice stimulation, which has shown that emergent properties of networks are more sensitive to electric fields than discrete neurons (Deans et al., 2007). As noted by Anastassiou and Koch (2015), the entrainment of spiking to field strengths as minimal as 0.5 mV/mm suggests that ephaptic entrainment to endogenous fields contributes to brain rhythms. Stronger ephaptic feedback also occurs after slower (<8 Hz) waves such as  $\theta$  and  $\delta$  compared to faster  $\gamma$  waves, suggesting that the non-synaptic electrical signals seen in ephaptic coupling contribute to neural computation.

As with ephaptic coupling, I would additionally like to propose cross-frequency coupling (CFC) as a core component of computation, as discussed above. It has been suggested that this generic operation coordinates spatiotemporal neural dynamics (Canolty and Knight, 2010; Lisman and Jensen, 2013), resolving a long-standing problem over how neural activity is synchronized. With larger neuronal populations oscillating at lower frequencies and smaller populations doing so at higher frequencies, CFC would enable their synchronization. In particular, it has been shown that via “phase-amplitude” CFC the phase of the lower frequency modulates the amplitude of the higher frequency component, a process claimed to be involved in information transfer for faculties such as memory (Tort et al., 2009, though see Aru et al., 2015 for current limitations of phase-amplitude modeling).

But while much is known about the biophysical substrates of individual frequency components, the cellular mechanisms behind frequency *interactions*—the origin of linguistic computation in the Basic Label model—remain opaque. Initial research leading to such an account has already been mentioned: Recall Korotkova et al. (2010) and their findings regarding hippocampal  $\theta \bullet \gamma$  coupling and its reliance on NMDA receptor-mediated PV interneuron excitation (see also Bi and Poo, 1998; Tort et al., 2008). Using laminar electrodes to



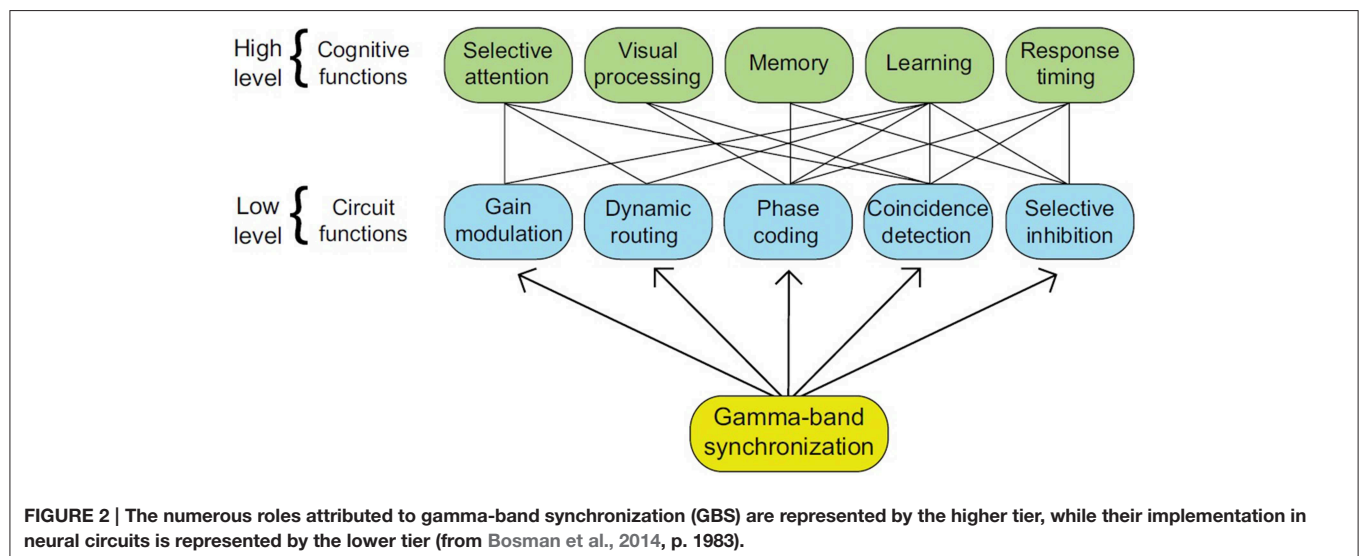
measure activity in monkey primate visual cortex, Spaak et al. (2012) found that  $\alpha$  phase in infragranular layers modulates  $\gamma$  amplitude in supergranular layers (see also Friston, 2008); similar to how thalamic nuclei oscillating at the  $\alpha$  band synchronize distant cortical regions oscillating at higher frequencies. As Aru et al. (2015) note, the most elegant theory to account for these findings is that periodic membrane potential fluctuations generate low frequency oscillations which subsequently gate the incidence of higher frequency activity in a phase-specific fashion. From a functional perspective, the above nested  $\gamma$  cycles could act as multiplexing mechanisms (Buzsáki, 2006, p. 356) for sustaining working memory representations by sending multiple representations as a single complex message to be recovered and “unpacked” downstream (see Hyafil et al., 2015 for empirical support, and Baddeley et al., 2014 for a review of working memory mechanisms); precisely as is seen in labeling and phasal transfer.

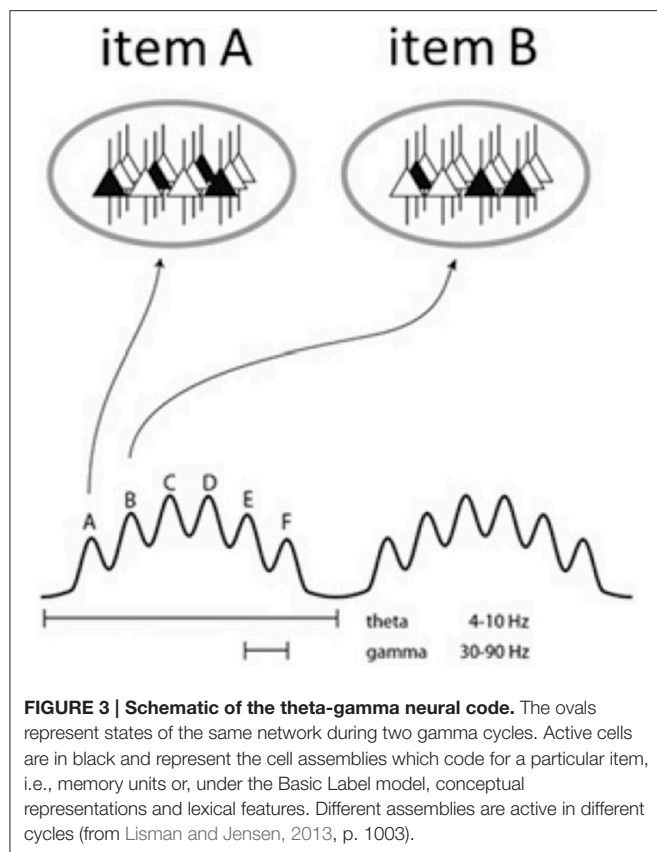
At a more general level, the cognome must operate within certain fundamental constraints on neuronal dynamics, such as the free-energy principle (following seminal insights from Friston, 2010) through which the homeostatic brain minimizes the dispersion (entropy) of interoceptive and exteroceptive states. If entropy is the average of “surprise” over time, then the brain will choose appropriate sensations to minimize surprise, and in so doing “the brain is implicitly maximizing the evidence for its own existence” (Bastos et al., 2012, p. 702); a notion not too far removed from Vaas’s assessment that the brain is “a self-referential, closed system, a functional reality emulator that constructs the world, rather than reconstruct it” (Vaas, 2001, p. 88). This form of “predictive coding” conforms to the free-energy principle and the image of the brain as a *constructive organ*, assembling and inferring linguistic representations. Studies of chaotic itinerancy (Tsuda, 2013, 2015), many-body physics and thermodynamics (Vitiello, 2015) may also prove indispensable in describing the high-dimensional state space of cortical activity implicated in computation (see the essays collected in Ohira and Uzawa, 2015 for discussion).

An emerging consensus regarding the validity of the communication-through-coherence (CTC) hypothesis lends further impetus to the claim that rhythms bring about linguistic computation (Bastos et al., 2015). CTC claims that rhythmic synchronization, especially in the  $\beta$  and  $\gamma$  bands, modulates the efficacy of anatomical connections, and that oscillations are necessary for long-distance assembly formation (König et al., 1995; Fries et al., 2008). CTC can be complemented with recent developments in the understanding of the functional role brain rhythms play, with assembly formation being the core operation at the connectome level necessary to establish the kinds of cross-modular representational structures seen in natural language (Lopes-dos-Santos et al., 2011).  $\gamma$  band activity, for instance, has been associated with numerous cognitive functions such as memory and selective attention (see **Figure 2** for examples of connectome-cognition links). With  $\gamma$  bands arising from an interplay of inhibition (produced by GABAergic neurons) and excitation (produced by glutamergic neurons), Bosman et al. (2014) propose that these bands have their origin in basic functional motifs conferring an advantage for low-level system processing and multiple cognitive functions (see also Bartos et al., 2007; Buzsáki and Wang, 2012).

The broad functionality of  $\gamma$  makes it an ideal candidate, along with the thalamus (discussed below), for being the conductor of language’s cross-modularity. The role of GBS in visual feature integration (Bosman et al., 2009), for instance, makes it a prime candidate for carrying out the forms of conceptual assimilation seen in any number of semantic phenomena. If linguistic computations are in fact responsible for this cross-modularity, then language can perhaps be more closely aligned to dominant descriptions of consciousness and working memory (Dehaene et al., 2014), even if we are forced to remain “virtually mute” (Chomsky, 1998, p. 440) about the nature of experiential content (Strawson, 2008, 2010).

In addition, GBS has been shown to support certain low-level functions in the hippocampus which may be vital to particular cognitive functions attributed to this region, such





as memory encoding and retrieval (Bosman et al., 2014). As mentioned, the hippocampus is the site of  $\gamma$ - $\theta$  coupling in that multiple  $\gamma$  waves are typically embedded within a single  $\theta$  cycle (Bragin et al., 1995). Along with the standard phase locking operation through which higher waves occur at stable phases in cycles of lower waves (Belluscio et al., 2012), this allows spike coordination and may consequently be partly responsible for low-level dynamical operations like phase coding (see Figure 3). As Lisman and Jensen (2013) review, the dual  $\gamma$  and  $\theta$  oscillations form a code for representing multiple items in an ordered way. Since each  $\theta$  cycle contains four to eight nested  $\gamma$  cycles, different forms of spatial information (such as a series of events from short-term memory, constituting an “episode”) can be represented and sequentially coordinated within a given cycle. This may in turn constrain the number of lexical items or features able to be transferred in a given phase. Through the coding scheme discussed by Lisman and Jensen, the cell assembly that fires during a given  $\gamma$  cycle forms a topographic pattern representing a particular item from memory. If this oscillatory mechanism is also responsible for syntactic computation, this would lend weight to the strong connection drawn in Murphy (2015b) between syntactic phases and episodic memory. The number of  $\gamma$  cycles able to be embedded within a  $\theta$  cycle may also be the reason why working memory is limited to its classic constraint of  $7 \pm 2$  (Kamifski et al., 2011). Roux and Uhlhaas (2014) make the related claim that oscillatory activity assures the maintenance of

working memory information. This explanation is of precisely the kind of granularity linguists should seek to capture syntactic operations like labeling, which involves storing conceptual roots in memory. In brief, and returning to issues outlined above, if intrinsic coupling across cortical oscillations is responsible for the hierarchical combination of computations at the syllabic and phonemic levels, “restoring the natural arrangement of phonemes within syllables” (Hyafil et al., 2015), then this leads to the possibility that hierarchical syntactic computations result from similar mechanisms.

These operations are all conserved from early in mammalian evolution, with the above interplay between excitation and inhibition being found in crustaceans (Nusbaum and Beenhakker, 2002) and major phyla dating back 350 million years (Katz and Harris-Warrick, 1999). Bosman et al. (2014) draw on such considerations in claiming that the evolutionary acquisition of this excitation-inhibition interplay led to the selection of these  $\gamma$  waves as a principal element of computation. If this GBS mechanism was a “direct, inevitable consequence of early circuitry organization” (Bosman et al., 2014, p. 1994), then it may be that it is an exaptation (being co-opted) in that it was later afforded a functional role in systems of memory and learning (see also Gould and Vrba, 1982). Further, top-down neocortical processes implicated in particular higher cognitive faculties like working memory (Buschman and Miller, 2007) and free-choice reach (Pesaran et al., 2008) also appear to be carried by interareal synchrony in the  $\beta$  rhythm (Bressler and Richter, 2015), increasing the electrophysiological validity of the functional roles attributed to this wave above.

### Cognomic Constraints and their Neurobiological Realizability

In the same way that  $\gamma$  oscillations “arise simultaneously and inevitably with inhibitory-excitatory interplay, and are neither an epiphenomenon nor a separate cause of the functionality beyond the underlying circuits” (Bosman et al., 2014, p. 1995), I would like to suggest that “linguistic” “computations” (which, as discussed, are neither purely linguistic nor thoroughly computational) are to be seen as identical to the operations of the connectome, which can be described in electrophysiological terms at the dynamical level and in still more abstract terms at the cognome level, in a similar way that *heat* and *energy* can be reduced to thermodynamics. While I hope to have shown that distinct oscillatory phases segregate discrete units of information (visual, olfactory, semantic, etc.), there remains the possibility that they also serve computations spanning multiple oscillatory cycles. Oscillatory phases may be the means through which different lexical features (e.g.,  $\phi$ , tense) are processed or time-locked with other features, leading to agreement relations, the resolution of filler-gap dependencies, feature inheritance/copying, and other familiar syntactic operations. Multiple  $\beta$  or  $\theta$  cycles could, for instance, employ dynamical operations like “cycle skipping” (Brandon et al., 2013) to control which cell assemblies are activated upon subsequent cycles to trigger different aspects of lexical and conceptual representations.

These remarks cover some basic computations, but what of their *constraints*? Consider Wurmbrand’s (2014) Merge Condition, stated below:

- (2) *Merge Condition*:  
Merge  $\alpha$  and  $\beta$  if  $\alpha$  can value a feature of  $\beta$ .

This condition ensures that set-formation via concatenation is licensed only under Agree, requiring also feature valuation. Leaving aside further details and the possibility that Merge applies freely, the scientist concerned with establishing linking hypotheses between linguistics and neuroscience is faced here with a number of challenges but also some surprising possibilities. For example, the cell assemblies implicated via cycle skipping in the features of  $\alpha$  and  $\beta$  may undergo phase-locking, leading to oscillatory synchronization of two discrete units of information. When this occurs, feature valuation takes place and the derivation can converge. If this process is barred in virtue of rhythmic coupling restrictions and the limits of assembly synchronization, feature valuation, and hence concatenation, does not take place. If the distribution of unvalued features,  $[uF]$ , also contributes to the demarcation of phases (Narita, 2014a), then the dynamics of feature valuation would likely align closely with the present Basic Label account of Transfer, since valuation, Agree and other copy-forming operations such as Internal Merge apply as a fundamental part of Transfer. Notice that this model at once implies specific neurobiological limitations, in that the hypothetical coupling responsible for feature valuation should occur after the cross-cortical  $\{\alpha(\gamma)\}$  embedding proposed to be the substrate of set-formation. This leads to clear, causally-addressable empirical predictions, to be investigated in future research.

As a secondary concern, I will assume that feature valuation (along with feature inheritance and Agree) are both cases of a more generalized Search operation, which forms relations between identical feature complexes (Ohta et al., 2013; Kato et al., 2014). Kato et al. (2014) even go as far as claiming that Search is in turn just an instance of Merge, and that the human language faculty may reduce to pure Merge. The Basic Label model and Kato et al. consequently yield different predictions about the *dynome*. From here, the matter is purely empirical, but these subtleties in distinct *cognome*-*dynome* hypotheses are yet to be investigated and are potentially of substantial interest to dynamic *cognomics*.

We are now in a position to outline a concrete research program. The first phase of dynamic *cognomics* will involve the above ongoing research into translating or reconstructing the

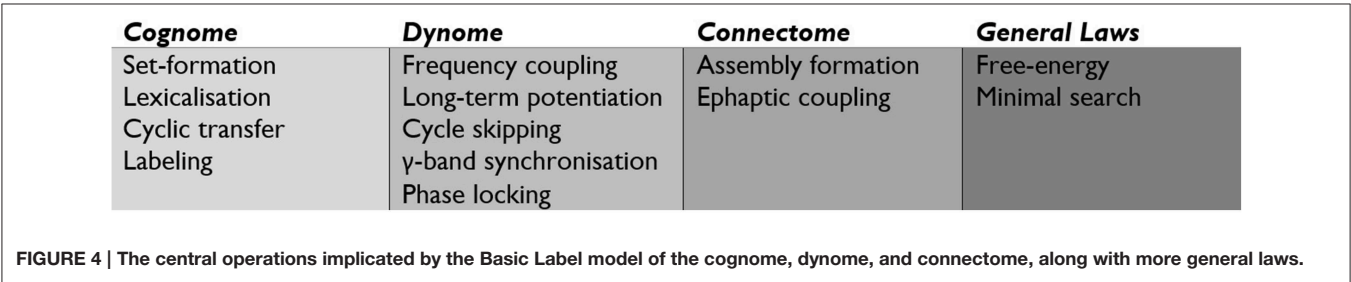
*operations* of syntax into oscillation terms. The second phase should center on translating the *constraints* of syntax, such as those concerning agreement, movement, and anti-locality. For instance, Richard’s (2010) Distinctness Condition, prohibiting the presence of multiple lexical units of the same label within a single phase complement, may be the consequence of how many distinct rhythms it is possible to couple in specific actions (Boeckx, 2013). These *\*XX*-like structures (e.g., structures containing multiple phase-internal nouns such as *\*John Mary ate apples*) may be ungrammatical because of the oscillatory patterns local language regions can sustain. These constraints may form the backdrop of what Narita (2014a, p. 26) identifies as a core aspect of minimal computation, the “Minimal Workspace” through which the construction and transferring of syntactic structures takes place. To put it more concretely, language-external systems (interfaces) may only be able to sustain a single rhythm from the  $\gamma$  and  $\beta$  bands due to the small size of localized regions, and would hence be incapable of interpreting multiple category-identical elements in a single cycle. The phase/non-phase rhythm of syntactic computation would thus arise from the limits of oscillatory sustainability, and the connection between syntactic phases and oscillatory phases becomes more than purely orthographic:  $[C [T v[V D/n [N]]]]$  emerges from  $[\beta [\gamma \beta [\gamma \beta [\gamma]]]]$  given the labeling role attributed to  $\beta$  above, which in turn explains *\*XX* violations. Narita’s (2014b) *\*{t,t}* constraint, which prohibits the transfer of syntactic objects whose two members are both traces/copies of movement, also strikes me as amenable to a similar, if not identical explanation. Objects of the  $\{t,t\}$  kind cannot be labeled, as in (3), and are hence illicit (Moro, 2006, p. 15):

- (3) *\*[which picture of the wall]<sub>i</sub> do you think that [the cause of the riot]<sub>j</sub> was  $\{t_i,t_j\}$ ?*

What is needed is consequently a re-conceptualization of language as not only a system of thought, planning and interpretation, but also a system of oscillatory and electrophysiological information synchronization. The computational constraints explored by Wurmbrand and others can direct inquiry into the possibilities of dynamic operations, although this process may require further elaboration of the nature of the role of oscillations in cognition.

Globularity and Cortico-centrism

Recent developments in systems neuroscience have identified large scale distributed brain networks, typically explored through fMRI and MEG (Brookes et al., 2012). Data from fMRI suggests



that the implication of a functionally specific set of neurons in any given computation is assisted by a backdrop of large-scale neural assembly inter-communication. These networks are composed of sub-networks with correlating and anti-correlating patterns, leading to a situation in which a single large-scale network may operate through overlapping but distinct neural sub-networks. **Figure 4** highlights the major operations at the level of the cognome, dynome, and connectome, along with general laws influencing such operations.

As the cognome-dynome-connectome linking hypotheses expand, it is important not to ignore the fundamental role of the genome. Consider briefly the genes *RUNX2*, the *DLX* suite and the *BMP* family, involved in skull and brain development (Perdomo-Sabotal et al., 2014). In a series of ongoing research, Boeckx and Benítez-Burraco (2014a,b, Benítez-Burraco and Boeckx, 2015) hypothesize that a modification in this gene network gave rise to a more “globular” head shape (relative to Neanderthals/Denisovans; Bruner, 2004; Gunz et al., 2012; Theofanopoulou, 2015)—approaching a level of sphericity unseen in our closest ancestors—and the consequent re-wiring of cortical and sub-cortical structures, permitting the construction of the forms of cross-modular representations well established in psychological, philosophical, and semantic theories of concepts (Spelke, 2010; Pietroski, Forthcoming). Globularity may also have contributed, as some have suggested, to an increase in wiring efficiency across the brain (Chklovskii et al., 2002). It is of outstanding interest for biolinguistics and dynamic cognomics that functional links of this kind are beginning to be drawn between genes and their cellular consequences for the human cognitive phenotype.

An evaluation of these observations can also be made alongside a consideration of what Piattelli-Palmarini and Uriagereka (2008) see as the optimizing role language has in building syntactic and phonological structures, which proceeds via minimal search and related principles of computational efficiency (Larson, 2015). This minimalist perspective leads to a separation of optimality from language’s proposed “function” of mapping structures to the interfaces, since similar optimizing principles are found elsewhere in the natural world, leading Piattelli-Palmarini and Uriagereka (2008, p. 209) to “suspect that the process behind the abstract form follow[s] from physico-chemical invariants.” But lacking a theory of brain dynamics, the authors are unable to ground these general proposals within any neurobiological framework. I suggest that the microcellular level and the dynome, operating within some general physical laws of neural organization such as free-energy, can provide a potential substrate of such “physico-chemical invariants.” The only human-unique aspect of the model pursued here, then, is the *context* in which the conserved and universal rhythms discussed above perform their operations of coupling and decoupling; namely, a globular brain case, which would have led to a decrease in the types of “spatial inequalities” (Salami et al., 2003) between cortical and subcortical regions which would prohibit long-distance coupling. This would imply that the numerous centuries-long approaches to human-uniqueness, ranging from philosophy to medicine, have approached the matter from the wrong perspective. Instead of asking what it is about humans which allows us to form complex systems of

symbolic interpretation, we should instead ask what it is about other animals which *prohibits* them from doing so.

Globularity may also have led to the expansion of the neo-cortex and the pulvinar, spurred on by the reduction of the large Neanderthal visual system (Pearce et al., 2013). As Benítez-Burraco and Boeckx (2015) point out, cross-modular concepts likely employ thalamic nuclei such as the pulvinar and the medio-dorsal nucleus, not least because of the thalamus’s role in modulating fronto-parietal activity, regulating cortical oscillations (Saalman et al., 2012) and enhancing the rhythmic range of different frequency bands (Singer, 2013). Controlling rhythmic behavior is also a function attributed to *RUNX2* (Reale et al., 2013, see also van der Lely and Pinker, 2014 for genetic discussion relating to phonological computations). A literature review leads Theofanopoulou and Boeckx (Forthcoming a,b) to claim that the thalamus is the brain region which tunes the oscillations of other subcortical structures (see also Boeckx, 2014b). The importance of the thalamus for higher cognition was also speculated in work by Campion and Elliot-Smith (1934), rejecting the dominant cortico-centrism and suggesting that cortico-thalamic impulse circulation was responsible for “thought.”

Relatedly, due to the few protein differences between humans and chimpanzees, the individuating computational factors may be attributed to cis- and trans-regulatory genes (Somel et al., 2013). Hominid-unique features which may have led to the higher mental faculties of humans include novel neuronal cell types and the duplication of developmental proteins such as *SRGAP2*, leading to unique dendritic spine density and form (Geschwind and Rakic, 2013). Synaptic and dendritic maturation also occurs in humans for a considerably longer time than in non-humans (Bianchi et al., 2013). If we also consider the conclusions of Harris’s review of cortical computation in mammals and birds, that the “human cortex appears to contain the same cell types, and their patterns of wiring and gene expression appear basically similar to well-studied model systems” (2015, p. 3184), the importance of subcortical investigations into linguistic computation becomes even clearer. While subcortical structures have often been derided as the “reptilian brain,” responsible for only primitive drives, far removed from the neocortex’s higher echelons of thought, the perspective of dynamic cognomics re-situates subcortical regions like the thalamus and the basal ganglia into the core areas responsible for linguistic phrase structure building (see also Johnson and Knight, 2015 for evidence that the thalamus plays a key role in neocortical oscillations involved in memory processes).

Summarizing these findings, it appears that the developed interneurons and dendritic spinal strength proposed by Geschwind and Rakic (2013) fortified long-distance assembly connections and, in turn, the mechanisms of ephaptic coupling, CFC and other neuronal processes (operating within the confines of the CTC hypothesis) necessary for the rhythmic interactions claimed above to be the source of computations like labeling and cyclic transfer. The targeting of the perisomatic region of pyramidal neurons by inhibitory interneurons in particular leads to the formation of  $\gamma$  rhythms and their concomitant properties of conceptual assimilation. Though many intervening



neurochemical processes need to be accounted for and explained, it seems that such processes, along with novel  $V_e$ -fluctuations, are the reason why we find the cyclic short-term memory storage capacities seen in labeling. Updating Darwin's claim of "He who understands baboon would do more toward metaphysics than Locke," we can conclude that he who understands brain rhythms would do more toward biolinguistics than Lenneberg.

## References

- Adger, D. (2013). *A Syntax of Substance*. Cambridge, MA: MIT Press.
- Allen, K., and Monyer, H. (2015). Interneuron control of hippocampal oscillations. *Curr. Opin. Neurobiol.* 31, 81–87. doi: 10.1016/j.conb.2014.08.016
- Amundson, R. (1998). Typology reconsidered: two doctrines on the history of evolutionary biology. *Biol. Philos.* 13, 153–177. doi: 10.1023/A:1006599002775
- Amundson, R. (2006). EvoDevo as cognitive psychology. *Biol. Theory* 1, 10–11. doi: 10.1162/biot.2006.1.1.10
- Anastassiou, C. A., and Koch, C. (2015). Ephaptic coupling to endogenous electric field activity: why bother? *Curr. Opin. Neurobiol.* 31, 95–103. doi: 10.1016/j.conb.2014.09.002
- Anastassiou, C. A., Perin, R., Markram, H., and Koch, C. (2011). Ephaptic coupling of cortical neurons. *Nat. Neurosci.* 14, 217–223. doi: 10.1038/nn.2727
- Aru, J., Aru, J., Priesemann, V., Wibral, M., Lana, L., Pipa, G., et al. (2015). Untangling cross-frequency coupling in neuroscience. *Curr. Opin. Neurobiol.* 31, 51–61. doi: 10.1016/j.conb.2014.08.002
- Baddeley, A., Eysenck, M. W., and Anderson, A. C. (2014). *Memory, 2nd Edn.* Abingdon: Psychology Press.
- Balari, S., Boeckx, C., and Lorenzo, G. (2012). On the feasibility of biolinguistics: Koster's word-based challenge and our 'natural computation' alternative. *Biolinguistics* 6, 205–221.
- Balari, S., and Lorenzo, G. (2013). *Computational Phenotypes: Towards an Evolutionary Developmental Biolinguistics*. Oxford: Oxford University Press.
- Bartos, M., Vida, I., and Jonas, P. (2007). Synaptic mechanisms of synchronized gamma oscillations in inhibitory interneuron networks. *Nat. Rev. Neurosci.* 8, 45–56. doi: 10.1038/nrn2044
- Bastiaansen, M., and Hagoort, P. (2015). Frequency-based segregation of syntactic and semantic unification during online sentence level language comprehension. *J. Cogn. Neurosci.* doi: 10.1162/jocn\_a\_00829. [Epub ahead of print].
- Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., and Friston, K. J. (2012). Canonical microcircuits for predictive coding. *Neuron* 76, 695–711. doi: 10.1016/j.neuron.2012.10.038
- Bastos, A. M., Vezoli, J., and Fries, P. (2015). Communication through coherence with inter-areal delays. *Curr. Opin. Neurobiol.* 31, 173–180. doi: 10.1016/j.conb.2014.11.001
- Belluscio, M. A., Mizuseki, K., Schmidt, R., Kempter, R., and Buzsáki, G. (2012). Cross-frequency phase-phase coupling between  $\theta$  and  $\gamma$  oscillations in the hippocampus. *J. Neurosci.* 32, 423–435. doi: 10.1523/JNEUROSCI.4122-11.2012
- Bemis, D. K., and Pykkänen, L. (2013). Basic linguistic composition recruits the left anterior temporal lobe and left angular gyrus during both listening and reading. *Cereb. Cortex* 23, 1859–1873. doi: 10.1093/cercor/bhs170
- Benítez-Burraco, A., and Boeckx, C. (2014a). FOXP2, retinoic acid, and language: a promising direction. *Front. Cell. Neurosci.* 8:387. doi: 10.3389/fncel.2014.00387
- Benítez-Burraco, A., and Boeckx, C. (2014b). Universal grammar and biological variation: an evo-devo agenda for comparative biolinguistics. *Biol. Theory* 9, 122–134. doi: 10.1007/s13752-014-0164-0
- Benítez-Burraco, A., and Boeckx, C. (2015). Possible functional links among brain- and skull-related genes selected in modern humans. *Front. Psychol.* 6:794. doi: 10.3389/fpsyg.2015.00794
- Berger, H. (1929). Über das elektroencephalogramm des menschen. *Archiv Psychiatr. Nervenkr.* 87, 527–570. doi: 10.1007/BF01797193
- Bertossa, R. (2011). Morphology and behaviour: functional links in development and evolution. *Philos. Trans. R. Soc. B* 366, 2056–2068. doi: 10.1098/rstb.2011.0035
- Berwick, R. (2011). "All you need is merge: biology, computation, and language from the bottom up," in *The Biolinguistic Enterprise: New Perspectives on the Evolution and Nature of the Human Language Faculty*, eds A. M. Di Sciullo and C. Boeckx (Oxford: Oxford University Press), 461–491.
- Bi, G. Q., and Poo, M. M. (1998). Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *J. Neurosci.* 18, 10464–10472.
- Bianchi, S., Stimpson, C. D., Duka, T., Larsen, M. D., Janssen, W. G., Collins, Z., et al. (2013). Synaptogenesis and development of pyramidal neuron dendritic morphology in the chimpanzee neocortex resembles humans. *Proc. Natl. Acad. Sci. U.S.A.* 110(Suppl. 2), 10395–10401. doi: 10.1073/pnas.1301224110
- Boeckx, C. (2013). Merge: biolinguistic considerations. *Engl. Linguist.* 30, 463–484.
- Boeckx, C. (2014a). *Elementary Syntactic Structures: Prospects of a Feature-Free Syntax*. Cambridge: Cambridge University Press.
- Boeckx, C. (2014b). Our brain's language-readiness. *El País*.
- Boeckx, C., and Benítez-Burraco, A. (2014a). The shape of the human language-ready brain. *Front. Psychol.* 5:282. doi: 10.3389/fpsyg.2014.00282
- Boeckx, C., and Benítez-Burraco, A. (2014b). Globularity and language-readiness: generating new predictions by expanding the set of genes of interest. *Front. Psychol.* 5:1324. doi: 10.3389/fpsyg.2014.01324
- Boeckx, C., and Theofanopoulou, C. (2014). A multidimensional interdisciplinary framework for linguistics: the lexicon as a case study. *J. Cogn. Sci.* 15, 403–420.
- Boeckx, C., and Theofanopoulou, C. (2015). Language, cognomes, and the challenges of building cognitive phylogenies. *Front. Psychol.* 6:784. doi: 10.3389/fpsyg.2015.00784
- Bolker, J. A. (2008). Developing a history of evo–devo. *BioScience* 58, 461–463. doi: 10.1641/B580515
- Bosman, C. A., Lansink, C. S., and Pennartz, C. M. A. (2014). Functions of gamma-band synchronization in cognition: from single circuits to functional diversity across cortical and subcortical systems. *Eur. J. Neurosci.* 39, 1982–1999. doi: 10.1111/ejn.12606
- Bosman, C. A., Womelsdorf, T., Desimone, R., and Fries, P. (2009). A microsaccadic rhythm modulates gamma-band synchronization and behavior. *J. Neurosci.* 29, 9471–9480. doi: 10.1523/JNEUROSCI.1193-09.2009
- Bouchard, K. E., and Chang, E. F. (2014). Neural decoding of spoken vowels from human sensory-motor cortex with high-density electrocorticography. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 2014, 6782–6785. doi: 10.1109/embsc.2014.6945185
- Bragin, A., Jandó, G., Nádasdy, Z., Hetke, J., Wise, K., and Buzsáki, G. (1995). Gamma (40–100 Hz) oscillation in the hippocampus of the behaving rat. *J. Neurosci.* 15, 47–60.
- Brandon, M. P., Bogaard, A. R., Schultheiss, N. W., and Hasselmo, M. E. (2013). Segregation of cortical head direction cell assemblies on alternating  $\theta$  cycles. *Nat. Neurosci.* 16, 739–748. doi: 10.1038/nn.3383
- Brandon, M. P., Koenig, J., Leutgeb, J. K., and Leutgeb, S. (2014). New and distinct hippocampal place codes are generated in a new environment during septal inactivation. *Neuron* 82, 789–796. doi: 10.1016/j.neuron.2014.04.013
- Bressler, S. L., and Menon, V. (2010). Large-scale brain networks in cognition: emerging methods and principles. *Trends Cogn. Sci.* 14, 277–290. doi: 10.1016/j.tics.2010.04.004
- Bressler, S. L., and Richter, C. G. (2015). Interareal oscillatory synchronization in top-down neocortical processing. *Curr. Opin. Neurobiol.* 31, 62–66. doi: 10.1016/j.conb.2014.08.010
- Brookes, M. J., Liddle, E. B., Hale, J. R., Woolrich, M. W., Luckhoo, H., Liddle, P. F., et al. (2012). Task induced modulation of neural oscillations

## Acknowledgments

My thanks go primarily to Cedric Boeckx for his persistent encouragement, and to Karl Friston, Alec Marantz, Jyrki Tuomainen, and Wing-Yee Chow for helpful discussion. This work was supported by an Economic and Social Research Council scholarship (1474910).

- in electrophysiological brain networks. *Neuroimage* 63, 1918–1930. doi: 10.1016/j.neuroimage.2012.08.012
- Bruner, E. (2004). Geometric morphometrics and paleoneurology: brain shape evolution in the genus homo. *J. Hum. Evol.* 47, 279–303. doi: 10.1016/j.jhevol.2004.03.009
- Buflin, E., and Carbonell, E. (2004). Are symbolic behavior and neuroplasticity an example of gene-culture evolution? *Rev. Neurol.* 39, 48–55.
- Buschman, T. J., and Miller, E. K. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science* 315, 1860–1862. doi: 10.1126/science.1138071
- Buzsáki, G. (2006). *Rhythms of the Brain*. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780195301069.001.0001
- Buzsáki, G. (2010). Neural syntax: cell assemblies, synapse ensembles, and readers. *Neuron* 68, 362–385. doi: 10.1016/j.neuron.2010.09.023
- Buzsáki, G., and Freeman, W. (2015). Editorial overview: brain rhythms and dynamic coordination. *Curr. Opin. Neurobiol.* 31, v–ix. doi: 10.1016/j.conb.2015.01.016
- Buzsáki, G., Logothetis, N., and Singer, W. (2013). Scaling brain size, keeping timing: evolutionary preservation of brain rhythms. *Neuron* 80, 751–764. doi: 10.1016/j.neuron.2013.10.002
- Buzsáki, G., and Wang, X. -J. (2012). Mechanisms of gamma oscillations. *Annu. Rev. Neurosci.* 35, 203–225. doi: 10.1146/annurev-neuro-062111-150444
- Calabrese, A., and Woolley, S. M. N. (2015). Coding principles of the canonical cortical microcircuit in the avian brain. *Proc. Natl. Acad. Sci. U.S.A.* 112, 3517–3522. doi: 10.1073/pnas.1408545112
- Campion, G., and Elliot-Smith, G. (1934). *The Neural Basis of Thought*. New York, NY: Harcourt Brace Jovanovich.
- Cannon, J., McCarthy, M. M., Lee, S., Lee, J., Börgers, C., Whittington, M. A., et al. (2014). Neurosystems: brain rhythms and cognitive processing. *Eur. J. Neurosci.* 39, 705–719. doi: 10.1111/ejn.12453
- Canolty, R. T., and Knight, R. T. (2010). The functional role of cross-frequency coupling. *Trends Cogn. Sci.* 14, 506–515. doi: 10.1016/j.tics.2010.09.001
- Carroll, S. B. (2006). *The Making of the Fittest: DNA and the Ultimate Forensic Record of Evolution*. New York, NY: W.W. Norton.
- Caton, R. (1875). The electric currents of the brain. *Br. Med. J.* 2, 278.
- Cervenka, M. C., Boatman-Reich, D. F., Ward, J., Franaszczuk, P. J., and Crone, N. E. (2011). Language mapping in multilingual patients: electrocorticography and cortical stimulation during naming. *Front. Hum. Neurosci.* 5:13. doi: 10.3389/fnhum.2011.00013
- Chesi, C. (2015). On the directionality of phrase structure building. *J. Psycholinguist. Res.* 44, 65–89. doi: 10.1007/s10936-014-9330-6
- Chklovskii, D. B., Schikorski, T., and Stevens, C. F. (2002). Wiring optimization in cortical circuits. *Neuron* 34, 341–347. doi: 10.1016/S0896-6273(02)00679-7
- Chomsky, N. (1968). Quine's empirical assumptions. *Synthese* 19, 53–68. doi: 10.1007/BF00568049
- Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, N. (1998). Comments: Galen Strawson, mental reality. *Philos. Phenomenol. Res.* 58, 437–441. doi: 10.2307/2653523
- Chomsky, N. (2000). *New Horizons in the Study of Language and Mind*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511811937
- Chomsky, N. (2008). "On phases," in *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*, eds R. Freidin, C. P. Otero, and M. L. Zubizarreta (Cambridge, MA: MIT Press), 133–166.
- Chomsky, N. (2012). *The Science of Language: Interviews with James McGilvray*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9781139061018
- Chomsky, N. (2015a). "Problems of projection: extensions," in *Structures, Strategies and Beyond: Studies in Honour of Adriana Belletti*, eds E. Di Domenico, C. Hamann, and S. Matteini (Amsterdam: John Benjamins), 1–16.
- Chomsky, N. (2015b). Some core contested concepts. *J. Psycholinguist. Res.* 44, 91–104. doi: 10.1007/s10936-014-9331-5
- Chow, B. Y., Han, X., Dobry, A. S., Qian, X., Chuong, A. S., Li, M., et al. (2010). High-performance genetically targetable optical neural silencing by light-driven proton pumps. *Nature* 463, 98–102. doi: 10.1038/nature08652
- Collins, J. (2015). "Naturalism without metaphysics," in *Experimental Philosophy, Rationalism, and Naturalism*, eds E. Fischer and J. Collins (London: Routledge), 85–109.
- Crunelli, V., David, F., Lörincz, M. L., and Hughes, S. W. (2015). The thalamocortical network as a single slow wave-generating unit. *Curr. Opin. Neurobiol.* 31, 72–80. doi: 10.1016/j.conb.2014.09.001
- Dawkins, R. (2006). *Climbing Mount Improbable*. Oxford: Oxford University Press.
- Dean, H. L., Hagan, M. A., and Pesaran, B. (2012). Only coherent spiking in posterior parietal cortex coordinates looking and reaching. *Neuron* 73, 829–841. doi: 10.1016/j.neuron.2011.12.035
- Deans, J. K., Powell, A. D., and Jefferys, J. G. R. (2007). Sensitivity of coherent oscillations in rat hippocampus to AC electric fields. *J. Physiol.* 583, 555–565. doi: 10.1113/jphysiol.2007.137711
- Dehaene, S., Charles, L., King, J.-R., and Marti, S. (2014). Toward a computational theory of conscious processing. *Curr. Opin. Neurobiol.* 25, 76–84. doi: 10.1016/j.conb.2013.12.005
- Deisz, R. A., and Prince, D. A. (1989). Frequency-dependent depression of inhibition in guinea-pig neocortex *in vitro* by GABAB receptor feed-back on GABA release. *J. Physiol.* 412, 513–541. doi: 10.1113/jphysiol.1989.sp017629
- de Pasquale, F., Della Penna, S., Snyder, A. Z., Marzetti, L., Pizzella, V., Luca Romani, G., et al. (2012). A cortical core for dynamic integration of functional networks in the resting human brain. *Neuron* 74, 753–764. doi: 10.1016/j.neuron.2012.03.031
- Diller, K. C., and Cann, R. L. (2013). "Genetics, evolution, and the innateness of language," in *The Evolutionary Emergence of Language: Evidence and Inference*, eds R. Botha and M. Everaert (Oxford: Oxford University Press), 244–258.
- Embick, D., and Poeppel, D. (2015). Towards a computational(ist) neurobiology of language: correlational, integrated and explanatory neurolinguistics. *Lang. Cogn. Neurosci.* 30, 357–366. doi: 10.1080/23273798.2014.980750
- Engel, A. K., and Fries, P. (2010). Beta-band oscillations - signalling the status quo? *Curr. Opin. Neurobiol.* 20, 156–165. doi: 10.1016/j.conb.2010.02.015
- Field, J. (2004). *Psycholinguistics: The Key Concepts*. London: Routledge.
- Fitch, W. T. (2009). Prolegomena to a future science of biolinguistics. *Biolinguistics* 3, 283–320.
- Freeman, W. J. (2015). Mechanism and significance of global coherence in scalp EEG. *Curr. Opin. Neurobiol.* 31, 199–205. doi: 10.1016/j.conb.2014.11.008
- Friederici, A. D. (2012). The cortical language circuit: from auditory perception to sentence comprehension. *Trends Cogn. Sci.* 5, 262–268. doi: 10.1016/j.tics.2012.04.001
- Friederici, A. D., Bahlmann, J., Heim, S., Schubotz, R. I., and Anwander, A. (2006). The brain differentiates human and non-human grammars: functional localization and structural connectivity. *Proc. Natl. Acad. Sci. U.S.A.* 103, 2458–2463. doi: 10.1073/pnas.0509389103
- Fries, P. (2009). Neuronal gamma-band synchronization as a fundamental process in cortical computation. *Annu. Rev. Neurosci.* 32, 209–224. doi: 10.1146/annurev-neuro.051508.135603
- Fries, P., Womelsdorf, T., Oostenveld, R., and Desimone, R. (2008). The effects of visual stimulation and selective visual attention on rhythmic neuronal synchronization in macaque area V4. *J. Neurosci.* 28, 4823–4835. doi: 10.1523/JNEUROSCI.4499-07.2008
- Friston, K. (2008). Hierarchical models in the brain. *PLoS Comput. Biol.* 4:e1000211. doi: 10.1371/journal.pcbi.1000211
- Friston, K. (2010). The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11, 127–138. doi: 10.1038/nrn2787
- Gallistel, C. R., and King, A. P. (2009). *Memory and the Computational Brain: Why Cognitive Science will Transform Neuroscience*. Malden, MA: Wiley-Blackwell. doi: 10.1002/9781444310498
- Geswind, D. H., and Rakic, P. (2013). Cortical evolution: judge the brain by its cover. *Neuron* 80, 633–647. doi: 10.1016/j.neuron.2013.10.045
- Giraud, A.-L., and Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nat. Neurosci.* 15, 511–517. doi: 10.1038/nn.3063
- Gould, S. J., and Vrba, E. S. (1982). Exaptation—a missing term in the science of form. *Paleobiology* 8, 4–15.
- Gunz, P., Neubauer, S., Golovanova, L., Doronichev, V., Maureille, B., and Hublin, J.-J. (2012). A uniquely modern human pattern of endocranial development: insights from a new cranial reconstruction of the Neandertal newborn from Mezmaiskaya. *J. Hum. Evol.* 62, 300–313. doi: 10.1016/j.jhevol.2011.11.013
- Hagoort, P. (2014). Nodes and networks in the neural architecture for language: Broca's region and beyond. *Curr. Opin. Neurobiol.* 28, 136–141. doi: 10.1016/j.conb.2014.07.013

- Harris, K. D. (2015). Cortical computation in mammals and birds. *Proc. Natl. Acad. Sci. U.S.A.* 112, 3184–3185. doi: 10.1073/pnas.1502209112
- Hauser, M. D., Yang, C., Berwick, R. C., Tattersall, I., Ryan, M. J., Watumull, J., et al. (2014). The mystery of language evolution. *Front. Psychol.* 5:401. doi: 10.3389/fpsyg.2014.00401
- Hebb, D. O. (1949). *The Organization of Behavior*. New York, NY: John Wiley and Sons.
- Honkanen, R., Rouhinen, S., Wang, S. H., Palva, J. M., and Palva, S. (2014). Gamma oscillations underlie the maintenance of feature-specific information and the contents of visual working memory. *Cereb. Cortex*. doi: 10.1093/cercor/bhu263. [Epub ahead of print].
- Hornstein, N. (2009). *A Theory of Syntax: Minimal Operations and Universal Grammar*. Cambridge: Cambridge University Press.
- Hornstein, N., and Pietroski, P. (2009). Basic operations: minimal syntax-semantics. *Catalan J. Linguist.* 8, 113–139.
- Hyafil, A., Fontolan, L., Kabdebon, C., Gutkin, B., and Giraud, A.-L. (2015). Speech encoding by coupled cortical theta and gamma oscillations. *eLife* 10, 7554. doi: 10.7554/eLife.06213
- Hyvärinen, A., and Oja, E. (2000). Independent component analysis: algorithms and applications. *Neural Netw.* 13, 411–430. doi: 10.1016/S0893-6080(00)00026-5
- Jefferys, J. G. (1995). Nonsynaptic modulation of neuronal activity in the brain: electric currents and extracellular ions. *Physiol. Rev.* 75, 689–723.
- Jensen, O., Gelfand, J., Kounios, J., and Lisman, J. E. (2002). Oscillations in the alpha band (9–12 Hz) increase with memory load during retention in a short-term memory task. *Cereb. Cortex* 12, 877–882. doi: 10.1093/cercor/12.8.877
- Johnson, E. L., and Knight, R. T. (2015). Intracranial recordings and human memory. *Curr. Opin. Neurobiol.* 31, 18–25. doi: 10.1016/j.conb.2014.07.021
- Jouen, A.-L., Verwey, W. B., van der Helden, J., Scheiber, C., Neveu, R., Dominey, P. F., et al. (2013). Discrete sequence production with and without a pause: the role of cortex, basal ganglia, and cerebellum. *Front. Hum. Neurosci.* 7:492. doi: 10.3389/fnhum.2013.00492
- Kaiser, G. (2010). *A Friendly Guide to Wavelets*. New York, NY: Springer.
- Kamiński, J., Brzezicka, A., and Wróbel, A. (2011). Short-term memory capacity ( $7 \pm 2$ ) predicted by theta to gamma cycle length ratio. *Neurobiol. Learn. Mem.* 95, 19–23. doi: 10.1016/j.nlm.2010.10.001
- Kato, T., Kuno, M., Narita, H., Zushi, M., and Fukui, N. (2014). Generalized search and cyclic derivation by phase. *Sophia Linguist.* 61, 203–222.
- Katz, P. S., and Harris-Warrick, R. M. (1999). The evolution of neuronal circuits underlying species-specific behavior. *Curr. Opin. Neurobiol.* 9, 628–633. doi: 10.1016/S0959-4388(99)00012-4
- Kayser, C., Wilson, C., Safaai, H., Sakata, S., and Panzeri, S. (2014). Rhythmic auditory cortex activity at multiple timescales. *J. Neurosci.* 35, 7750–7762. doi: 10.1523/JNEUROSCI.0268-15.2015
- Kinzel, K. D., and Spelke, E. S. (2007). Core systems in human cognition. *Prog. Brain Res.* 164, 257–264. doi: 10.1016/S0079-6123(07)64014-X
- Klausberger, T., Marton, L. F., O'Neill, J., Huck, J. H., Dalezios, Y., Fuentealba, P., et al. (2005). Complementary roles of cholecystokinin- and parvalbumin-expressing GABAergic neurons in hippocampal network oscillations. *J. Neurosci.* 25, 9782–9793. doi: 10.1523/JNEUROSCI.3269-05.2005
- Klimesch, W., Sauseng, P., and Hanslmayr, S. (2007). EEG alpha oscillations: the inhibition/timing hypothesis. *Brain Res. Rev.* 53, 63–88. doi: 10.1016/j.brainresrev.2006.06.003
- Klimesch, W., Schack, B., Schabus, M., Doppelmayr, M., Gruber, W., and Sauseng, P. (2004). Phase-locked alpha and theta oscillations generate the P1-N1 complex and are related to memory performance. *Cogn. Brain Res.* 19, 302–316. doi: 10.1016/j.cogbrainres.2003.11.016
- Koch, C. (1999). *Biophysics of Computation: Information Processing in Single Neurons*. Oxford: Oxford University Press.
- König, P. (1994). A method for the quantification of synchrony and oscillatory properties of neuronal. *J. Neurosci. Methods* 54, 31–37. doi: 10.1016/0165-0270(94)90157-0
- König, P., Engel, A. K., and Singer, W. (1995). Relation between oscillatory activity and long-range synchronization in cat visual cortex. *Proc. Natl. Acad. Sci. U.S.A.* 92, 290–294. doi: 10.1073/pnas.92.1.290
- Kopell, N. J., Gritton, H. J., Whittington, M. A., and Kramer, M. A. (2014). Beyond the connectome: the dynamo. *Neuron* 83, 1319–1328. doi: 10.1016/j.neuron.2014.08.016
- Korotkova, T., Fuchs, E. C., Ponomarenko, A., von Engelhardt, J., and Monyer, H. (2010). NMDA receptor ablation on parvalbumin-positive interneurons impairs hippocampal synchrony, spatial representations, and working memory. *Neuron* 68, 557–569. doi: 10.1016/j.neuron.2010.09.017
- Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., and Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science* 320, 110–113. doi: 10.1126/science.1154735
- Lapray, D., Laszotz, B., Lagler, M., Viney, T. J., Katona, L., Valenti, O., et al. (2012). Behavior-dependent specialization of identified hippocampal interneurons. *Nat. Neurosci.* 15, 1265–1271. doi: 10.1038/nn.3176
- Larson, B. (2015). Minimal search as a restriction on merge. *Lingua* 156, 57–69. doi: 10.1016/j.lingua.2014.12.009
- Larson-Prior, L. J., Oostenveld, R., Della Penna, S., Michalareas, G., Prior, F., Babajani-Feremi, A., et al. (2013). Adding dynamics to the human connectome project with MEG. *Neuroimage* 80, 190–201. doi: 10.1016/j.neuroimage.2013.05.056
- Lau, E. F., Phillips, C., and Poeppel, D. (2008). A cortical network for semantics: (de)constructing the N400. *Nat. Rev. Neurosci.* 9, 920–933. doi: 10.1038/nrn2532
- Lenneberg, E. H. (1964). “A biological perspective of language,” in *New Directions in the Study of Language*, ed E. H. Lenneberg (Cambridge, MA: MIT Press), 65–88.
- Leventhal, D. K., Gage, G. J., Schmidt, R., Pettibone, J. R., Case, A. C., and Berke, J. D. (2012). Basal ganglia beta oscillations accompany cue utilization. *Neuron* 73, 523–536. doi: 10.1016/j.neuron.2011.11.032
- Lewis, S., and Phillips, C. (2015). Aligning grammatical theories and language processing models. *J. Psycholinguist. Res.* 44, 27–46. doi: 10.1007/s10936-014-9329-z
- Lieberman, P. (2006). *Toward an Evolutionary Biology of Language*. Cambridge, MA: Harvard University Press.
- Lieberman, P. (2015). Language did not spring forth 100,000 years ago. *PLoS Biol.* 13:e1002064. doi: 10.1371/journal.pbio.1002064
- Lisman, J. E., and Jensen, O. (2013). The theta–gamma neural code. *Neuron* 77, 1002–1016. doi: 10.1016/j.neuron.2013.03.007
- Liu, Z., Fukunaga, M., de Zwart, J. A., and Duyn, J. H. (2010). Large-scale spontaneous fluctuations and correlations in brain electrical activity observed with magnetoencephalography. *Neuroimage* 51, 102–111. doi: 10.1016/j.neuroimage.2010.01.092
- Lopes-dos-Santos, V., Conde-Ocazone, S., Nicoletis, M. A. L., Ribeiro, S. T., and Tort, A. B. L. (2011). Neuronal assembly detection and cell membership specification by principal component analysis. *PLoS ONE* 6:e20996. doi: 10.1371/journal.pone.0020996
- Luck, S. J. (2014). *An Introduction to the Event-Related Potential Technique, 2nd Edn*. Cambridge, MA: MIT Press.
- Maini, P. K. (2004). The impact of Turing’s work on pattern formation in biology. *Math. Today* 40, 140–141.
- Makeig, S., Westerfield, M., Jung, T. P., Enghoff, S., Townsend, J., Courchesne, E., et al. (2002). Dynamic brain sources of visual evoked responses. *Science* 295, 690–694. doi: 10.1126/science.1066168
- Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. New York, NY: Freeman.
- Martin, C., and Ravel, N. (2014). Beta and gamma oscillatory activities associated with olfactory memory tasks: different rhythms for different functional networks? *Front. Behav. Neurosci.* 8:218. doi: 10.3389/fnbeh.2014.00218
- Martins, P. T., and Boeckx, C. (2014). Attention mechanisms and the mosaic evolution of speech. *Front. Psychol.* 5:1463. doi: 10.3389/fpsyg.2014.01463
- McCormick, D. A., McGinley, M. J., and Salkoff, D. B. (2015). Brain state dependent activity in the cortex and thalamus. *Curr. Opin. Neurobiol.* 31, 133–140. doi: 10.1016/j.conb.2014.10.003
- McGivray, J. (2013). “The philosophical foundations of biolinguistics,” in *The Cambridge Handbook of Biolinguistics*, eds C. Boeckx and K. Grohmann (Cambridge: Cambridge University Press), 22–46.
- Moro, A. (2006). “Copular sentences,” in *The Blackwell Companion to Syntax II*, eds M. Everaert and H. van Riemsdijk (Oxford: Blackwell), 1–23.



- Müller, G. B. (2008). "EvoDevo as a discipline," in *Evolving Pathways: Key Themes in Evolutionary Developmental Biology*, eds A. Minelli and G. Fusco (Cambridge: Cambridge University Press), 3–29.
- Murphy, E. (Forthcoming). *Neural Substrates of Copredication: When an Unstoppable Scan Meets an Impossible Object*. York Papers in Linguistics.
- Murphy, E. (2015a). Labels, cognomes and cyclic computation: an ethological perspective. *Front. Psychol.* 6:715. doi: 10.3389/fpsyg.2015.00715
- Murphy, E. (2015b). Reference, phases and individuation: Topics at the labeling-interpretive interface. *Opticon* 17, 1–13. doi: 10.5334/opt.cn
- Muthukumaraswamy, S. D., Edden, R. A. E., Jones, D. K., Swettenham, J. B., and Singh, K. D. (2009). Resting GABA concentration predicts peak gamma frequency and fMRI amplitude in response to visual stimulation in humans. *Proc. Natl. Acad. Sci. U.S.A.* 106, 8356–8361. doi: 10.1073/pnas.0900728106
- Narita, H. (2014a). *Endocentric Structuring of Projection-free Syntax*. Amsterdam: John Benjamins.
- Narita, H. (2014b). "{\*t,t}" in *Poster presented at the 32nd West-Coast Conference on Formal Linguistics (WCCFL 32)*, California: University of Southern California.
- Narita, H., and Fujita, K. (2010). A naturalist reconstruction of minimalist and evolutionary biolinguistics. *Biolinguistics* 4, 356–376.
- Neeleman, A. (2013). Comments on Pullum. *Mind Lang.* 28, 522–531. doi: 10.1111/mila.12030
- Nóbrega, V. A., and Miyagawa, S. (2015). The precedence of syntax in the rapid emergence of human language in evolution as defined by the integration hypothesis. *Front. Psychol.* 6:271. doi: 10.3389/fpsyg.2015.00271
- Nosarti, C., Rushe, T. M., Woodruff, P. W. R., Stewart, A. L., Rifkin, L., and Murray, R. M. (2004). Corpus callosum size and very preterm birth: relationship to neuropsychological outcome. *Brain* 127, 2080–2089. doi: 10.1093/brain/awh230
- Nusbaum, M. P., and Beenhakker, M. P. (2002). A small-systems approach to motor pattern generation. *Nature* 417, 343–350. doi: 10.1038/417343a
- Ohira, T., and Uzawa, T. (eds.). (2015). *Mathematical Approaches to Biological Systems: Networks, Oscillations, and Collective Motions*. New York, NY: Springer.
- Ohta, S., Fukui, N., and Sakai, K. (2013). Syntactic computation in the human brain: the degree of merger as a key factor. *PLoS ONE* 8:e56230. doi: 10.1371/journal.pone.0056230
- Ossandón, T., Jerbi, K., Vidal, J. R., Bayle, D. J., Henaff, M. A., Jung, J., et al. (2011). Transient suppression of broadband gamma power in the default-mode network is correlated with task complexity and subject performance. *J. Neurosci.* 31, 14521–14530. doi: 10.1523/JNEUROSCI.2483-11.2011
- Parker, S. T., and McKinney, M. L., (1999). *Origins of Intelligence: The Evolution of Cognitive Development in Monkeys, Apes, and Humans*. Baltimore, MD: Johns Hopkins University Press.
- Pearce, E., Stringer, C., and Dunbar, R. I. M. (2013). New insights into differences in brain organization between Neanderthals and anatomically modern humans. *Proc. R. Soc. B.* 280:20130168. doi: 10.1098/rspb.2013.0168
- Perdomo-Sabotal, A., Kanton, S., Walter, M. B., and Nowick, K. (2014). The role of gene regulatory factors in the evolutionary history of humans. *Curr. Opin. Genet. Dev.* 29C, 60–67. doi: 10.1016/j.gde.2014.08.007
- Pesaran, B., Nelson, M. J., and Andersen, R. A. (2008). Free choice activates a decision circuit between frontal and parietal cortex. *Nature* 453, 406–409. doi: 10.1038/nature06849
- Petersson, K. M., Folia, V., and Hagoort, P. (2012). What artificial grammar learning reveals about the neurobiology of syntax. *Brain Lang.* 120, 83–95. doi: 10.1016/j.bandl.2010.08.003
- Pfenning, A. R., Hara, E., Whitney, O., Rivas, M. V., Wang, R., Roulhac, P. L., et al. (2014). Convergent transcriptional specializations in the brains of humans and song-learning birds. *Science* 346:1256846. doi: 10.1126/science.1256846
- Piattelli-Palmarini, M., and Uriagereka, Y. (2008). Still a bridge too far? Biolinguistic questions for grounding language on brains. *Phys. Life Rev.* 5, 207–224. doi: 10.1016/j.plrev.2008.07.002
- Piattelli-Palmarini, M., and Vitiello, G. (2015). Linguistics and some aspects of its underlying dynamics. *arXiv*:1506.08663.
- Pietroski, P. (Forthcoming). *Conjoining Meanings: Semantics without Truth Values*. Oxford: Oxford University Press.
- Pigliucci, M., and Müller, G. B. (2010). "Elements of an extended evolutionary synthesis," in *Evolution – The Extended Synthesis*, eds M. Pigliucci and G. B. Müller (Cambridge, MA: MIT Press), 3–17.
- Poeppel, D. (1996). Neurobiology and linguistics are not yet unifiable. *Behav. Brain Sci.* 19, 642–643. doi: 10.1017/S0140525X00043430
- Poeppel, D. (2011). Genetics and language: a neurobiological perspective on the missing link(-ing hypothesis). *J. Neurodev. Disord.* 3, 381–387. doi: 10.1007/s11689-011-9097-0
- Poeppel, D. (2012). The maps problem and the mapping problem: two challenges for a cognitive neuroscience of speech and language. *Cogn. Neuropsychol.* 29, 34–55. doi: 10.1080/02643294.2012.710600
- Poeppel, D. (2014). The neuroanatomic and neurophysiological infrastructure for speech and language. *Curr. Opin. Neurobiol.* 28C, 142–149. doi: 10.1016/j.conb.2014.07.005
- Poeppel, D., and Embick, D. (2005). "Defining the relation between linguistics and neuroscience," in *Twenty-First Century Psycholinguistics: Four Cornerstones*, ed A. Cutler (Mahwah, NJ: Lawrence Erlbaum), 103–118.
- Principe, J. C., and Brockmeier, A. J. (2015). Representing and decomposing neural potential signals. *Curr. Opin. Neurobiol.* 31, 13–17. doi: 10.1016/j.conb.2014.07.023
- Pulvermüller, F., Preissl, H., Eulitz, C., Pantev, C., Lutzenberger, W., Elbert, T., et al. (1994). Brain rhythms, cell assemblies and cognition: evidence from the processing of words and pseudowords. *Psychology* 5, 1–30.
- Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., and Shulman, G. L. (2001). A default mode of brain function. *Proc. Natl. Acad. Sci. U.S.A.* 98, 676–682. doi: 10.1073/pnas.98.2.676
- Rakic, P., and Kornack, D. R. (2001). "Neocortical expansion and elaboration during primate evolution," in *Evolutionary Anatomy of the Primate Cerebral Cortex*, eds D. Falk and K. R. Gibson (Cambridge: Cambridge University Press), 30–56.
- Reale, M. E., Webb, I. C., Wang, X., Baltazar, R. M., Coolen, L. M., and Lehman, M. N. (2013). The transcription factor Runx2 is under circadian control in the suprachiasmatic nucleus and functions in the control of rhythmic behavior. *PLoS ONE* 8:e54317. doi: 10.1371/journal.pone.0054317
- Reimann, M. W., Muller, E. B., Ramaswamy, S., and Markram, H. (2015). An algorithm to predict the connectome of neural microcircuits. *Front. Neural Circuits* 9:28. doi: 10.3389/fncir.2015.00028
- Richards, M. (2011). Deriving the edge: what's in a phase? *Syntax* 14, 74–95. doi: 10.1111/j.1467-9612.2010.00146.x
- Richards, N. (2010). *Uttering Trees*. Cambridge, MA: MIT Press. doi: 10.7551/mitpress/9780262013765.001.0001
- Roux, F., and Uhlhaas, P. J. (2014). Working memory and neural oscillations: alpha-gamma versus theta-gamma codes for distinct WM information? *Trends Cogn. Sci.* 18, 16–25. doi: 10.1016/j.tics.2013.10.010
- Royer, S., Zemelman, B. V., Losonczy, A., Kim, J., Chance, F., Magee, J. C., et al. (2012). Control of timing, rate and bursts of hippocampal place cells by dendritic and somatic inhibition. *Nat. Neurosci.* 15, 769–775. doi: 10.1038/nn.3077
- Saalmann, Y. B., Pinsk, M. A., Wang, L., Li, X., and Kastner, S. (2012). The pulvinar regulates information transmission between cortical areas based on attention demands. *Science* 337, 753–756. doi: 10.1126/science.1223082
- Salami, M., Itami, K., Tsumoto, T., and Kimura, F. (2003). Change of conduction velocity by regional myelination yields constant latency irrespective of distance between thalamus and cortex. *Proc. Nat. Acad. Sci.* 100, 6174–6179. doi: 10.1073/pnas.0937380100
- Santi, A., Friederici, A. D., Makuuchi, M., and Grodzinsky, Y. (2015). An fMRI study dissociating distance measures computed by Broca's area in movement processing: clause boundary vs. identity. *Front. Psychol.* 6:654. doi: 10.3389/fpsyg.2015.00654
- Santi, A., and Grodzinsky, Y. (2010). fMRI adaptation dissociates syntactic complexity dimensions. *Neuroimage* 51, 1285–1293. doi: 10.1016/j.neuroimage.2010.03.034
- Sauseng, P., Klimesch, W., Gruber, W. R., Hanslmayr, S., Freunberger, R., and Doppelmayr, M. (2007). Are event-related potential components generated by phase resetting of brain oscillations? A critical discussion. *Neuroscience* 146, 1435–1444. doi: 10.1016/j.neuroscience.2007.03.014
- Schoenemann, P. T. (2012). Evolution of brain and language. *Prog. Brain Res.* 195, 443–459. doi: 10.1016/B978-0-444-53860-4.00022-2



- Seung, S. (2012). *Connectome: How the Brain's Wiring Makes Us Who We Are*. Boston, MA: Houghton Mifflin Harcourt.
- Singer, W. (2013). Cortical dynamics revisited. *Trends Cogn. Sci.* 17, 616–626. doi: 10.1016/j.tics.2013.09.006
- Somel, M., Liu, X., and Khaitovich, P. (2013). Human brain evolution: transcripts, metabolites and their regulators. *Nat. Rev. Neurosci.* 14, 112–127. doi: 10.1038/nrn3372
- Somogyi, P., and Klausberger, T. (2005). Defined types of cortical interneurone structure space and spike timing in the hippocampus. *J. Physiol.* 562, 9–26. doi: 10.1113/jphysiol.2004.078915
- Spaak, E., Bonnefond, M., Maier, A., Leopold, D. A., and Jensen, O. (2012). Layer-specific entrainment of gamma-band neural activity by the alpha rhythm in monkey visual cortex. *Curr. Biol.* 22, 2313–2318. doi: 10.1016/j.cub.2012.10.020
- Spelke, E. (2010). “Innateness, choice, and language,” in *Chomsky Notebook*, eds J. Bricmont, and J. Franck (New York, NY: Columbia University Press), 203–210.
- Sporns, O. (2013). The human connectome: origins and challenges. *Neuroimage* 80, 53–61. doi: 10.1016/j.neuroimage.2013.03.023
- Strauss, A., Henry, M. J., Scharinger, M., and Obleser, J. (2015). Alpha phase determines successful lexical decision in noise. *J. Neurosci.* 35, 3256–3262. doi: 10.1523/JNEUROSCI.3357-14.2015
- Strawson, G. (2008). *Real Materialism and Other Essays*. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199267422.001.0001
- Strawson, G. (2010). *Mental Reality, 2nd Edn*. Cambridge, MA: MIT Press.
- Swaab, T. Y., LeDoux, K., Camblin, C. C., and Boudewyn, M. A. (2012). “Language-related ERP components,” in *The Oxford Handbook of Event Related Potential Components*, eds S. J. Luck and E. S. Kappenman (New York, NY: Oxford University Press), 397–439.
- Szathmáry, E. (1996). From RNA to language. *Curr. Biol.* 6, 764. doi: 10.1016/S0960-9822(02)00585-7
- Tallon-Baudry, C., and Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends Cogn. Sci.* 3, 151–162. doi: 10.1016/S1364-6613(99)01299-1
- Tass, P. A. (2000). Stochastic phase resetting: a theory for deep brain stimulation. *Prog. Theor. Phys. Suppl.* 139, 301–313. doi: 10.1143/PTPS.139.301
- Tettamanti, M., and Weniger, D. (2006). Broca's area: a supramodal hierarchical processor? *Cortex* 42, 491–494. doi: 10.1016/S0010-9452(08)70384-8
- Theofanopoulou, C. (2015). Brain asymmetry in the white matter making and globularity. *Front. Psychol.* 6:1355. doi: 10.3389/fpsyg.2015.01355
- Theofanopoulou, C., and Boeckx, C. (Forthcoming a). “The central role of the thalamus in language and cognition,” in *Advances in Biolinguistics: The Human Language Faculty and its Biological Basis*, eds C. Boeckx and K. Fujita (London: Routledge).
- Theofanopoulou, C., and Boeckx, C. (Forthcoming b). “(Neural) syntax,” in *Language, Syntax, and the Natural Sciences*, eds R. Martin and A. Gallego (Cambridge: Cambridge University Press).
- Tort, A. B., Komorowski, R. W., Manns, J. R., Kopell, N. J., and Eichenbaum, H. (2009). Theta-gamma coupling increases during the learning of item-context associations. *Proc. Natl. Acad. Sci. U.S.A.* 106, 20942–20947. doi: 10.1073/pnas.0911331106
- Tort, A. B., Kramer, M. A., Thorn, C., Gibson, D. J., Kubota, Y., Graybiel, A. M., et al. (2008). Dynamic cross-frequency couplings of local field potential oscillations in rat striatum and hippocampus during performance of a T-maze task. *Proc. Natl. Acad. Sci. U.S.A.* 105, 20517–20522. doi: 10.1073/pnas.0810524105
- Tsuda, I. (2013). Chaotic itinerancy. *Scholarpedia* 8:4459. doi: 10.4249/scholarpedia.4459
- Tsuda, I. (2015). Chaotic itinerancy and its roles in cognitive neurodynamics. *Curr. Opin. Neurobiol.* 31, 67–71. doi: 10.1016/j.conb.2014.08.011
- Turing, A. M. (1952). The chemical basis of morphogenesis. *Philos. Trans. R. Soc. Lond. B* 237, 37–73. doi: 10.1098/rstb.1952.0012
- Turken, A. U., and Dronkers, N. F. (2011). The neural architecture of the language comprehension network: converging evidence from lesion and connectivity analyses. *Front. Syst. Neurosci.* 5:1. doi: 10.3389/fnsys.2011.00001
- Uriagereka, J. (2012). *Spell-out and the Minimalist Program*. Oxford: Oxford University Press.
- Vaas, R. (2001). It binds, therefore I am! Review of Rodolfo Llinás's *I of the Vortex*. *J. Conscious. Stud.* 8, 85–88.
- van der Lely, H. K., and Pinker, S. (2014). The biological basis of language: insight from developmental grammatical impairments. *Trends Cogn. Sci.* 18, 586–595. doi: 10.1016/j.tics.2014.07.001
- Varga, C., Golshani, P., and Soltesz, I. (2012). Frequency-invariant temporal ordering of interneuronal discharges during hippocampal oscillations in awake mice. *Proc. Natl. Acad. Sci. U.S.A.* 109, E2726–E2734. doi: 10.1073/pnas.1210929109
- Vitiello, G. (2015). The use of many-body physics and thermodynamics to describe the dynamics of rhythmic generators in sensory cortices engaged in memory and learning. *Curr. Opin. Neurobiol.* 31, 7–12. doi: 10.1016/j.conb.2014.07.017
- Viventi, J., Kim, D. H., Vigeland, L., Frechette, E. S., Blanco, J. A., Kim, Y. S., et al. (2011). Flexible, foldable, actively multiplexed, high-density electrode array for mapping brain activity *in vivo*. *Nat. Neurosci.* 14, 1599–1605. doi: 10.1038/nn.2973
- Westerlund, M., and Pykkänen, L. (2014). The role of the left anterior temporal lobe in semantic composition vs. semantic memory. *Neuropsychologia* 57, 59–70. doi: 10.1016/j.neuropsychologia.2014.03.001
- Whittington, M. A., and Traub, R. D. (2003). Interneuron diversity series: inhibitory interneurons and network oscillations *in vitro*. *Trends Neurosci.* 26, 676–682. doi: 10.1016/j.tins.2003.09.016
- Wipf, D. P., Owen, J. P., Attias, H. T., Sekihara, K., and Nagarajan, S. S. (2010). Robust Bayesian estimation of the location, orientation, and time course of multiple correlated neural sources using MEG. *Neuroimage* 49, 641–655. doi: 10.1016/j.neuroimage.2009.06.083
- Wurmbrand, S. (2014). “The Merge condition: a syntactic approach to selection,” in *Minimalism and Beyond: Radicalizing the Interfaces*, eds L. Franks, T. Radeva-Bork, and L. Schürcks (Amsterdam: John Benjamins), 130–166.
- Zhang, Y. E., Landback, P., Vibration, M. D., and Long, M. (2011). Accelerated recruitment of new brain development genes into the human genome. *PLoS Biol.* 9:e1001179. doi: 10.1371/journal.pbio.1001179

**Conflict of Interest Statement:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Murphy. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# The “Globularization Hypothesis” of the Language-ready Brain as a Developmental Frame for Prosodic Bootstrapping Theories of Language Acquisition

Aritz Irurtzun \*

CNRS, IKER (UMR 5478), Bayonne, France

## OPEN ACCESS

### Edited by:

Cedric Boeckx,  
Catalan Institute for Research and  
Advanced Studies (ICREA) and  
Universitat de Barcelona, Spain

### Reviewed by:

Silvia Martínez Ferreiro,  
University of Barcelona and University  
of Groningen, Spain  
Monika T. Molnar,  
Basque Center on Cognition, Brain  
and Language, Spain

### \*Correspondence:

Aritz Irurtzun  
aritz.irurtzun@iker.cnrs.fr

### Specialty section:

This article was submitted to  
Language Sciences,  
a section of the journal  
Frontiers in Psychology

**Received:** 30 July 2015

**Accepted:** 10 November 2015

**Published:** 09 December 2015

### Citation:

Irurtzun A (2015) The “Globularization Hypothesis” of the Language-ready Brain as a Developmental Frame for Prosodic Bootstrapping Theories of Language Acquisition.  
*Front. Psychol.* 6:1817.  
doi: 10.3389/fpsyg.2015.01817

In recent research (Boeckx and Benítez-Burraco, 2014a,b) have advanced the hypothesis that our species-specific language-ready brain should be understood as the outcome of developmental changes that occurred in our species after the split from Neanderthals-Denisovans, which resulted in a more globular braincase configuration in comparison to our closest relatives, who had elongated endocrania. According to these authors, the development of a globular brain is an essential ingredient for the language faculty and in particular, it is the centrality occupied by the thalamus in a globular brain that allows its modulatory or regulatory role, essential for syntactico-semantic computations. Their hypothesis is that the syntactico-semantic capacities arise in humans as a consequence of a process of globularization, which significantly takes place postnatally (cf. Neubauer et al., 2010). In this paper, I show that Boeckx and Benítez-Burraco’s hypothesis makes an interesting developmental prediction regarding the path of language acquisition: it teases apart the onset of phonological acquisition and the onset of syntactic acquisition (the latter starting significantly later, after globularization). I argue that this hypothesis provides a developmental rationale for the prosodic bootstrapping hypothesis of language acquisition (cf. *i.a.* Gleitman and Wanner, 1982; Mehler et al., 1988, *et seq.*; Gervain and Werker, 2013), which claim that prosodic cues are employed for syntactic parsing. The literature converges in the observation that a large amount of such prosodic cues (in particular, rhythmic cues) are already acquired *before* the completion of the globularization phase, which paves the way for the premises of the prosodic bootstrapping hypothesis, allowing babies to have a rich knowledge of the prosody of their target language *before* they can start parsing the primary linguistic data syntactically.

**Keywords:** globularization, prosodic bootstrapping, language development, language acquisition, postnatal development

## 1. INTRODUCTION: THE GLOBULARIZATION HYPOTHESIS

According to a recent article in this journal by Boeckx and Benítez-Burraco (2014a), “much work in neurolinguistics has unintentionally emphasized the externalization component of language, since morpho-phonology is perhaps the easiest aspect to single out linguistic tasks, even if the word “syntax” was said to be the target of the relevant works. In so doing, work on neuroimaging biased

the results toward the Broca-Wernicke model, and all too quickly attributed “syntax” to Broca’s area.” In contrast, Boeckx and Benítez-Burraco (2014a,b) have advanced the hypothesis that our species-specific language-ready brain (a brain which is suited for acquiring natural languages) should be understood as the outcome of developmental changes that occurred in our species after the split from Neanderthals-Denisovans, and which resulted in a more globular braincase configuration in comparison to our closest relatives, who had elongated endocrania. They propose that even if factors like brain lateralization are important, the development of a globular brain is at the outset of our language faculty, and in particular, it is the centrality of the thalamus in a globular brain that allows its modulatory or regulatory role, essential for syntactico-semantic computations (cf. *i.a.* Wahl et al., 2008).

Importantly, globularization takes place postnatally (cf. Lieberman et al., 2002; Neubauer et al., 2010; Gunz et al., 2012), therefore, according to Boeckx and Benítez-Burraco’s hypothesis, even if innately specified, the combinatorial syntactic ability of humans is not innate *stricto sensu*, but the outcome of a postnatal developmental phase. After globularization a new brain configuration is obtained whereby the thalamus occupies a central position (and a central role). As Boeckx and Benítez-Burraco (2014a) put it, “a proper characterization of the language-ready brain that does not recognize a central role to the thalamus is unlikely to be correct, for it would miss the critical engagement of the thalamus in regulating cortical activity. By providing low-frequency oscillations capable of embedding higher-frequency oscillations across distant brain regions, the thalamus provides the crucial regulation needed to form the sort of meaningful cross-modular conceptual structures that are characteristic of language.”

In this paper I discuss the developmental dimension of Boeckx and Benítez-Burraco’s hypothesis and relate it to one of the most prominent hypotheses in early language acquisition studies; the “prosodic bootstrapping hypothesis” (cf. *i.a.* Mehler et al., 1988; Christophe et al., 2003; Bernard and Gervain, 2012; Gervain and Werker, 2013; Lungus and Nespor, 2013).

The argument is presented as follows: Section 2 gives a brief overview of the development of the ability for phonological discrimination in human infants (an essential prerequisite for the identification and acquisition of the prosodic patterns of the target language). Section 3 presents the basic tenets of the prosodic bootstrapping hypothesis (a hypothesis that claims that language-acquiring children use prosody as a guide for inferring the basic syntactic pattern of their target language). Last, Section 4 argues for a natural combination of the globularization hypothesis and the prosodic bootstrapping hypothesis. In a nutshell, the globularization hypothesis proposes that the ability for syntactic computations is not innate, but that it rather develops *after* the postnatal globularization phase. In contrast, as the studies of early phonological development show, babies a few months old have already a rich knowledge of the prosodic patterns of their target language. Therefore, and in line with the prosodic bootstrapping hypothesis, language-acquiring babies will be able to use their early-acquired prosodic knowledge as a guiding principle for inferring the syntax

of their target language the moment the syntactic ability develops.

## 2. EARLY PHONOLOGICAL ABILITIES IN HUMAN INFANTS

Some essential ingredients for language acquisition are already present at birth. Since the seventies, a wide range of studies have shown infants’ capacity for very early phonological parsing and discrimination (for an overview, see Panneton and Newman, 2012; Vihman, 2014). For instance, Eimas et al. (1971) found that infants as young as 1 month of age are able to discriminate the voice onset time (VOT) of synthetic stop consonants like /p/-/b/ in a manner approximating adult categorical perception. Similar results were obtained by Moffitt (1971) with 20- to 24-week-old infants in a study attesting the discrimination in place of articulation of different consonants. Given the limited exposure of newborn infants to speech, these results suggest that this categorical perception in a linguistic mode may be innate, and in the general debate on language nature vs. nurture, scholars such as J. Mehler have built upon these early capacities to argue for innatist “selectionist” theories of language learning whereby the baby “learns” her target language by “forgetting” others (cf. *i.a.* Mehler, 1974; Mehler and Dupoux, 1990).

What is more, the earliest fetal responses to auditory stimuli are reported at 19 weeks of gestation, long before the development of the fetal ear is complete (cf. Hepper and Shahidullah, 1994; Abdala and Keefe, 2012), and effects of very early auditory categorization have also been found *in utero*: a number of experiments have shown that third-trimester fetuses’ auditory experience can influence their postnatal auditory preferences: newborns tend to quiet in response to their mothers’ voice (and touch), Marx and Nagy (2015), and they also tend to prefer their mother’s voice over other female’s voice (cf. Mehler et al., 1978; DeCasper and Fifer, 1980; Fifer, 1981; Querleu et al., 1984; Spence and DeCasper, 1987; Ockleford et al., 1988; Hepper et al., 1993)<sup>1</sup>. Besides, as reported by DeCasper and Spence (1986), newborns also tend to be more reinforced by the audition of speech passages they heard *in utero* over passages they were not exposed to (and they can remember them for over a month; Granier-Deferre et al., 2011). Finally, Mampe et al. (2009) provide evidence that even the cry melodies of newborns of around 3 days of age are shaped in accordance with the intonational contours of the language they were exposed to prenatally (German vs. French). All this confirms evidence of a very early ability for the discrimination and memorization of complex sounds in newborn infants.

Regarding prosody and rhythm, there is ample evidence that newborns also have the ability for discrimination between inputs varying in different suprasegmental properties (see *i.a.* Morse, 1972; Olsho et al., 1982; Mehler et al., 1988; Karzon and Nicholas, 1989; Shahidullah and Hepper, 1994; Sansavini et al., 1997; Nazzi et al., 1998a,b; Carral et al., 2005). In

<sup>1</sup>However, infants do not seem to show preferences for their father’s voice (cf. Ward and Cooper, 1999).

particular, studies like Nazzi et al. (1998a) show that babies can discriminate between languages pertaining to different rhythmic classes [such as Japanese (mora-timed) or British English (stress-timed)] when exposed to low-pass filtered speech signals. The setting in this type of experiment shows that babies discriminate between rhythmic classes because by low-pass filtering (e.g., under 400 Hz) the speech signal, it gets a dramatic degradation of its phonemic content (i.e., the vast majority of its formant structure is removed), while it retains its rhythmic structure. Other studies employing this type of low-pass filtered stimuli (like Byers-Heinlein et al., 2010) provide evidence that language discrimination in neonates which were surrounded by a bilingual environment prenatally is robust, and that that language preference reflects previous listening experience (see also Gervain and Werker, 2013; Molnar et al., 2014a,b). Besides, other types of studies show that at 4 1/2 months babies tend to listen longer to speech samples that include prosodic pauses corresponding to syntactic units, as opposed to speech samples with pauses that break syntactic units (cf. Jusczyk and Nelson, 1996 and references therein).

All these results are to be framed in the fast (pre- and post-natal) development of the basic structures for sound discrimination in humans (whereby infants already possess an adult-like dedicated neuronal network for phonological processing at 3 months of age (cf. Dehaene-Lambertz and Baillet, 1998 as well as Peña et al., 2003; Dehaene-Lambertz et al., 2006 or Dubois et al., 2015)<sup>2</sup>.

Interestingly, however, early acoustic discrimination is not a human-specific ability, for it is also observed in a wide variety of other animals like guinea pigs (Vince, 1979), sheep (Vince et al., 1982) or chinchillas (Kuhl and Miller, 1975), and discrimination of languages of different prosodic types is also mastered by different species like cotton-top tamarins (cf. Ramus et al., 2000), or rats (cf. Toro et al., 2003).

Nonetheless, there is a growing amount of literature arguing that human infants go well beyond mere acoustic pattern-recognition and learning; evidence suggests that babies use the prosodic patterns of their target language in order to infer the syntactic structure underneath them in a sort of “reverse engineering.” That is, part of the knowledge obtained by babies from categorical perception is restricted to a specific area (say, learning of the vowel space or the consonantal inventory of the target language), but a subpart of the learning obtained with this innate capacity is more consequential: learning the tunes of the surrounding language helps the child making informed guesses about the syntactic structure of the language [this is so because the prosodic pattern of a language partially reflects the syntactic structure underneath (cf. Gussenhoven, 2004; Truckenbrodt, 2007; Selkirk, 2011)]. This is in a nutshell the proposal of the “prosodic bootstrapping hypothesis.”

<sup>2</sup>See Telkemeyer et al. (2009) for a near-infrared spectroscopy and EEG study showing that a right hemispheric lateralization for slow acoustic modulations (characteristic of prosodic features) is present at birth (see also Telkemeyer et al., 2011).

### 3. THE “PROSODIC BOOTSTRAPPING” HYPOTHESIS

Prosody and rhythm are essential ingredients of natural language (cf. *i.a.* Brentari, 1999; Gussenhoven, 2004; Pfau and Quer, 2010) and a growing number of scholars argue that they have a close connection with other aspects of human cognition like musical aesthetics and computation, or our mathematical abilities (cf. *i.a.* Rebuschat et al., 2011; Arbib, 2013; Asano and Boeckx, 2015)<sup>3</sup>. Current literature converges in the idea that beyond the early ability for prosodic discrimination, “prosodic segmentation abilities emerge crosslinguistically some time around 8 months” (Nazzi et al., 2006, p. 296).

The rationale under the rapid acquisition of prosody could be seen as emerging from the combination of the following two factors:

- (i) First, babies develop very early the necessary brain structures for adequately parsing acoustic inputs—and in particular human language inputs (see references above and Pang and Taylor, 2000, among others)—, and a growing number of works is emphasizing the natural “tuning up” between speech rhythm and endogenous oscillatory auditory cortical properties (cf. *i.a.* Drullman et al., 1994; Smith et al., 2002; Lakatos et al., 2005; Giraud et al., 2007; VanRullen and Dubois, 2011; Leong, 2012)<sup>4</sup>. In particular, neuronal oscillatory activity in the Theta band (3–7 Hz) is thought to track syllable patterns, whereas slower oscillations in the Delta band (1–3 Hz) track phrasal and intonational patterns (cf. Giraud and Poeppel, 2012; Peelle and Davis, 2012). Ghazanfar and Takahashi (2014) have argued that the same oscillatory cycles are present in macaques’ lip smacking, suggesting that “lip smacking may have been an ancestral expression linked to vocal output to produce the original rhythmic audiovisual speech-like utterances in the human lineage” (see also Fitch, 2013; Martins and Boeckx, 2014, for discussion, as well as Theofanopoulou, 2015, for a recent evo-devo hypothesis according to which the myelination of the Corpus Callosum, brain asymmetry and globularity “are conjectured to make up the angles of a co-evolutionary triangle that gave rise to our language-ready brain”).
- (ii) Second, children are exposed to a very particular input (infant-directed speech), which has very specific linguistic and paralinguistic properties. Following the traditional view, infant-directed speech has a set of hyperarticulated features that help the child develop her linguistic capacities and acquire her language. For instance, infant-directed speech is typically associated with an exaggerated pitch, which is covered with emotional prosody to capture the child’s attention (cf. *i.a.* Fernald, 1984; Cooper and Aslin, 1989; Fernald and Mazzie, 1991; Katz et al., 1996).

<sup>3</sup>See also Wang et al. (2015) for a proposal about the brain areas in charge of the human-specific ability for the integration of multiple features in abstract pattern learning.

<sup>4</sup>A reviewer rightly notes that these studies are performed with adult subjects, not with babies. Indeed, we need infant studies to assess the development of rhythmic brain oscillations (see Musacchia et al., 2015, for some of the first reported evoked oscillations analyses in infants).



Actually, a recent study of the spectral amplitude modulation in the speech rhythm shows that (Australian English) infant-directed speech “exaggerates” the synchronization between syllable-rate modulations and stress-rate modulations, whereas adult-directed speech is dominated by syllable-time modulations. This is taken as evidence showing that infant-directed speech “is primarily stress-dominant, which could “tune” the infant brain toward stress-based speech segmentation—an adaptive strategy for bootstrapping early language learning” (Leong et al., 2014). Such infant-directed speech hyperarticulations are taken to help the child acquire the relevant phonological distinctions in her language (Kuhl et al., 1997; Cristia, 2013), a knowledge that is mostly acquired during the first year of life (cf. *i.a.* Kuhl et al., 1992; Werker and Tees, 2002)<sup>5</sup>. Incidentally, it has to be noted that recent studies have shown that the characteristic “hyperarticulation” of infant-directed speech may be restricted to these suprasegmental levels of prosody, given that rather than hyperarticulated, phonemic contrasts can be hypoarticulated in infant-directed speech, i.e., that mothers hyperarticulate their infant-directed speech in prosodic aspects, but in segmental aspects mothers may “speak less clearly to infants than to adults” (cf. Martin et al., 2014).

Now, several authors have proposed that the early acquired rhythmic properties of languages are not idiosyncratic and isolated properties, but rather that they are strongly correlated with the particular syntactic properties of the particular languages (i.e., that there are correlations between rhythmic patterns and syntactic patterns in that languages tend to cluster with the same rhythmic and syntactic properties, conforming linguistic typologies). Furthermore, the explanation of this typological clustering is proposed to derive from the fact that rhythmic patterns serve to bootstrap or catalyze the acquisition of the specific syntactic patterns of each language (cf. *i.a.* Mehler et al., 1988; Christophe et al., 2003; Bernard and Gervain, 2012; Gervain and Werker, 2013; Langus and Nespor, 2013)<sup>6</sup>. In particular, a number of authors have proposed that the relative order between heads and their complements strongly correlates with the rhythmic type of the language. A number of experiments have shown that languages whose correlates of phrasal accent are increases in duration and intensity tend to be head-initial (with a Verb-Object word order) whereas languages that realize stress through a combination of higher pitch and intensity (and possibly also duration) tend to be head-final (with an Object-Verb word order)<sup>7</sup>. This generalization is known as the ‘iambic-trochaic law’ (cf. *i.a.* Hayes, 1995; Nespor et al., 2008; Shukla and Nespor, 2010), which is taken to be a basic law of grouping based on general auditory perception (i.e., not specific to language) that states that units (language or music) that differ in intensity

tend to be grouped as constituents in which the most prominent element comes first, and units that differ in duration are grouped as constituents in which the most prominent element comes last<sup>8</sup>. As Nespor et al. (2008) put it, “if [their] proposal is on the right track, one of the basic properties of syntax can be learned through a general mechanism of perception.”

This line of reasoning is reinforced by recent studies such as Gordon et al. (2015) suggesting that there is a correlation between rhythm perception skills and morpho-syntactic production in children with typical language development (and note also that a strong association between reading skills and meter perception and rhythm processing has been found; Flaunacco et al., 2014; Leong and Goswami, 2014). Likewise, studies like Zumbansen et al. (2014), Leong and Goswami (2014) report the beneficial effects of both pitch and rhythm in the clinical therapy for patients with Broca’s aphasia.

In the next section, I argue for the natural combination of the “globularization” and “prosodic bootstrapping” hypotheses.

#### 4. SYNTHESIS: THE “GLOBULARIZATION HYPOTHESIS” AS A DEVELOPMENTAL FRAME FOR THE “PROSODIC BOOTSTRAPPING” HYPOTHESIS

Let us focus on the two main ideas that we have seen so far, which are that (i) according to the “globularization hypothesis” of Boeckx and Benítez-Burraco (2014a,b), the postnatal globularization of the brain is an essential ingredient for the development of our syntactic capacities, and that (ii) according to the “prosodic bootstrapping hypothesis” of Mehler et al. (1988), Christophe et al. (2003), Bernard and Gervain (2012), Gervain and Werker (2013), Langus and Nespor (2013) and others, children use prosody in order to infer the syntactic pattern of the language they are acquiring.

The combination of these two hypotheses brings about an interesting picture regarding language acquisition: it leaves room for a delay in the acquisition of syntax with respect to prosody. If the “globularization hypothesis” is correct, syntactic capacities develop some months after birth and if the “prosodic bootstrapping hypothesis” is correct, children use prosody as a guiding principle for acquiring syntax. That is, babies may have a rich knowledge of prosody (as pure melodic patterns, unrelated to any syntactic structure) by the moment they develop the capacity to start parsing syntax. Crucially, all the data discussed in Sections 2 and 3 point in that direction: after some months of pre- and post-natal experience with linguistic input, babies have a fairly good knowledge of the prosodic properties of the language(s) spoken around them, this knowledge being arguably well established by the time they develop the structures necessary for parsing syntax. Therefore, babies will be able to use all this phonological knowledge as a guiding principle to discover the syntax behind the acoustic

<sup>5</sup>In fact, lack of adequate acquisition of the phonology of the target language can generate disorders such as dyslexia (Paulescu et al., 2001; Goswami, 2011; Saralegui et al., 2014; see also Benítez-Burraco, 2013)

<sup>6</sup>Donegan and Stampe (1983, 2004) have even proposed a “holistic typology” based on rhythmic grounds in order to account for the polarized structural divergence of languages like Munda and Mon Khmer.

<sup>7</sup>In turn, speakers of languages with different rhythmic patterns like English vs. Japanese tend to behave differently in the way in which they group nonlinguistic stimuli (Iversen et al., 2008).

<sup>8</sup>In a recent study de la Mora et al. (2013) observed that rats group sequences based on pitch variations as trochees, but that they do not group sequences varying in duration as iambs.

signals. As a matter of fact, the hypothesis by Boeckx and Benítez-Burraco (2014a,b) can provide a developmental rationale for the prosodic bootstrapping hypothesis of early language acquisition. Given Boeckx and Benítez-Burraco's hypothesis, it is natural for a rich phonological knowledge to be established *before* the syntactic ability develops, for the necessary mechanisms for phonological acquisition are present at birth. Then, endowed with a rich prosodic knowledge, language-acquiring children will be able to use it as a bias for hypothesizing the syntactic pattern of the target language (which in a Bayesian model could take the form of an informed prior). In a nutshell, the prosodic bootstrapping hypothesis claims that beyond the observed typological correlation between prosodic and syntactic patterns, there is a causal developmental connection between them: babies use prosody to guess the syntactic pattern of their target language and my proposal is that the globularization hypothesis provides a natural developmental frame for the prosodic bootstrapping hypothesis, for it presents

a relatively late syntactic development *vis à vis* the prosodic development.

As a last remark, it should be noted that the globularization hypothesis—besides capturing the fact that prosodic knowledge precedes syntactic knowledge—also leaves room for explaining why first language acquisition is fast, but not immediate, for not all the necessary neurocognitive machinery would be established from birth (cf. Boeckx and Benítez-Burraco, 2014a,b). Even if innately specified, some maturation is in order for a fully language-ready brain.

## FUNDING

This work was partially funded by the following agencies: the European Union's Seventh Framework Program (AThEME 613465), the Basque Government (IT769-13), and the Spanish MINECO (FFI2012-38064-C02-01, FFI2014-53675-P, FFI2013-41509-P).

## REFERENCES

- Abdala, C., and Keefe, D. H. (2012). "Morphological and functional ear development," in *Human Auditory Development*, eds L. A. Werner, R. R. Fay, and A. N. Popper (Dordrecht: Springer), 19–60.
- Arbib, M. A., editor (2013). *Language, Music, and the Brain: A Mysterious Relationship*. Cambridge; London: MIT Press.
- Asano, R., and Boeckx, C. (2015). Syntax in language and music: what is the right level of comparison? *Front. Psychol.* 6:942. doi: 10.3389/fpsyg.2015.00942
- Benítez-Burraco, A. (2013). "Genetics of language: roots of specific language deficits," in *The Cambridge Handbook of Biolinguistics*, eds C. Boeckx and K. K. Grohmann (Cambridge: Cambridge University Press), 375–412.
- Bernard, C., and Gervain, J. (2012). Prosodic cues to word order: what level of representation? *Front. Psychol.* 3:451. doi: 10.3389/fpsyg.2012.00451
- Boeckx, C., and Benítez-Burraco, A. (2014a). Globularity and language-readiness: generating new predictions by expanding the set of genes of interest. *Front. Psychol.* 5:1324. doi: 10.3389/fpsyg.2014.01324
- Boeckx, C., and Benítez-Burraco, A. (2014b). The shape of the language-ready brain. *Front. Psychol.* 5:282. doi: 10.3389/fpsyg.2014.00282
- Brentari, D. (1999). *A Prosodic Model of Sign Language Phonology*. Cambridge; London: MIT Press.
- Byers-Heinlein, K., Burns, T. C., and Werker, J. F. (2010). The roots of bilingualism in newborns. *Psychol. Sci.* 21, 343–348. doi: 10.1177/0956797609360758
- Carral, V., Huotilainen, M., Ruusuvirta, T., Fellman, V., Näättänen, R., and Escera, C. (2005). A kind of auditory 'primitive intelligence' already present at birth. *Eur. J. Neurosci.* 21, 3201–3204. doi: 10.1111/j.1460-9568.2005.04144.x
- Christophe, A., Nespore, M., Teresa Guasti, M., and Van Ooyen, B. (2003). Prosodic structure and syntactic acquisition: the case of the head-direction parameter. *Dev. Sci.* 6, 211–220. doi: 10.1111/1467-7687.00273
- Cooper, R. P., and Aslin, R. N. (1989). The language environment of the young infant: implications for early perceptual development. *Can. J. Psychol.* 43, 247–265. doi: 10.1037/h0084216
- Cristia, A. (2013). Input to Language: the phonetics and perception of infant-directed speech. *Lang. Linguist. Compass* 7, 157–170. doi: 10.1111/lnc3.12015
- de la Mora, D. M., Nespore, M., and Toro, J. M. (2013). Do humans and nonhuman animals share the grouping principles of the iambic-trochaic law? *Atten. Percept. Psychophys.* 75, 92–100. doi: 10.3758/s13414-012-0371-3
- DeCasper, A. J., and Fifer, W. (1980). Of human bonding: newborns prefer their mothers' voices. *Science* 208, 1174–1176. doi: 10.1126/science.7375928
- DeCasper, A. J., and Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behav. Dev.* 9, 133–150. doi: 10.1016/0163-6383(86)90025-1
- Dehaene-Lambertz, G., and Baillet, S. (1998). A phonological representation in the infant brain. *Neuroreport* 9, 1885–1888. doi: 10.1097/00001756-199806010-00040
- Dehaene-Lambertz, G., Hertz-Pannier, L., Dubois, J., Mériaux, S., Roche, A., Sigman, M., et al. (2006). Functional organization of perisylvian activation during presentation of sentences in preverbal infants. *Proc. Natl. Acad. Sci. U.S.A.* 103, 14240–14245. doi: 10.1073/pnas.0606302103
- Donegan, P. J., and Stampe, D. (1983). "Rhythm and the holistic organization of language structure," in *Papers from the Parasession on the Interplay of Phonology, Morphology and Syntax*, eds J. F. Richardson, M. Marks, and A. Chukerman (Chicago, IL: University of Chicago Press), 337–353.
- Donegan, P. J., and Stampe, D. (2004). "Rhythm and the synthetic drift of munda," in *The Yearbook of South Asian Languages and Linguistics 2004*, eds U. N. Singh and S. Singh (Berlin; New York, NY: Mouton de Gruyter), 3–36.
- Drullman, R., Festen, J. M., and Plomp, R. (1994). Effect of temporal envelope smearing on speech reception. *J. Acoust. Soc. Am.* 95, 1053–1064. doi: 10.1121/1.408467
- Dubois, J., Poupon, C., Thirion, B., Simonnet, H., Kulikova, S., Leroy, F., et al. (2015). Exploring the early organization and maturation of linguistic pathways in the human infant brain. *Cereb. Cortex*. doi: 10.1093/cercor/bhv082. [Epub ahead of print].
- Eimas, P. D., Siqueland, E. R., Jusczyk, P., and Vigorito, J. (1971). Speech perception in infants. *Science* 171, 303–306. doi: 10.1126/science.171.3968.303
- Fernald, A. (1984). "The perceptual and affective salience of mothers' speech to infants," in *The Origins and Growth of Communication*, eds L. Feagans, R. M. Golinkoff, and C. Garvey (Norwood, NJ: Ablex), 5–29.
- Fernald, A., and Mazzie, C. (1991). Prosody and focus in speech to infants and adults. *Dev. Psychol.* 27, 209–221. doi: 10.1037/0012-1649.27.2.209
- Fifer, W. (1981). *Early Attachment: Maternal-voice Preference in One- and Three-day Old Infants*. Ph.D. thesis, University of North Carolina, Greensboro, NC.
- Fitch, W. T. (2013). Speech science: tuned to the rhythm. *Nature* 494, 434–435. doi: 10.1038/494434a
- Fluagnacco, E., Lopez, L., Terribili, C., Zoia, S., Buda, S., Tilli, S., et al. (2014). Rhythm perception and production predict reading abilities in developmental dyslexia. *Front. Hum. Neurosci.* 8:392. doi: 10.3389/fnhum.2014.00392
- Gervain, J., and Werker, J. F. (2013). Prosody cues word order in 7-month-old bilingual infants. *Nat. Commun.* 4:1490. doi: 10.1038/ncomms2430
- Ghazanfar, A. A., and Takahashi, D. Y. (2014). Facial expressions and the evolution of the speech rhythm. *J. Cogn. Neurosci.* 26, 1196–1207. doi: 10.1162/jocn.a.00575
- Giraud, A.-L., Kleinschmidt, A., Poeppel, D., Lund, T. E., Frackowiak, R. S., and Laufs, H. (2007). Endogenous cortical rhythms determine cerebral

- specialization for speech perception and production. *Neuron* 56, 1127–1134. doi: 10.1016/j.neuron.2007.09.038
- Giraud, A.-L., and Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nat. Neurosci.* 15, 511–517. doi: 10.1038/nn.3063
- Gleitman, L., and Wanner, E. (1982). “Language acquisition: the state of the art,” in *Language Acquisition: The State of the Art*, eds E. Wanner and L. Gleitman (Cambridge: Cambridge University Press), 3–48.
- Gordon, R. L., Shivers, C. M., Wieland, E. A., Kotz, S. A., Yoder, P. J., and Devin McAuley, J. (2015). Musical rhythm discrimination explains individual differences in grammar skills in children. *Dev. Sci.* 18, 635–644. doi: 10.1111/desc.12230
- Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends Cogn. Sci.* 15, 3–10. doi: 10.1016/j.tics.2010.10.001
- Granier-Deferre, C., Bassereau, S., Ribeiro, A., Jacquet, A.-Y., and DeCasper, A. J. (2011). A melodic contour repeatedly experienced by human near-term fetuses elicits a profound cardiac reaction one month after birth. *PLoS ONE* 6:e17304. doi: 10.1371/journal.pone.0017304
- Gunz, P., Neubauer, S., Golovanova, L., Doronichev, V., Maureille, B., and Hublin, J.-J. (2012). A uniquely modern human pattern of endocranial development. Insights from a new cranial reconstruction of the Neandertal newborn from Mezmaiskaya. *J. Hum. Evol.* 62, 300–313. doi: 10.1016/j.jhevol.2011.11.013
- Gussenhoven, C. (2004). *The Phonology of Tone and Intonation*. Cambridge; New York, NY: Cambridge University Press.
- Hayes, B. (1995). *Metrical Stress Theory: Principles and Case Studies*. Chicago, IL: University of Chicago Press.
- Hepper, P., Scott, D., and Shahidullah, S. (1993). Newborn and fetal response to maternal voice. *J. Reprod. Infant Psychol.* 11, 147–153. doi: 10.1080/02646839308403210
- Hepper, P. G., and Shahidullah, B. S. (1994). Development of fetal hearing. *Arch. Disease Child.* 71, F81–F87. doi: 10.1136/fn.71.2.F81
- Iversen, J. R., Patel, A. D., and Ohgushi, K. (2008). Perception of rhythmic grouping depends on auditory experience. *J. Acoust. Soc. Am.* 124, 2263–2271. doi: 10.1121/1.2973189
- Jusczyk, P. W., and Nelson, D. G. K. (1996). “Syntactic units, prosody, and psychological reality during infancy,” in *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*, eds J. L. Morgan and K. Demuth (Hillsdale: Lawrence Erlbaum), 389–410.
- Karzon, R. G., and Nicholas, J. G. (1989). Syllabic pitch perception in 2- to 3-month-old infants. *Percept. Psychophys.* 45, 10–14. doi: 10.3758/BF03208026
- Katz, G. S., Cohn, J. F., and Moore, C. A. (1996). A combination of Vocal f0 dynamic and summary features discriminates between three pragmatic categories of infant-directed speech. *Child Dev.* 67, 205–217. doi: 10.1111/j.1467-8624.1996.tb01729.x
- Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., et al. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science* 277, 684–686. doi: 10.1126/science.277.5326.684
- Kuhl, P. K., and Miller, J. D. (1975). Speech perception by the chinchilla: voiced-voiceless distinction in alveolar plosive consonants. *Science* 190, 69–72. doi: 10.1126/science.1166301
- Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., and Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 255, 606–608. doi: 10.1126/science.1736364
- Lakatos, P., Shah, A. S., Knuth, K. H., Ulbert, I., Karmos, G., and Schroeder, C. E. (2005). An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *J. Neurophysiol.* 94, 1904–1911. doi: 10.1152/jn.00263.2005
- Langus, A., and Nespore, M. (2013). Language development in infants: what do humans hear in the first months of life? *Hear. Balance Commun.* 11, 121–129. doi: 10.3109/21695717.2013.817133
- Leong, V. (2012). *Prosodic Rhythm in the Speech Amplitude Envelope - Amplitude Modulation Phase Hierarchies (AMPHs) and AMPH Models*. Ph.D. thesis, University of Cambridge, Cambridge.
- Leong, V., and Goswami, U. (2014). Impaired extraction of speech rhythm from temporal modulation patterns in speech in developmental dyslexia. *Front. Hum. Neurosci.* 8:96. doi: 10.3389/fnhum.2014.00096
- Leong, V., Kalashnikova, M., Burnham, D., and Goswami, U. (2014). “Infant-directed speech enhances temporal rhythmic structure in the envelope,” in *Proceedings of INTERSPEECH 2014* (Singapore), 2563–2567.
- Lieberman, D. E., McBratney, B. M., and Krovitz, G. (2002). The evolution and development of cranial form in *homo sapiens*. *Proc. Natl. Acad. Sci. U.S.A.* 99, 1134–1139. doi: 10.1073/pnas.022440799
- Mampe, B., Friederici, A. D., Christophe, A., and Wermke, K. (2009). Newborns’ cry melody is shaped by their native language. *Curr. Biol.* 19, 1994–1997. doi: 10.1016/j.cub.2009.09.064
- Martin, A., Schatz, T., Versteegh, M., Miyazawa, K., Mazuka, R., Dupoux, E., et al. (2014). Mothers speak less clearly to infants than to adults: a comprehensive test of the hyperarticulation hypothesis. *Psychol. Sci.* 26, 341–347. doi: 10.1177/0956797614562453
- Martins, P. T., and Boeckx, C. (2014). Attention mechanisms and the mosaic evolution of speech. *Front. Psychol.* 5:1463. doi: 10.3389/fpsyg.2014.01463
- Marx, V., and Nagy, E. (2015). Fetal behavioural responses to maternal voice and touch. *PLoS ONE* 10:e0129118. doi: 10.1371/journal.pone.0129118
- Mehler, J. (1974). “Connaissance par désapprentissage,” in *L'Unité de l'Homme II: Le Cerveau Humain*, eds E. Morin and M. Piattelli-Palmarini (Paris: Editions du Seuil), 25–37.
- Mehler, J., Bertoncini, J., and Barriere, M. (1978). Infant recognition of mother’s voice. *Perception* 7, 491–497.
- Mehler, J., and Dupoux, E. (1990). *Naitre Humain*. Paris: Odile Jacob.
- Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., and Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition* 29, 143–178. doi: 10.1016/0010-0277(88)90035-2
- Moffitt, A. R. (1971). Consonant cue perception by twenty- to twenty-four-week-old infants. *Child Dev.* 42, 717–731. doi: 10.2307/1127443
- Molnar, M., Gervain, J., and Carreiras, M. (2014a). Within-rhythm class native language discrimination abilities of basque-spanish monolingual and bilingual infants at 3.5 months of age. *Infancy* 19, 326–337. doi: 10.1111/inf.12041
- Molnar, M., Lallier, M., and Carreiras, M. (2014b). The amount of language exposure determines nonlinguistic tone grouping biases in infants from a bilingual environment. *Lang. Learn.* 64, 45–64. doi: 10.1111/lang.12069
- Morse, P. A. (1972). The discrimination of speech and nonspeech stimuli in early infancy. *J. Exp. Child Psychol.* 14, 477–492. doi: 10.1016/0022-0965(72)90066-5
- Musacchia, G., Ortiz-Mantilla, S., Realpe-Bonilla, T., Roesler, C. P., and Benasich, A. A. (2015). Infant auditory processing and event-related brain oscillations. *J. Vis. Exp.* e52420. doi: 10.3791/52420
- Nazzi, T., Bertoncini, J., and Mehler, J. (1998a). Language discrimination by newborns: toward an understanding of the role of the role of rhythm. *J. Exp. Psychol.* 24, 756–766. doi: 10.1037/0096-1523.24.3.756
- Nazzi, T., Floccia, C., and Bertoncini, J. (1998b). Discrimination of pitch contours by neonates. *Infant Behav. Dev.* 21, 779–784. doi: 10.1016/S0163-6383(98)90044-3
- Nazzi, T., Iakimova, G., Bertoncini, J., Frdonie, S., and Alcantara, C. (2006). Early segmentation of fluent speech by infants acquiring French: emerging evidence for crosslinguistic differences. *J. Mem. Lang.* 54, 283–299. doi: 10.1016/j.jml.2005.10.004
- Nespor, M., Shukla, M., van de Vijver, R., Avesani, C., Schrauldorf, H., and Donati, C. (2008). Different phrasal prominence realizations in VO and OV languages. *Lingue e Linguaggio* VII-2, 139–168. doi: 10.1418/28093
- Neubauer, S., Gunz, P., and Hublin, J.-J. (2010). Endocranial shape changes during growth in chimpanzees and humans: a morphometric analysis of unique and shared aspects. *J. Hum. Evol.* 59, 555–566. doi: 10.1016/j.jhevol.2010.06.011
- Ockelford, E., Vince, M., Layton, C., and Reader, M. (1988). Responses of neonates to parents’ and others’ voices. *Early Hum. Dev.* 18, 27–36. doi: 10.1016/0378-3782(88)90040-0
- Olsho, L. W., Schoon, C., Sakai, R., Turpin, R., and Sperduto, V. (1982). Auditory frequency discrimination in infancy. *Dev. Psychol.* 18, 721–726. doi: 10.1037/0012-1649.18.5.721
- Pang, E. W., and Taylor, M. J. (2000). Tracking the development of the N1 from age 3 to adulthood: an examination of speech and non-speech stimuli. *Clin. Neurophysiol.* 111, 388–397. doi: 10.1016/S1388-2457(99)00259-X
- Panneton, R., and Newman, R. (2012). “Development of speech perception,” in *Human Auditory Development*, eds L. A. Werner, R. R. Fay, and A. N. Popper (New York, NY: Springer), 197–222.



- Paulescu, E., Déonot, J.-F., Fazio, F., McCrory, E., Chanoine, V., Brunswick, N., et al. (2001). Dyslexia: cultural diversity and biological unity. *Science* 291, 2165–2167. doi: 10.1126/science.1057179
- Peña, M., Maki, A., Kovacic, D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, F., et al. (2003). Sounds and silence: an optical topography study of language recognition at birth. *Proc. Natl. Acad. Sci. U.S.A.* 100, 11702–11705. doi: 10.1073/pnas.1934290100
- Peelle, J. E., and Davis, M. H. (2012). Neural oscillations carry speech rhythm through to comprehension. *Lang. Sci.* 3:320. doi: 10.3389/fpsyg.2012.00320
- Pfau, R., and Quer, J. (2010). “Nonmanuals: their grammatical and prosodic roles,” in *Sign Languages*, ed D. Brentari (Cambridge: Cambridge University Press), 381–402.
- Querleu, D., Lefebvre, C., Titran, M., Renard, X., Morillion, M., and Crepin, G. (1984). Reaction of the newborn infant less than 2 hours after birth to the maternal voice. *J. Gynecol. Obstet. Biol. Reprod.* 13, 125–134.
- Ramus, F., Hauser, M. D., Miller, C., Morris, D., and Mehler, J. (2000). Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science* 288, 349–351. doi: 10.1126/science.288.5464.349
- Rebuschat, P., Rohrmeier, M., Hawkins, J. A., and Cross, I. (eds.). (2011). *Language and Music as Cognitive Systems*. Oxford: Oxford University Press.
- Sansavini, A., Bertocchini, J., and Giovanelli, G. (1997). Newborns discriminate the rhythm of multisyllabic stressed words. *Dev. Psychol.* 33, 3–11. doi: 10.1037/0012-1649.33.1.3
- Saralegui, I., Ontañón, J. M., Fernandez-Ruanova, B., Garcia-Zapirain, B., Basterra, A., and Sanz-Arigita, E. J. (2014). Reading networks in children with dyslexia compared to children with ocular motility disturbances revealed by fMRI. *Front. Hum. Neurosci.* 8:936. doi: 10.3389/fnhum.2014.00936
- Selkirk, E. (2011). “The syntax-phonology interface,” in *The Handbook of Phonological Theory*, 2nd Edn., eds J. Goldsmith, J. Riggle, and A. C. L. Yu (Oxford: Wiley-Blackwell), 435–484.
- Shahidullah, S., and Hepper, P. G. (1994). Frequency discrimination by the fetus. *Early Hum. Dev.* 36, 13–26. doi: 10.1016/0378-3782(94)90029-9
- Shukla, M., and Nespor, M. (2010). “Rhythmic patterns cue word order,” in *The Sound Patterns of Syntax*, eds N. Erteschik-Shir and L. Rochman (Oxford: Oxford University Press), 174–188.
- Smith, Z. M., Delgutte, B., and Oxenham, A. J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature* 416, 87–90. doi: 10.1038/416087a
- Spence, M. J., and DeCasper, A. J. (1987). Prenatal experience with low-frequency maternal-voice sounds influence neonatal perception of maternal voice samples. *Infant Behav. Dev.* 10, 133–142. doi: 10.1016/0163-6383(87)90028-2
- Telkemeyer, S., Rossi, S., Koch, S. P., Nierhaus, T., Steinbrink, J., Poeppel, D., et al. (2009). Sensitivity of newborn auditory cortex to the temporal structure of sounds. *J. Neurosci.* 29, 14726–14733. doi: 10.1523/JNEUROSCI.1246-09.2009
- Telkemeyer, S., Rossi, S., Nierhaus, T., Steinbrink, J., Obrig, H., and Wartenburger, I. (2011). Acoustic processing of temporally modulated sounds in infants: evidence from a combined near-infrared spectroscopy and EEG study. *Lang. Sci.* 1:62. doi: 10.3389/fpsyg.2011.00062
- Theofanopoulou, C. (2015). Brain asymmetry in the white matter making and globularity. *Front. Psychol.* 6:1355. doi: 10.3389/fpsyg.2015.01355
- Toro, J. M., Trobalon, J. B., and Sebastián-Gallés, N. (2003). The use of prosodic cues in language discrimination tasks by rats. *Anim. Cogn.* 6, 131–136. doi: 10.1007/s10071-003-0172-0
- Truckenbrodt, H. (2007). “The syntax-phonology interface,” in *The Cambridge Handbook of Phonology*, ed P. de Lacy (Cambridge: Cambridge University Press), 435–456.
- VanRullen, R., and Dubois, J. (2011). The Psychophysics of brain rhythms. *Front. Psychol.* 2:203. doi: 10.3389/fpsyg.2011.00203
- Vihman, M. M. (2014). *Phonological Development: The First Two Years*, 2 Edn. Malden, MA; Oxford: Wiley Blackwell.
- Vince, M. A. (1979). Postnatal effects of prenatal sound stimulation in the guinea pig. *Anim. Behav.* 27, 908–918. doi: 10.1016/0003-3472(79)90029-0
- Vince, M. A., Armitage, S. E., Walser, E. S., and Reader, M. (1982). Postnatal consequences of prenatal sound stimulation in the sheep. *Behaviour* 81, 128–139. doi: 10.1163/156853982X00111
- Wahl, M., Marzinzik, F., Friederici, A. D., Hahne, A., Kupsch, A., Schneider, G.-H., et al. (2008). The human thalamus processes syntactic and semantic language violations. *Neuron* 59, 695–707. doi: 10.1016/j.neuron.2008.07.011
- Wang, L., Uhrig, L., Jarraya, B., and Dehaene, S. (2015). Representation of numerical and sequential patterns in macaque and human brains. *Curr. Biol.* 25, 1966–1974. doi: 10.1016/j.cub.2015.06.035
- Ward, C. D., and Cooper, R. P. (1999). A lack of evidence in 4-month-old human infants for paternal voice preference. *Dev. Psychobiol.* 35, 49–59. doi: 10.1002/(SICI)1098-2302(199907)35:1<49::AID-DEV7>3.0.CO;2-3
- Werker, J. F., and Tees, R. C. (2002). Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behav. Dev.* 25, 121–133. doi: 10.1016/S0163-6383(02)00093-0
- Zumbansen, A., Peretz, I., and Hébert, S. (2014). The combination of rhythm and pitch can account for the beneficial effect of melodic intonation therapy on connected speech improvements in Broca's aphasia. *Front. Hum. Neurosci.* 8:592. doi: 10.3389/fnhum.2014.00592

**Conflict of Interest Statement:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer, Silvia Martínez Ferreiro, and handling editor, Cedric Boeckx, declared their shared affiliation, and the handling editor states that the process nevertheless met the standards of a fair and objective review.

Copyright © 2015 Irurtzun. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# Temporal Attention as a Scaffold for Language Development

Ruth de Diego-Balaguer<sup>1,2,3\*</sup>, Anna Martinez-Alvarez<sup>2,3</sup> and Ferran Pons<sup>3,4</sup>

<sup>1</sup> Institució Catalana de Recerca i Estudis Avançats, Barcelona, Spain, <sup>2</sup> Cognition and Brain Plasticity Unit, Institut d'Investigació Biomèdica de Bellvitge, Barcelona, Spain, <sup>3</sup> Department of Basic Psychology, University of Barcelona, Barcelona, Spain, <sup>4</sup> Department of Basic Psychology, Institute for Brain, Cognition and Behavior (IR3C), University of Barcelona, Barcelona, Spain

## OPEN ACCESS

### Edited by:

Antonio Benítez-Burraco,  
University of Huelva, Spain

### Reviewed by:

LouAnn Gerken,  
University of Arizona, USA  
Andriy Myachkov,  
Northumbria University, UK

### \*Correspondence:

Ruth de Diego-Balaguer  
ruth.dediego@ub.edu

### Specialty section:

This article was submitted to  
Language Sciences,  
a section of the journal  
Frontiers in Psychology

**Received:** 15 September 2015

**Accepted:** 11 January 2016

**Published:** 02 February 2016

### Citation:

de Diego-Balaguer R,  
Martinez-Alvarez A and Pons F (2016)  
Temporal Attention as a Scaffold  
for Language Development.  
Front. Psychol. 7:44.  
doi: 10.3389/fpsyg.2016.00044

Language is one of the most fascinating abilities that humans possess. Infants demonstrate an amazing repertoire of linguistic abilities from very early on and reach an adult-like form incredibly fast. However, language is not acquired all at once but in an incremental fashion. In this article we propose that the attentional system may be one of the sources for this developmental trajectory in language acquisition. At birth, infants are endowed with an attentional system fully driven by salient stimuli in their environment, such as prosodic information (e.g., rhythm or pitch). Early stages of language acquisition could benefit from this readily available, stimulus-driven attention to simplify the complex speech input and allow word segmentation. At later stages of development, infants are progressively able to selectively attend to specific elements while disregarding others. This attentional ability could allow them to learn distant non-adjacent rules needed for morphosyntactic acquisition. Because non-adjacent dependencies occur at distant moments in time, learning these dependencies may require correctly orienting attention in the temporal domain. Here, we gather evidence uncovering the intimate relationship between the development of attention and language. We aim to provide a novel approach to human development, bridging together temporal attention and language acquisition.

**Keywords:** language development, infancy, attention, temporal orienting, statistical learning, rule learning, morphosyntactic development, word segmentation

## INTRODUCTION

Speech is a complex auditory stimulation. A single word in speech can be perceived as a sequence of phonemes, as a whole word, as a stem and a suffix or as having a specific meaning, depending on the level of processing. In order to face this complexity infants do not learn all of this information at once but rather in an incremental fashion. In particular, two main linguistic milestones -word segmentation and non-adjacent rule acquisition- appear in a sequential fashion. During the first months, infants are able to segment speech into words and recognize them; however, it is not until after the first year that they are able to understand and detect the subtle changes carried by different rule transformations (Gómez and Maye, 2005; Christophe et al., 2008). As a matter of fact, brain development in general is not uniformly distributed through infancy. Different brain structures and their white matter connections do not develop homogeneously, nor do all cognitive functions develop at the same speed (Gogtay et al., 2004; Diamond, 2007). In particular attention shows also a developmental progression with different mechanisms arising at different moments

of development. *Exogenous attention*, captured by salient events in the environment, is functional much earlier than *endogenous attention*, which allows for selecting which information to process and which to ignore (Posner and Cohen, 1984). The progressive general cognitive development may be seen as affecting all functions independently. However, cognitive functions do not work in isolation. In particular, the attentional system acts as a filter to any incoming stimulation, influencing perception, and therefore may affect learning, which suggests that the development of the attention system is likely to shape the way language is processed and how it develops along with the available attention resources.

Extensive literature has previously reported a link between attention and different aspects of language processing. Adult speakers of different languages tend to adapt the syntactic structure of their productions as a function of their focus of attention on the visual information they describe (Myachykov et al., 2005; Ibbotson et al., 2013; Tomblin and Myachykov, 2015). Similarly, focus and topicalization are naturally used to draw attention to relevant elements in the sentence (Jackendoff, 2002). Indeed, using focus to comprehend sentences activates left inferior parietal lobe (IPL) overlapping with the attention-orienting network used in visual attention (Kristensen et al., 2012). Even babies produce isolated words in an attempt to attract the listener's attention to their own focus of interest (Jackendoff, 2002). This behavior is closely linked to the development of joint attention present at the end of the first year of life (Bruner, 1983; Carpenter et al., 1998; Brooks and Meltzoff, 2002). From joint attention interaction infants begin to link words with objects and events (Baldwin, 1991, 1993).

This previous literature converges with our proposal with the close link between attention and language, the source of the link is closely related in these studies to the communicative roots of language, given that attention is used as a tool to drive the listener's attention to the focus of attention of the speaker (Smith et al., 2011). However, our proposal, in contrast, is intimately related to processing. The development of attention affects how the input is processed because it filters the input received, independently of the presence of an interlocutor and in the absence of a message to be transmitted. This is what shapes language learning from this point of view, in the same way as it shapes learning of other sources of information.

The present proposal presents an integrative approach of language acquisition, in which the powerful and dynamic interplay of exogenous and endogenous attention mechanisms allows infants to focus on different aspects of speech at different moments in development. In particular, during the initial stages of language acquisition, attention is captured by salient elements of speech, such as prosodic cues (e.g., pitch, rhythm, or pauses) because the infant perceptual system is guided by stimulus-driven attention. As months pass and endogenous attention progressively develops, this more flexible mechanism can be used to learn non-adjacent linguistic dependencies. This allows filtering out irrelevant information and selectively focusing on relevant elements that reliably predict forthcoming information. As we will argue and support with evidence from infant

development, changes in these two different aspects of cognition are not independent. In other words, the attention mechanisms available early on limit the type of linguistic information that infants can extract from speech. The delay in development of more controlled mechanisms of endogenous attention may not indicate a disadvantage in language acquisition but rather an advantage at early linguistic stages. That is, in agreement with Newport's "Less is More" hypothesis (Newport, 1990), this delay in the development of endogenous attention and the initial use of more automatic exogenous attention mechanisms may allow young infants to face a perceptual simplification of the complex speech stream early in the learning process. Therefore, relying in prosodic cues may be crucial and beneficial during the first months of life, when exogenous attention is the main mechanism available. Such a pattern would lead to the observed early segmentation and acquisition of words during the first months, which is followed by a shift later in infancy to focus on the upcoming information indicated by relevant cues (as in non-adjacent dependency learning) when the infant is able to select which information to attend and which to ignore.

Crucially, the other important difference from the previous theoretical approaches is that our proposal is based on the allocation of attention in the *temporal domain*. Whereas previous proposals have focused purely on visual attention and how it influences the conceptualization of the message to be conveyed (Levelt, 1989) and how this is reflected in the linguistic output, we are interested in how attention affects the processing of the ongoing auditory stimulation -the speech flow. Given that speech is a sequence of sounds that unfolds in time, attention to speech is necessarily oriented in the temporal domain. Because temporal-selective attention directs resources to certain moments in time, enhancing perception (Correa et al., 2006; Nobre et al., 2007), it can allow for the extraction of different events in speech (e.g., consonants, vowels, words, and phrases) that have different durations and appear in a certain order and moments in ongoing speech. Recently, more general proposals have also underscored that cerebral mechanisms for timing and ordinal knowledge are in charge of the neural representation of sequences in different domains, including language (Dehaene et al., 2015). More precisely, speech has temporally rhythmic and salient prosodic cues that capture attention automatically when they appear, helping, for example, to locate boundaries to segment words. On the other hand, segments carrying cues for rule dependencies may have different durations and different onset times (such as suffixes and pronouns). Attention can be progressively tuned to focus on these cues when they are progressively noticed to predict later upcoming dependencies in a sequence of words. This tuning requires the engagement of endogenous attention in the temporal domain. Therefore, acquiring words and rules may require the engagement of different attentional systems. A dynamic shift between systems should develop in the course of learning. Indeed, recent data in adults show that the same prosodic cues can lead to exogenous effects related to segmentation, even in the absence of any possible learning and endogenous effects when the prosodic information can be used as a cue to extract non-adjacent rules (de Diego-Balaguer et al., 2015).

By studying the attention mechanisms involved in language learning, we also pave the way to understand some of the sources of language learning disabilities. In the following sections, we first describe the typically developmental trajectory of attention and language functions and their underlying brain development. We review evidence supporting the hypothesis that maturation of the attention mechanisms may serve as a scaffold for language development, and we review evidence indicating the close relationship between attention deficits and impairments in language acquisition.

## STAGES OF DEVELOPMENT OF THE ATTENTION SYSTEM

Before entering into the details of attention development, some conceptual clarifications can be helpful concerning the terms that are used throughout this paper. In the attention literature, a distinction between exogenous (bottom-up) and endogenous (top-down) attention has been classically proposed, and a plethora of studies have dissected the effects that characterize each of these systems and their interactions (Chica et al., 2013). In brief, both types of attention have been proven to facilitate processing. However, exogenous orienting appears even when a secondary task is performed, and it can be voluntarily attenuated but not completely suppressed. Endogenous attention is often voluntary, but it can also appear with no effort and even when participants are not aware of the relationship between the cue and the target. Some models of attention (Corbetta and Shulman, 2002; Corbetta et al., 2008) propose a distinction in terms of *stimulus-driven* vs. *goal-directed* attention, which partially overlap with *exogenous* vs. *endogenous* attention but have important discrepancies that are worth mentioning here. Within the frontal and parietal brain regions involved in attention, stimulus-driven attention involves a more ventral fronto-parietal network (Corbetta et al., 2008), including the inferior parietal cortex, the ventral and inferior prefrontal cortex (PFC), and insula, as well as subcortically, the superior colliculus. Goal-directed attention, in contrast, involves a more dorsal fronto-parietal network, including the middle PFC and the superior parietal lobe (Corbetta et al., 2008), and subcortically, the pulvinar of the thalamus.

An important distinction within this framework that might help to understand the attentional systems in a less dichotomic way is the distinction between saliency when (i) no task or goal is present (i.e., exogenous saliency) compared to when (ii) elements are salient because they share some feature that is relevant for the task or goal of the subject despite not being the target of the task (i.e., task-relevant saliency). In other words, a red circle surrounded by green squares will attract attention in the absence of any task due to their exogenous saliency; however, a green circle will drive our attention if our task is to detect a green square because the circle shares a relevant feature (i.e., green) to our task. This distinction is important not only because neuroimaging data show that when performing a task, sensory-salient and task-relevant stimuli induce the activation of different brain networks but also because in the absence

of a task, these different types of stimulation do engage the ventral network (Corbetta and Shulman, 2002; Chica et al., 2013). Because very young infants do not have a goal-directed system available, salient stimuli in the environment may trigger the ventral network and subcortical areas. With incremental learning and the progressive availability of the goal-directed system, the relevant elements in the environment attract the ventral attention system in a more *task-relevant* manner. In terms of what this might mean for language, infants may first be attracted by any change in pitch, pauses or in voice onset time in speech sounds due to their intrinsic saliency, whereas later in development, once prosodic characteristics, representations of the phonemes and words of their native language are learned, only those prosodic variations and speech sounds that correspond to the contrast of their language will attract their attention. To avoid misunderstandings in the course of the paper, we will refer to *saliency* to designate only those stimuli that attract attention due to their sensory characteristics irrespective of their relevance.

Turning back to the development of the attentional system, three main attentional mechanisms have been described in the literature on attention development: alertness, orienting, and endogenous attention (Colombo, 2001). The first two characterize exogenous attention. Rudimentary forms of each of the functions of attention are already present to some degree at birth, but each exhibits progressive maturity during the first years of life.

The arousal system is already present at early stages of development. This attention system is associated with an infant's level of alertness and readiness to process stimuli from the environment. From birth to 2 months of age, alertness is commonly initiated by exogenous stimulation (Wolff, 1965). Very young infants show "obligatory attention" (Stechler and Latz, 1966) or "sticky fixation" (Hood, 1995), a difficulty in interrupting gaze from a given stimulus they are fixating in order to shift attention to a different one. The efficiency of disengaging from and shifting gaze to a stimulus increases during the first months after birth (Hunnius and Geuze, 2004). This phenomenon is tied to the neurological maturation of the visual pathway and associated with subcortical structures (Richards et al., 2010) that, although present at birth, are still developing in terms of their connectivity to most cortical areas (Casey et al., 2004; Uylings, 2006). Between 2 and 3 months of age, maturation moderates the inhibition mechanisms that limit eye-movements, which start to gain cortical control. Effective visual exploration requires disengaging and shifting gaze across different locations, and the perseveration of this sticky phenomenon at 7 months of age is an early feature of later emerging autism (Elison et al., 2013) that persists with disengagement difficulties in childhood (Landry and Bryson, 2004). Therefore, this attentional mechanism might be more related to communicative and social aspects of language development.

At later stages, infants start developing the ability to orient attention toward a particular stimulus in space (Courage and Richards, 2008). The infant's visual behavior in the first year of life is dominated by an orienting system of attention with two main components (Ruff and Rothbart, 1996). On the one hand, the

spatial-orienting network, which includes the posterior parietal cortex and several subcortical systems, mediates attentional functions, such as engagement, disengagement, shifting, and inhibition of return. On the other hand, the object recognition network, which includes pathways from the primary visual cortex to the parietal cortex and the inferior temporal cortex, orients attention to object features. A remarkable developmental progression of this orienting system occurs between 3 and 9 months of age (Ruff and Rothbart, 1996). During this period infant flexibly and quickly orient attention to stimuli in the environment in terms of experiential factors (e.g., novelty or complexity) rather than their exogenous salience (Courage et al., 2006).

Finally, the endogenous orienting of attention shows a slower and later developmental time course than other attention systems, showing a remarkable change during the later parts of the first year and beyond (Colombo, 2001). It is not until the end of the first year that more complex features arise and endogenous control of attention starts acquiring an executive component to a greater extent (Courage and Richards, 2008), which is closely related to the initial maturation of the dorsal prefrontal and the anterior cingulate cortices (Posner and Petersen, 1990). Age-related attentional improvements are related to changes in structural and functional connectivity (Rueda et al., 2015). Data reveal that increased attentional performance is related to greater information transfer in the brain, which involves distributed brain nodes and paths that connect these nodes. Crucially, the anterior cingulate cortex does not begin to develop long-range connectivity with other brain areas until after the first year of life, developing progressively during childhood (Fair et al., 2009; Gao et al., 2009). However, although the structural connectivity pattern in children resembles that of adults, the functional connectivity of attentional networks shows different patterns. While adults' orienting and executive attentional systems exhibit separate functional networks, these systems are more unified in children (Fair et al., 2007). In summary, whereas exogenous attention shows an earlier maturation course, endogenous attention develops later and slower, continuing its development through childhood and until adolescence (Colombo and Cheatham, 2006; Johnson et al., 2015).

Despite the apparent dichotomy between exogenous and endogenous attention, the appearance of the latter does not imply an inhibition of the exogenous mechanism but a better interaction between both. As it has been highlighted in the adult literature (Corbetta et al., 2008), stimulus-driven attention is able to break the engagement of goal-directed attention, highlighting the close interaction between systems. However, goal-directed attention can attenuate the interference from distractors by decreasing the activation of stimulus-driven attention. The ability to ignore salient distractors to support learning is observed at 8 and 12 months (Althaus and Mareschal, 2012; Tummelshammer et al., 2014) but is not present earlier, indicating a stronger influence of exogenous factors on young infants' attention. From this point of view, it is clear that the development of the exogenous and endogenous mechanisms of attention do not show a strict sequential order but rather a smooth overlap with subtle signs of endogenous attention appearing

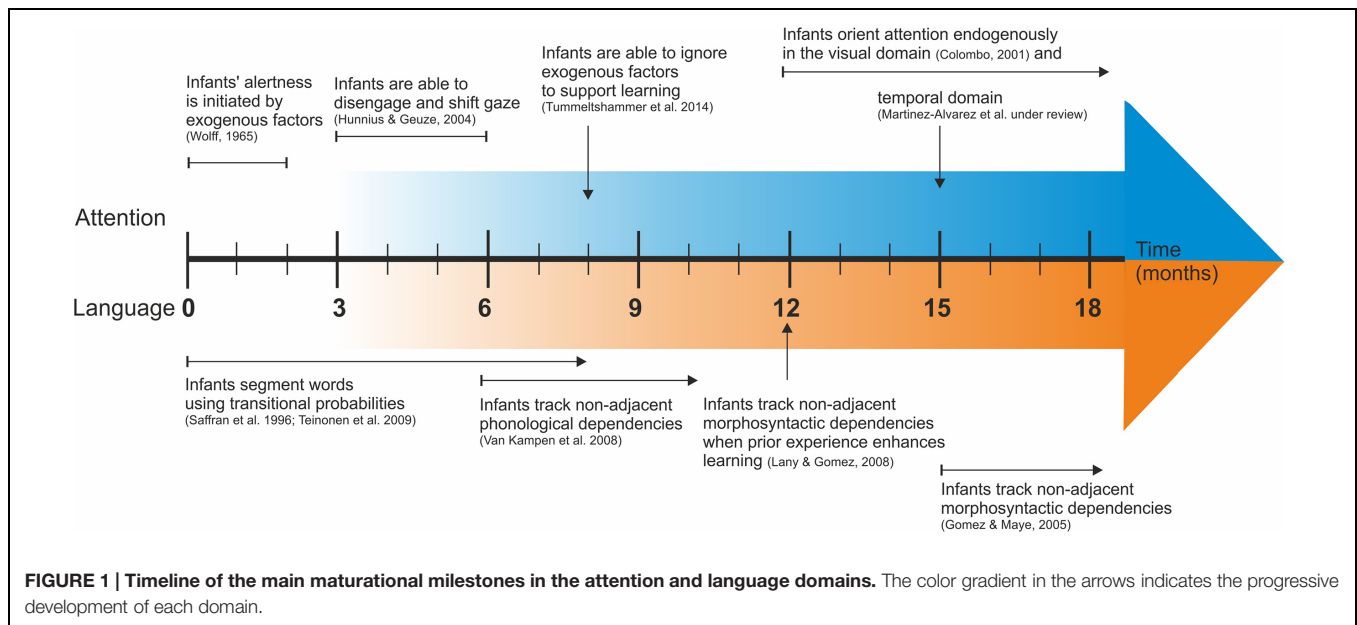
before 8 months of age but with poor command observed to progressive control reached at the end of the first year (see **Figure 1**).

Most of these developmental descriptions of infants have been based on studies of visuospatial attention. Although this description is useful in understanding the developmental progression of attention in general, tracking auditory information in time is critical when we consider language learning due to the intrinsic temporal characteristics of speech. Generally speaking, to orient attention in time, different capacities need to be in place. Infants need to be able to perceive the difference between distinct temporal lags and be sensitive to the order of elements in a sequence. Once these perceptual capacities are available, the attentional system should be able to orient to these different elements in the auditory domain to extract information from speech. Temporal processing is supported by a cortico-subcortical network, including the premotor cortex, basal ganglia and cerebellum that have been proposed to also be involved in speech processing (Kotz and Schwartze, 2010). As previously mentioned, subcortical structures are functional and are used by infants from birth (Casey et al., 2004; Uylings, 2006), and sensorimotor cortices are the first to develop (Dehaene-Lambertz and Spelke, 2015). The early availability of these structures may allow infants to use temporal information in an exogenous manner in the early stages of development.

Studies exploring infants' ability to perceive time mostly focus on regular temporal structure perception (e.g., rhythm and regular isochronic sequences), i.e., focusing in attention mechanisms that are mainly exogenous and stimulus-driven (Demany et al., 1977; Haith et al., 1988; Adler et al., 2008). These studies show that in the first months of age, infants are sensitive and can orient their attention in time following regular patterns. vanMarle and Wynn (2006) reported that 6-month-old infants can discriminate event durations between 2 and 4 s. Brannon et al. (2004, 2008) additionally showed that 10-month-old infants can detect changes in temporal rhythm by detecting a temporal deviation in a stream of tones formed by a regular inter-stimulus interval. In terms of infants' ability to benefit from rhythmic and regular patterns, their behavior is similar to that observed in adult research (Large and Jones, 1999; Barnes and Jones, 2000; Sanabria et al., 2011).

In contrast, the ability to orient attention in time endogenously has not been reported in infancy. Recent data (Martinez-Alvarez et al., under review) indicate that the ability to endogenously orient attention in time appears after the ability to orient attention in space. More precisely, whereas 12-month-olds show only spatial orienting abilities, 15-month-olds are able to adapt their anticipatory behavior according to both spatial and temporal predictive cues. A recent study with children revealed also that the developmental trajectory of voluntarily use temporal cues is delayed relative to the use of spatial cues. However, this study showed that 11-year-olds were only able to implicitly but not voluntarily orient attention in time (Johnson et al., 2015), which apparently seems to contradict the results with infants (Martinez-Alvarez et al., under review). Different explanations of these results with children are possible. As Johnson and colleagues explain, one possibility is that the temporal cues were





conceptually more demanding than the spatial cues. Another possibility is that the spatial uncertainty of target appearance in their paradigm diminished the utility of the temporal cue. Electrophysiological and behavioral investigations have shown that temporal predictability is most successful when joined with spatial predictability (Doherty et al., 2005; Rohenkohl et al., 2014). Indeed, preliminary evidence from Coull's lab shows that children can use temporal cues when the spatial location of the target is known in advance (Johnson et al., 2015). Further investigations are needed for a better understanding the development of temporal attention at the functional and anatomical levels.

In sum, the development of attention is characterized by a shift from exogenous, stimulus-driven orienting of attention, particularly during the first 3 months of age, to a smooth progression to greater endogenous control, the first hints observable before 8 months and showing a marked dominance after the first year (Johnson, 1990; Ruff and Rothbart, 1996). Although little evidence is available from attention orienting in the temporal domain, clear effects of sensitivity to temporal differences and rhythmic cues are present early on, whereas endogenous orienting of attention in time appears later.

## LANGUAGE DEVELOPMENT IN INFANCY IN RELATION TO ATTENTION

In the current section, we review the developmental trajectory of linguistic abilities in infants, focusing on studies on word segmentation and non-adjacent rule learning, which are the two main milestones of interest for our hypothesis. Throughout the review, we point out the role of the attention mechanisms related to the language data available at each stage.

## Early Stages of Language Learning and Exogenous Mechanisms of Attention

The early capacities of infants to acquire their native language have been extensively reported. Even before they begin to produce their first words, infants have already acquired an important amount of linguistic knowledge. One very early ability is their sensitivity to perceive the rhythmic characteristics of language at birth, showing discrimination of the stress patterns in different languages and an early preference for their native language stress pattern (Nazzi et al., 1998). Importantly, prosodic characteristics, such as intonation, stress and pitch variations, are salient perceptual cues that can easily attract infant's exogenous attention. These prosodic cues play a key role in word segmentation because infants exploit these even before they use other cues to locate word boundaries (Mattys et al., 1999).

During the first months of life, infants are capable of extracting words in spoken language by detecting and exploiting other perceptual cues. For example, neonates (Teinonen et al., 2009) and 8 months old infants can make use of statistical regularities between adjacent syllables (also known as transitional probabilities, TP) to locate word boundaries and extract words from both artificial (Saffran et al., 1996) and natural languages (Pelucchi et al., 2009). On the other hand, the combination of both prosodic and statistical cues shows how predominant are the acoustic features of infant-directed (ID) speech (e.g., exaggerated pitch contours) compared to adult-directed (AD) speech to attract infants' attention (Fernald, 1985; Cooper and Aslin, 1990) and to facilitate infants' word segmentation (Trainor and Desjardins, 2002; Thiessen et al., 2005).

Statistical learning is a remarkable ability, and numerous studies have been developed to understand the mechanisms underlying and the factors affecting this type of learning. Indeed, several important features of growing literature with the same paradigms are important to mention in relation to the hypothesis

outlined here. One critical factor is that statistical learning is a simple adaptive capacity that can also be found in other animals that have much less developed prefrontal cortices. For example, rodents exposed to the same type of linguistic speech streams are able to correctly segment it, albeit with a somewhat different computation (Toro and Trobalón, 2005). Another important feature is that the presence of statistical regularities in the input captures attention (Turk-Browne et al., 2005). Thus, even when no effort to learn is given, regularities can be extracted from the input (Saffran et al., 1997), capturing our attention in an automatic manner, which is consistent with the fact that even newborns are able to detect these statistical regularities (Teinonen et al., 2009). Therefore, the development of endogenous attention is not necessary for this learning to occur. In the same vein, electrophysiological evidence indicates that once words are segmented, the recognition/detection of a known word within the speech stream also captures attention (Sanders et al., 2002; Parise et al., 2010; de Diego-Balaguer et al., 2015), enhancing the long-term memorization of the segmented word forms.

Another important fact underscores the importance of exogenous attention in these early learning stages and highlights the adaptive function of the unavailability of the endogenous system in young infants. In adults, the manipulation of diverted attention, orienting endogenous attention outside the speech stream, can interfere with teach (Toro et al., 2005). Adults and older infants can orient their attention endogenously, and although this can be helpful when it converges to track the critical information for learning, it can also interfere with learning when it diverts from the correct focus of attention. For example, attention diverted from the dependency by the attraction of novel words that need to be ignored prevents non-adjacent learning. Infants' ability to generalize the detection of non-adjacent dependency to nonsense stems (e.g., *These meeps*) occurs only if they are first presented with familiar stems (e.g., *These chairs*). When attention is captured by a novel intervening element, learning of non-adjacent dependencies is altered (Soderstrom et al., 2002). In contrast, if salient information automatically captures infants' attention and this information is helpful for learning, the absence of endogenous attention prevents infants from disengaging and reorienting their attention to a different focus of attention that may interfere with the correct computation. In this way, the early dominance of this automatic exogenous mechanism can make learning more likely to occur. Other salient features, such as adjacent repetitions, can also act as important attentional attractors improving learning. Already present at birth, infants possess an automatic perceptual mechanism to detect repetitions in the auditory domain (Endress et al., 2009). This is reflected in greater activation in the temporal and left frontal brain areas when tested for recognition after exposure to simple repetition-based structures (ABB; e.g., "mubaba," "penana") than to random sequences (ABC; e.g., "mubage," "penaku").

Overall, the evidence indicates that the characteristics of speech with their statistical regularities and the salient prosodic cues are perfectly adapted to make the most of the early availability of exogenous attention. By engaging exogenous stimulus-driven attention, available since birth, learning can be

achieved. The absence of control of voluntary attention at these early stages of development does not limit infants' ability to acquire language but rather helps them by allowing infants to follow their stimulus-driven mechanism to capture the relevant information for learning automatically.

## Early Signs of Non-Adjacent Dependency Learning in the Rise of Endogenous Attention

Although the ability to segment and extract words from speech is a critical milestone of language acquisition, to acquire the grammar of their language infants must also track non-adjacent relationships. Importantly, the extraction of hierarchical structures relies on temporally distant relationships and is fundamental to capture the properties of language (Chomsky, 1957). Nonadjacent dependencies refer to cases in which two elements co-occur over one or more intervening elements. In natural languages, for example, in English, there is an association between auxiliaries and inflectional morphemes, irrespective of the intervening verb stem (e.g., **is walking; is running; is eating**). Infants must dismiss the variable irrelevant information and focus instead on the invariant relevant cues that predict the non-adjacent dependency (Gómez and Maye, 2005).

Because the endogenous system appears progressively in the course of development, its initial use in its earliest stages depends on the convergent presence of exogenous cues. The first signs of non-adjacent tracking in language are observed in the phonological domain where the presence of exogenous cues helps infants to track the dependencies grouped by their high similarity (for a review, see Sandoval and Gómez, 2013). For example, infants as young as 7 months can use harmony on vowel which are more salient than consonants and linked to prosodic variations as a cue to find word boundaries (Mintz and Walker, 2006; Kanpam et al., 2008) but cannot use consonantal harmony (Nazzi et al., 2009; Gonzalez-Gomez and Nazzi, 2012). They need to reach 10 months of age before they can, an age where endogenous attention starts to be more prominent. In a similar vein, newborns can discriminate adjacent rules based on the repetition of the same syllable but not when rules are non-adjacent (ABA; e.g., "bamuba," "napena") (Gervain et al., 2008). In contrast, 7-month-old and older infants track non-adjacent dependencies but only under some circumstances; when non-adjacent syllables are identical and the interleaved syllables are different (e.g., *le di le, ga po ga*) (Marcus et al., 1999; Gerken, 2006). Unexpectedly, a more recent study demonstrated that German infants as young as 4 months of age could discriminate between grammatical and ungrammatical non-adjacent dependencies in Italian (Friederici et al., 2011). As the authors indicate, Italian morphosyntactic dependencies also contain phonological dependencies. Given that phonological, non-adjacent dependencies are tracked from very early stages in development, it has been proposed that 4-month-olds may be tracking the phonological aspects to discriminate these non-adjacent dependencies. In other words, the exogenous attentional resources already available to 4-month-olds could have driven the success of such young infants in this task.

## Learning more Challenging Non-Adjacent Dependencies with Greater Maturation of Endogenous Attention

Although young infants can track non-adjacent linguistic rules under certain learning conditions (e.g., when the dependent units are similar and the intermediate elements are dissimilar), learning of morphosyntactic dependency appears several months after phonological dependency learning has occurred. This non-adjacent dependency learning that appears in more challenging perceptual and linguistic arrangements, requires greater involvement of endogenous mechanisms. Simply tracking non-adjacent dependencies can be used to locate word boundaries but is not enough to extract and generalize the underlying rule that entails the creation of abstract categories for generalization (Peña et al., 2002).

In this context, prosodic information in natural languages provides reliable cues not only for word segmentation but also for rule learning (Juszyk, 2002) because prosodic pauses tend to co-occur with syntactic boundaries. Nevertheless, although prosodic cues play a role in word segmentation from birth, it is not until the first year that infants start exploiting these cues for rule extraction (Johnson, 2008; Seidl and Johnson, 2008). The presence of this prosodic information in an artificial language enhances the extraction of non-adjacent dependencies compared to continuous speech streams without pauses (Peña et al., 2002). An important point to highlight is that the presence of pauses *per se* does not improve learning of non-adjacent dependencies. Those pauses need to occur at the boundaries of the position of the dependencies to be useful (Endress et al., 2005; Mueller et al., 2010). The use of these cues (stress pattern or prosodic pauses) for word segmentation requires only orienting attention to the position of the prosodic information that captured attention, that is, in an exogenous fashion. However, the use of prosodic pauses for rule extraction additionally requires the use of this cue to selectively focus attention on concurrent phonological information at this specific position. This cue then has to be used as a relevant predictor of forthcoming information to extract the rule dependency, which implies focusing attention to this cue and the predicted element while disregarding the intervening irrelevant information.

Within the morphosyntactic domain, nonadjacent relationships are often found between subject and verb agreement (**he** walks) or between auxiliary and verb agreement (**he is** walking). Learning of morphosyntactic, non-adjacent dependencies emerges after the first year of life (Gómez and Maye, 2005). This developmental course is reasonable when considering the challenge of the task, that is, in order to track the dependency among non-adjacent elements, infants must first identify the morphemes without involving any given similarity and then track the dependency between them across intervening elements irrelevant to the rule dependency (**he** walks; **he** runs; **he** eats). In one of the first studies exploring infants' ability to learn verb-tense agreement (Santelmann and Juszyk, 1998), researchers reported that 18-month-olds accepted grammatical phrases in English, such as "is running,"

and rejected ungrammatical phrases, such as "can running," whereas 15-month-olds were not able to differentiate between the phrases. Moreover, learning was possible only under certain conditions, with infants succeeding when the intervening element extended three syllables or less (e.g., *Grandma is always singing*, but not *Grandma is almost always singing*).

In addition, in order to learn a non-adjacent relation of the form "**these** cats," infants must track a dependency between two elements that occur over an intervening element and create different categories (e.g., determiner, noun, verb). Several lines of research have explored the mechanisms underlying the ability of grouping elements into categories. For example, it has been proposed that *frequent frames* (e.g., "these × are") yield category formation by their frequent co-occurrence with intervening content words and constitute the basis for the creation of grammatical categories (Mintz, 2003). Gómez and Maye (2005) showed that 15- and 18-month-old succeed when frames have high variability in the intervening word inside the frame but failed with low variability which is in agreement with the frame-based categorization proposed by Mintz (2003). Similarly, increasing the variability of the irrelevant intervening information makes adjacent relations less statistically informative and the non-adjacent dependency more prominent allowing learners to focus on the relevant and reliable relationship among non-adjacent elements (Gómez, 2002). Interestingly these different studies converge in a similar age between 15 and 18 months old as Gerken et al. (2011) where these authors found that infants use selective attention to focus on languages having learnable grammatical patterns. These studies converge to the parallelism between the development of non-adjacent dependency learning, category formation and endogenous control of attention in the second year of life. The importance of correct tuning in attention for the acquisition of non-adjacent rules is also seen in the Lany and Gómez (2008) study, where infants younger than those of the previous studies were able to track non-adjacent dependencies if the correct attention focus was guided by training them first on the dependencies between categories. Infants later discriminated grammatical and ungrammatical items involving non-adjacent dependencies with the same category words.

Thus, the overall pattern in agreement with the progressive ability to orient attention endogenously and the close collaboration between exogenous and endogenous attention. Early on, infants need more concurrent exogenous cues such as high degree of similarity, same identity between the dependent pairs (Creel et al., 2004; Onnis et al., 2005) or prior exposure to them (Lany and Gómez, 2008; Lai and Poletiek, 2011) (for a review, see Perruchet et al., 2012), to help them to orient their attention to the relevant information (Pacton and Perruchet, 2008; Pacton et al., 2015), allowing a greater interaction between exogenous and endogenous attention. At later stages of development, after the first year of age, the improved endogenous system allows infants to rely less on the availability of these salient features to orient their attention to the relevant information.

An important point to consider is that signs of discrimination of more complex non-adjacent dependencies at a very early age have been observed only in electrophysiological studies

(Mueller et al., 2012). Online EEG measures may not reflect the same knowledge as more overt behavioral responses that require greater explicit knowledge. In that sense, it is worth considering that indicators of prediction present from birth are reflected in mismatched responses in the EEG at the presentation of unexpected events, and these early online effects reflect these more automatic prediction mechanisms. However, recent research has shown that electrophysiological indexes of conscious access, equivalent to the P300 in adults, that show a non-linear pattern, can only be tracked clearly at the end of the first year. This response associated to consciousness was visible and sustained from 12 to 15 months of age (750 ms) and may serve to amplify the sensory input through selective attention (Kouider et al., 2013). Conscious access before the first year of age may not be possible because even if the structural architecture is in place, its immaturity may not allow an adequate flux of information for conscious availability (Dehaene-Lambertz and Spelke, 2015). From the perspective presented here, this conscious access may be required for these predictions to reach a long-lasting representation that may allow the infant to show behavioral effects. More studies are needed to examine early computation of different types of non-adjacent dependencies in infancy. New research should take into account the role that variables attracting attention may have in their acquisition in order to understand when the capacity actually arises in development.

## BRAIN DEVELOPMENT OF THE ATTENTION AND LANGUAGE NETWORKS

In terms of brain development, the parallel maturation of the attention and language network is also evident. This is in part unavoidable given the partial overlap between those two networks. As we have previously mentioned, a fronto-parietal network with either more ventral or dorsal distribution is related to stimulus-driven and goal-directed attention mechanisms, respectively. These areas are connected through the superior longitudinal fascicle (SLF), and the ventral and dorsal connectivity is ensured through the SLF III branch and I branch of this fascicle, respectively; these two connections have been proposed to interact through the SLF II, which connects the dorsal regions of the PFC to the ventral regions of the parietal lobe (de Schotten et al., 2011; **Figure 2**, left).

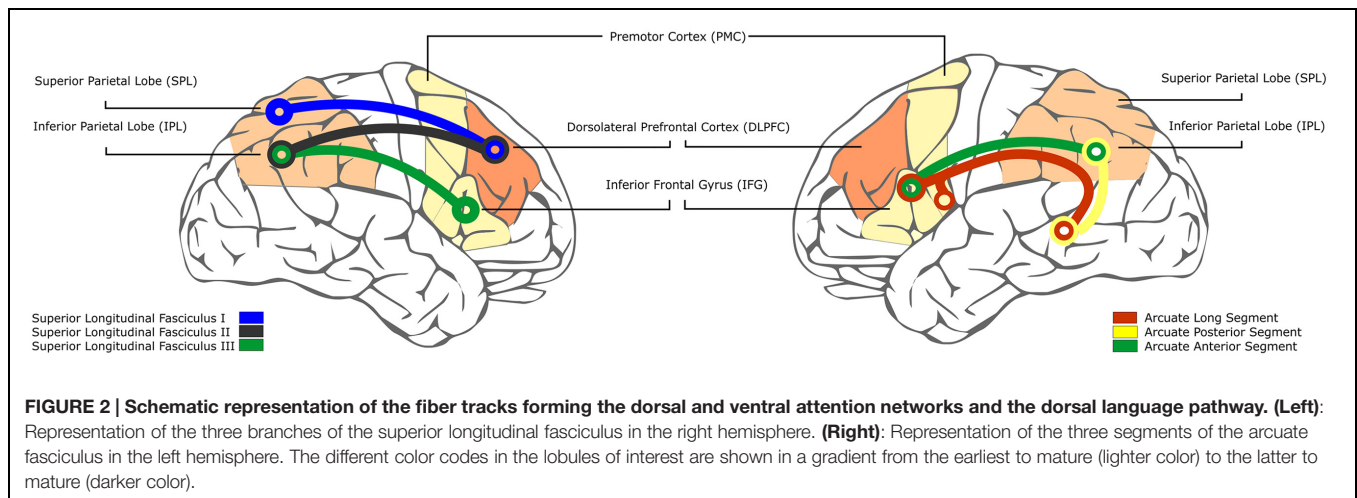
For language, a division in ventral and dorsal pathways ensures audio-motor integration and language production dorsally and language comprehension and semantic processing ventrally (Hickok and Poeppel, 2007). Direct connections between the language-related areas in the left frontal and temporal cortices are sustained dorsally through the arcuate fasciculus (AF) and ventrally by paths running through the extreme capsule (Saur et al., 2008; Brauer et al., 2011, 2013). The dorsal connection is also assured indirectly through the parietal lobe with shorter segments (Catani et al., 2005): an anterior segment connecting the premotor and inferior frontal regions with the IPL and a posterior segment connecting inferior

parietal and temporal cortices (**Figure 2**, right). There is some controversy concerning the terminations of the AF (Dick and Tremblay, 2012, for a review). Interestingly, this bundle overlaps with the SLF III, previously mentioned in relation to the ventral attention network, the anterior segment of the AF and the SLF III having a greater right lateralization (Catani et al., 2005; López-Barroso et al., 2013). Although the same nomenclature is used for attention and language in terms of ventral and dorsal streams, only the ventral attention and dorsal language stream overlap (see **Figure 2**). Although we have based this section on models of attention based on visual attention, Corbetta et al. (2008) did mention that the ventral attention network responds to different modalities. The overlap between the ventral attention and dorsal language networks is even greater if we consider that the temporal attention network shows a greater left functional lateralization, pointing again to the importance of temporal attention in speech processing (Coull et al., 2011).

During development, studies of whole ventral language connections demonstrated that newborns exhibit an adult-like ventral connection between the frontal and temporal lobes, and even children at 7 years of age have a preferential use of this pathway for sentence comprehension, in contrast to adults, who preferentially use the dorsal pathway (Brauer et al., 2011). In contrast, the dorsal pathway follows different developmental courses, with two subparts maturing at different rates. Whereas the dorsal connections reaching the premotor cortex are functional at birth, the terminations of the dorsal pathway reaching the posterior portion of Broca's area (BA 44) are still underdeveloped (Perani et al., 2011) and are not fully myelinated at the age of seven (Brauer et al., 2011, 2013). Adult studies on the learning of new languages indicate that whereas the audio-motor subpart is related to word learning (López-Barroso et al., 2013), the processing of non-adjacent elements relies on the latter subpart running from frontal BA 44 (Friederici, 2011). This pathway, which may support hierarchical (non-local) dependencies (Boeckx et al., 2014), follows a later and slower rate of development, similarly to non-adjacent, dependency rule learning.

Rapid changes are observed during the first year of life in terms of maturation. Sensory and motor systems myelinate earlier than brain systems serving higher level functions (Flechsig, 1920). Myelination starts at different times and occurs at different rates in different areas. At birth, there is little differentiation between gray and white matter in cortical areas. The primary visual cortex rapidly matures during the first 3 months with parallel myelination of optical radiations, whereas the primary auditory cortex and acoustic radiations extend over the first 3 years of life. The frontal areas and cortico-cortical connections continue to mature until puberty, but myelination is already observed during the first year in all associative regions. Diffusion measures increase with the compactness/myelination of the tracts in the left lower part of the cortico-spinal tract and in the parietal part of the AF relative to the right during the first 3 months of life. During this period the maturation of the right hemisphere is generally faster than the left (i.e., superior temporal sulcus, STS), but the inferior frontal gyrus (IFG) shows earlier left than right development. The left AF matures faster than the right





and correlates with the maturation of BA 44 and the posterior part of the STS (Dehaene-Lambertz and Spelke, 2015 for a review).

Postnatal maturation shows subcortical white matter expansion in the connections to the frontal, anterior temporal, and parietal cortices, as measured by diffusion imaging and volume expansion (Hill et al., 2010). Surface expansion reflects an underlying change in synaptogenesis, dendritic arborization, gliogenesis, and intracortical myelination. The lateral temporal and parietal lobes and the dorsal and medial prefrontal regions are functionally and structurally not mature at birth. They show high expansion in cortical folding in both hemispheres in infants compared with adults. The latest maturation in synaptic density, peak cortical thickness, and mature values of gray matter density are reached in the dorsolateral prefrontal cortex. The comparison between human and macaque monkey cortices reveals that these dorsal, medial frontal and lateral parietal cortices show correlated high postnatal and evolutionary expansion (Hill et al., 2010). This pattern suggests similar patterns of cortical expansion in the development and evolution of these areas, which points to the importance of these areas for human specific functions.

These changes in connectivity at the structural level are also reflected in functional connectivity. Graph-theoretic measures of infants' brains (Power et al., 2010) indicate that the developing functional networks are in some respects similar to adult networks. The necessary connections are present; however, the brain connectivity compared to adults tends to have strong resting state functional connectivity MRI (rs-fcMRI) signal correlations with nearby regions, even during childhood. The progressively local correlations tend to weaken, whereas correlations with more distant regions, such as those between the frontal and parietal cortices, tend to increase. This trend stems from synaptic pruning that contributes to reduced local rs-fcMRI correlation, and myelination that could facilitate increased long-range connectivity.

Considering the overall attention and language networks, brain regions and connections of overlap are observed between the ventral attention network and the dorsal language network

(Figure 2). Ventral prefrontal and insular regions integrating the ventral attention network and the anterior segment of the AF show an early availability, whereas the IPL and their connections show a later and more progressive development. This delayed development also affects the dorsal attention network with the dorsal prefrontal regions having a slow maturation extending to childhood and with delayed maturation of the parietal lobe (Casey et al., 2000; Fuster, 2001). There is evidence showing that the left IFG is engaged in the extraction of TPs (Karuza et al., 2013) as well as the PMC (Cunillera et al., 2009) when no other cue is available to segment speech (McNealy et al., 2006; Scott-Van Zeeland et al., 2010). The early functionality of the left IFG and the premotor cortex (PMC) allows early use of TPs and stimulus-driven attention to orient to salient prosodic information and to segment speech. When the dorsal prefrontal cortex starts to be maturationally functional during the second year of life (Colombo and Cheatham, 2006), the dorsal fronto-parietal network allows for more proficient control of attention. The later maturation of the dorsal prefrontal cortex (DLPFC) and part of the ventral attention network (i.e., IPL), including the temporo-parietal junction (TPJ) (see Figure 2), allows progressively to (i) orient the ventral attention network to task-relevant representations (e.g., phonemes of the native language and segmented words) created in the earlier stages of development, (ii) recruit goal-directed attention, (iii) optimal functioning of the attention system, that requires the effective interaction between the two networks through the TPJ and the DLPFC (Corbetta and Shulman, 2002; Corbetta et al., 2008), necessary to accurately and selectively attend to specific stimuli and shift the focus of attention when relevant stimulation appears.

## ATTENTION DEFICITS AND LANGUAGE DEVELOPMENT DISORDERS

The proposal delineated here makes a straightforward prediction in relation to the effects of attention deficits in language development. If control of attention is a function used for the

optimal acquisition of non-adjacent rule dependencies, then impairments in the development of this function should interfere with the acquisition of these rules. In contrast, early language development relying on more automatic attention mechanisms should not be affected.

Commonly, children acquire language rapidly and effortlessly. However, some children show problems acquiring language. Specific language impairment (SLI) is classically defined as a developmental disorder of language characterized by difficulty in acquiring language in the absence of neurological damage, hearing deficits, or intellectual disabilities (Bishop, 1992; Leonard, 1998). The prevalence of SLI in pre-school children is approximately 7% (Tomblin et al., 1997; Law et al., 2000). Longitudinal studies reveal that more than 70% of diagnosed cases of SLI in kindergarten persist into adulthood (Johnson et al., 1999). SLI children have been shown to have difficulties in the acquisition of non-adjacent dependencies (Hsu et al., 2014) and in the use of prosodic information for syntactic processing (Sabisch et al., 2009). In a longitudinal study, impaired prosodic processing of word stress during early development was shown to be an early marker of risk for SLI (Weber et al., 2005).

Linguistic impairments often co-occur with non-linguistic deficits, including attention-deficit/hyperactivity disorder (ADHD). Both SLI and ADHD frequently overlap within the same children, that is, comorbidity between the two disorders is commonly found (Baker and Cantwell, 1992; Benasich et al., 1993; Coster et al., 1999; Noterdaeme and Amorosa, 1999; Tomblin et al., 2000; Lindsay et al., 2007). ADHD is the most frequent diagnosis among children with language impairments (Cohen et al., 2000). Longitudinal studies suggest that SLI children have a profound risk for ADHD (Baker and Cantwell, 1987; Cantwell and Baker, 1987; Beitchman et al., 1989; Benasich et al., 1993; Redmond and Rice, 1998, 2002). More precisely, deficits in selective attention (Stevens et al., 2008) and sustained attention (Spaulding et al., 2008; Finneran et al., 2009) have been found in children with SLI. ADHD is a common childhood disorder characterized by a persistent pattern of inattention and/or developmentally inappropriate levels of hyperactivity/impulsivity (American Psychiatric Association, 2000). ADHD prevalence is approximately 10% in children (Faraone et al., 2003; Pastor and Reuben, 2008). As with SLI, children with ADHD are a highly heterogeneous group. ADHD is commonly divided into three subtypes: ADHD-Inattentive (ADHD-I), ADHD-Hyperactive-Impulsive (ADHD-H/I), and ADHD-Combined type (ADHD-C). Whereas children in the ADHD-I subgroup usually show difficulties with attention control, sustained attention and are often inattentive, ADHD-H children exhibit high levels of activity and poor impulse control. ADHD-I children do poorly in tasks requiring sustained attention, covert shifting of attention and selective attention. Thus, individual differences in the control of selective attention in infancy may be related to ADHD-I outcomes. Children in the ADHD-I group are more probable to meet criteria for learning disability than ADHD-H children (Willcutt and Pennington, 2000).

Similar to the findings on attention deficits found in SLI children, a similar pattern is present in ADHD children. Between

50 and 90% of children with ADHD have co-occurring language difficulties (Gualtieri et al., 1983; Camarata et al., 1988; Love and Thompson, 1988; Tirosh and Cohen, 1998). However, the overlap between these disorders shows an asymmetrical pattern, that is, more ADHD children have co-occurring SLI than SLI children have co-occurring ADHD (Tannock and Schachar, 1996). Higher order cognitive functions (e.g., executive functions, working memory, and attention) have been explored as possible causal deficits for SLI and ADHD disorders (e.g., Cardy et al., 2010; Hutchinson et al., 2012).

In SLI, abnormal diffusion measures are observed systematically in the SLF and AF (Verhoeven et al., 2012; Roberts et al., 2014). A more recent study showed also differences in the ventral language network (i.e., the inferior fronto-occipital fasciculus, IFOF) (Vydrova et al., 2015). The discrepancies between studies may stem from the heterogeneity of the disease with children with more semantico-pragmatic profiles that are more likely to show differences in the IFOF function and those with and without associated ADHD, which may cause an associated SLF abnormality in addition to the AF. The brain structures supporting cognitive functions commonly associated with ADHD have also been investigated. Gross anatomical changes in brain dimensions are often associated with ADHD, specifically, reduced dimensions of the caudate nucleus, the prefrontal cortex, the corpus callosum, and the cerebellar vermis (see Bush et al., 2005 for a review) and in the parietal lobes (Sowell et al., 2003) are found in ADHD. Evidence from pathophysiology research has shown that ADHD physiology involves dopaminergic and noradrenergic pathway dysfunction in the prefrontal cortex and subcortical regions of the brain (Barkley et al., 1992; Castellanos et al., 1996; Faraone and Biederman, 1998; Konrad et al., 2006). This network partially overlaps with both goal-directed attention and temporal processing. DA dysfunction affects mainly the dorsal regions of the PFC, which are those required for goal-directed attention. The subcortical regions affected (i.e., striatum) and the cerebellum are important structures for temporal processing (Coull et al., 2011).

Recent studies have provided the first evidence that temporal selective attention during speech perception predicts language outcome in preschool children. Children who selectively allocate attention to informative moments during speech, such as word onsets, demonstrate better metalinguistic capacity (Astheimer et al., 2014).

## CONCLUSION

Infants acquire language exceptionally fast and without any given instruction. But, how can infants so easily achieve such a remarkable landmark, whereas adults struggle to do so? Following Kuhl's view (Kuhl, 2004), understanding how the early brain is committed to the statistical and prosodic patterns experienced early in life helps to explain the long-standing puzzle of why infants are better language learners than adults. One of the possible answers is the way their cognitive

development is structured, with functions, such as attention, appearing in an incremental fashion and assisting language learning.

Based on the characteristics of the developmental trajectory of the attention and language systems, we have outlined the hypothesis that attention development, characterized by an initial phase when attention is stimulus-driven, followed by a progressive ability to endogenously control the focus of attention, shapes the developmental trajectory of language. In the evidence reviewed here, we have seen that the learning trajectory of two types of linguistic learning (words and rules) shows a different profile in infant language development. Whereas words in fluent speech are already segmented and extracted at early stages, non-adjacent dependencies occurring over temporally distant elements are learned many months later (see **Figure 1**).

More precisely, the early segmentation and word learning abilities is profoundly influenced by the salient characteristics of the speech signal, with an important role of prosodic information. Later acquisition of more complex information associated with the extraction of more distant dependencies is influenced by variables that help infants to focus attention on the relevant elements carrying the dependency and to disregard the information that is not relevant for the acquisition of the dependencies. This trajectory goes hand in hand with the development of the ability to progressively orient attention endogenously. Early in this phase, infants require more concurrent salient cues, such as phonological similarity or identity repetition, to help them to orient their attention to the relevant information. A greater development of goal-directed attention allows infants to learn less salient, non-adjacent dependencies by relying more on endogenous cues. In terms of brain development, whereas the initial stages of development rely on the availability of some areas of the ventral attention network, including the ventral prefrontal regions and the premotor cortex, the latter stages require the maturation of more dorsal prefrontal and parietal regions (see **Figure 2**).

We consider that this development of attention in different stages allows for an earlier simplification of learning. This early learning is driven by the automatic capture of attention, creating the first building blocks that learning can lean on when control of attention allows for the extraction of more complex relations between non-adjacent elements in speech. Data from adults show that they can track both adjacent and non-adjacent information at the same time, and one information can interfere with the other (Romberg and Saffran, 2013). Thus, the inability to reorient attention away from the automatic attractors of attention is valuable in the early stages of acquisition, allowing for incremental learning.

Moreover, the same exogenous system that allows young infants to extract words using salient cues may also help them to extract complex rules. Young infants are able to succeed in non-adjacent learning that otherwise would not be available after the first year of life. In these early stages, this success of non-adjacent dependency tracking

occurs only under certain conditions. Applying our present proposal to this developmental scenario, two main conditions should be fulfilled to extract non-adjacent dependencies in the early stages of development: (1) a rudimentary mechanism of endogenous attention should be available to select certain predictive elements and to disregard irrelevant information, and (2) stimulus-driven factors should be present in the linguistic input (e.g., certain degree of similarity or saliency) to automatically capture the exogenous attention system.

The implications of our hypothesis are clear in terms of the parallelism between the development of the endogenous attention system and the rule learning abilities in healthy infants. This relation is seen not only in healthy development but also in the effects of attention deficits in relation with impairments in language development. The importance of being able to exploit the available information given by exogenous cues, such as prosodic information, to orient attention endogenously is crucial not only in infant healthy development but also in studies with different pathologies.

Comprehending the cognitive processes involved in language development is of critical importance for our understanding of why, under certain conditions, language development impairment occurs. However, research in the field of language development often offers limited explanations bounded within the language domain, ignoring the importance of other cognitive functions. The present proposal overcomes these limits and presents an integrative approach to understand the role of attentional tuning during language acquisition. By reviewing the main stages of attention and language development and possible impairments, we have strengthened the importance of taking an interdisciplinary approach to the study of human development. We believe that this integrative approach exploring the role of temporal attention as a scaffold for language development can lead to a wider scope than previous proposals, allowing the development of a precise model of language and cognitive function interaction during learning that has important clinical and developmental consequences, hence providing an important contribution to the language learning and language rehabilitation fields.

## AUTHOR CONTRIBUTIONS

RD-B, AM-A provided the ideas and wrote the article. FP contributed to discussions, writing and checked that the infant research review was accurate.

## ACKNOWLEDGMENTS

The authors have been funded by an ERC StG (TuningLang 313841) to RD-B. They are thankful to Pablo Ripollés for his help with the figures and to Joan Orpella for helpful discussions during manuscript preparation.



## REFERENCES

- Adler, S. A., Haith, M. M., Arehart, D. M., and Lanthier, E. C. (2008). Infants' visual expectations and the processing of time. *J. Cogn. Dev.* 9, 1–25. doi: 10.1080/15248370701836568
- Althaus, N., and Mareschal, D. (2012). Using saliency maps to separate competing processes in infant visual cognition. *Child Dev.* 83, 1122–1128. doi: 10.1111/j.1467-8624.2012.01766.x
- American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders*, 4th Edn. Washington, DC: American Psychiatric Association.
- Astheimer, L., Janus, M., Moreno, S., and Bialystok, E. (2014). Electrophysiological measures of attention during speech perception predict metalinguistic skills in children. *Dev. Cogn. Neurosci.* 7, 1–12. doi: 10.1016/j.dcn.2013.10.005
- Baker, L., and Cantwell, D. P. (1987). A prospective psychiatric follow-up of children with speech/language disorders. *J. Am. Acad. Child Adolesc. Psychiatry* 26, 546–553. doi: 10.1097/00004583-198707000-00015
- Baker, L., and Cantwell, D. P. (1992). Attention deficit disorder and speech/language disorders. *Compr. Ment. Health Care* 2, 3–16.
- Baldwin, D. A. (1991). Infants' contribution to the achievement of joint reference. *Child Dev.* 62, 875–890. doi: 10.2307/1131140
- Baldwin, D. A. (1993). Infants' ability to consult the speaker for clues to word reference. *J. Child Lang.* 20, 395–418. doi: 10.1017/S0305000900008345
- Barkley, R. A., Grodzinsky, G., and DuPaul, G. J. (1992). Frontal lobe functions in attention deficit disorder with and without hyperactivity: a review and research report. *J. Abnorm. Child Psychol.* 20, 163–188. doi: 10.1007/BF00916547
- Barnes, R., and Jones, M. R. (2000). Expectancy, attention, and time. *Cogn. Psychol.* 41, 254–311. doi: 10.1006/cogp.2000.0738
- Beitchman, J. H., Hood, J., Rochon, J., and Peterson, M. (1989). Empirical classification of speech/language impairment in children. II. Behavioral characteristics. *J. Am. Acad. Child Adolesc. Psychiatry* 28, 118–123. doi: 10.1097/00004583-198901000-00022
- Benasich, A. A., Curtiss, S., and Tallal, P. (1993). Language, learning, and behavioral disturbances in childhood: a longitudinal perspective. *J. Am. Acad. Child Adolesc. Psychiatry* 32, 585–594. doi: 10.1097/00004583-199305000-00015
- Bishop, D. V. (1992). The underlying nature of specific language impairment. *J. Child Psychol. Psychiatry* 33, 3–66. doi: 10.1111/j.1469-7610.1992.tb00858.x
- Boeckx, C., Martínez-Alvarez, A., and Leivada, E. (2014). The functional neuroanatomy of serial order in language. *J. Neurolinguistics* 32, 1–15. doi: 10.1016/j.jneuroling.2014.07.001
- Brannon, E. M., Libertus, M. E., Meck, W. H., and Woldorff, M. G. (2008). Electrophysiological measures of time processing in infant and adult brains: Weber's Law holds. *J. Cogn. Neurosci.* 20, 193–203. doi: 10.1162/jocn.2008.20016
- Brannon, E. M., Roussel, L. W., Meck, W. H., and Woldorff, M. (2004). Timing in the baby brain. *Brain Res. Cogn. Brain Res.* 21, 227–233. doi: 10.1016/j.cogbrainres.2004.04.007
- Brauer, J., Anwender, A., and Friederici, A. D. (2011). Neuroanatomical prerequisites for language functions in the maturing brain. *Cereb. Cortex* 21, 459–466. doi: 10.1093/cercor/bhq108
- Brauer, J., Anwender, A., Perani, D., and Friederici, A. D. (2013). Dorsal and ventral pathways in language development. *Brain Lang.* 127, 289–295. doi: 10.1016/j.bandl.2013.03.001
- Brooks, R., and Meltzoff, A. N. (2002). NIH public access. *Dev. Psychol.* 38, 958–966. doi: 10.1016/j.micinf.2011.07.011
- Bruner, J. (1983). *Child's Talk: Learning to Use Language*. New York: Academic Press.
- Bush, G., Valera, E. M., and Seidman, L. J. (2005). Functional neuroimaging of attention-deficit/hyperactivity disorder: a review and suggested future directions. *Biol. Psychiatry* 57, 1273–1284. doi: 10.1016/j.biopsych.2005.01.034
- Camarata, S. M., Hughes, C. A., and Ruhl, K. L. (1988). Mild/moderate behaviorally disordered students: a population at risk for language disorders. *Lang. Speech Hear. Serv. Sch.* 19, 191–200. doi: 10.1044/0161-1461.1902.191
- Cantwell, D. P., and Baker, L. (1987). Prevalence and type of psychiatric disorder and developmental disorders in three speech and language groups. *J. Commun. Disord.* 20, 151–160. doi: 10.1016/0021-9924(87)90006-2
- Cardy, J. E. O., Tannock, R., Johnson, A. M., and Johnson, C. J. (2010). The contribution of processing impairments to SLI: insights from attention-deficit/hyperactivity disorder. *J. Commun. Disord.* 43, 77–91. doi: 10.1016/j.jcomdis.2009.09.003
- Carpenter, M., Nagell, K., and Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monogr. Soc. Res. Child Dev.* 63, i–vi, 1–143. doi: 10.2307/1166214
- Casey, B. J., Davidson, M. C., Hara, Y., Thomas, K. M., Martinez, A., Galvan, A., et al. (2004). Early development of subcortical regions involved in non-cued attention switching. *Dev. Sci.* 7, 534–542. doi: 10.1111/j.1467-7687.2004.00377.x
- Casey, B. J., Giedd, J. N., and Thomas, K. M. (2000). Structural and functional brain development and its relation to cognitive development. *Biol. Psychol.* 54, 241–257. doi: 10.1016/S0301-0511(00)00058-2
- Castellanos, F. X., Giedd, J. N., Marsh, W. L., Hamburger, S. D., Vaituzis, A. C., Dickstein, D. P., et al. (1996). Quantitative brain magnetic resonance imaging in attention-deficit hyperactivity disorder. *Arch. Gen. Psychiatry* 53, 607–616. doi: 10.1001/archpsyc.1996.01830070053009
- Catani, M., Jones, D. K., and Ffytche, D. H. (2005). Perisylvian language networks of the human brain. *Ann. Neurol.* 57, 8–16. doi: 10.1002/ana.20319
- Chica, A. B., Bartolomeo, P., and Lupiáñez, J. (2013). Two cognitive and neural systems for endogenous and exogenous spatial attention. *Behav. Brain Res.* 237, 107–123. doi: 10.1016/j.bbr.2012.09.027
- Chomsky, N. (1957). *Syntactic Structures*. The Hague/Paris: Mouton.
- Christophe, A., Millotte, S., Bernal, S., and Lidz, J. (2008). Bootstrapping lexical and syntactic acquisition. *Lang. Speech* 51, 61–75. doi: 10.1177/00238309080510010501
- Cohen, N. J., Vallance, D. D., Barwick, M., Im, N., Menna, R., Horodezky, N. B., et al. (2000). The interface between ADHD and language impairment: an examination of language, achievement, and cognitive processing. *J. Child Psychol. Psychiatry* 41, 353–362. doi: 10.1111/1469-7610.00619
- Colombo, J. (2001). The development of visual attention in infancy. *Annu. Rev. Psychol.* 52, 337–367. doi: 10.1146/annurev.psych.52.1.337
- Colombo, J., and Cheatham, C. L. (2006). The emergence and basis of endogenous attention in infancy and early childhood. *Adv. Child Dev. Behav.* 34, 283–322. doi: 10.1016/S0065-2407(06)80010-8
- Cooper, R. P., and Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Dev.* 61, 1584–1595. doi: 10.2307/1130766
- Corbetta, M., Patel, G., and Shulman, G. L. (2008). The reorienting system of the human brain: from environment to theory of mind. *Neuron* 58, 306–324. doi: 10.1016/j.neuron.2008.04.017
- Corbetta, M., and Shulman, G. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215. doi: 10.1038/nrn755
- Correa, Á., Lupiáñez, J., Madrid, E., and Tudela, P. (2006). Temporal attention enhances early visual processing: a review and new evidence from event-related potentials. *Brain Res.* 1076, 116–128. doi: 10.1016/j.brainres.2005.11.074
- Coster, F. W., Goorhuis-Brouwer, S. M., Nakken, H., and Spelberg, H. C. (1999). Specific language impairments and behavioural problems. *Folia Phoniatr. Logop.* 51, 99–107. doi: 10.1159/000021484
- Coull, J. T., Cheng, R.-K., and Meck, W. H. (2011). Neuroanatomical and neurochemical substrates of timing. *Neuropsychopharmacology* 36, 3–25. doi: 10.1038/npp.2010.113
- Courage, M. L., Reynolds, G. D., and Richards, J. E. (2006). Infants' attention to patterned stimuli: developmental change from 3 to 12 months of age. *Child Dev.* 77, 680–695. doi: 10.1111/j.1467-8624.2006.00897.x
- Courage, M. L., and Richards, J. E. (2008). "Attention," in *Encyclopedia of Infant and Early Childhood Development*, eds M. M. Haith and J. B. Benson (Oxford: Elsevier), 106–117.
- Creel, S. C., Newport, E. L., and Aslin, R. N. (2004). Distant melodies: statistical learning of nonadjacent dependencies in tone sequences. *J. Exp. Psychol. Learn. Mem. Cogn.* 30, 1119–1130. doi: 10.1037/0278-7393.30.5.1119
- Cunillera, T., Càmar, E., Toro, J. M., Marco-Pallares, J., Sebastián-Gallés, N., Ortiz, H., et al. (2009). Time course and functional neuroanatomy of speech segmentation in adults. *Neuroimage* 48, 541–553. doi: 10.1016/j.neuroimage.2009.06.069
- de Diego-Balaguer, R., Rodríguez-Fornells, A., and Bachoud-Lévi, A. C. (2015). Prosodic cues enhance rule learning by changing speech



- segmentation mechanisms. *Front. Psychol.* 6:1478. doi: 10.3389/fpsyg.2015.01478
- de Schotten, M. T., Dell'Acqua, F., Forkel, S. J., Simmons, A., Vergani, F., Murphy, D. G. M., et al. (2011). A lateralized brain network for visuospatial attention. *Nat. Neurosci.* 14, 1245–1246. doi: 10.1038/nn.2905
- Dehaene, S., Meyniel, F., Wacongne, C., Wang, L., and Pallier, C. (2015). The neural representation of sequences: from transition probabilities to algebraic patterns and linguistic trees. *Neuron* 88, 2–19. doi: 10.1016/j.neuron.2015.09.019
- Dehaene-Lambertz, G., and Spelke, E. S. (2015). The infancy of the human brain. *Neuron* 88, 93–109. doi: 10.1016/j.neuron.2015.09.026
- Demany, L., Mckenzie, B., and Vurpillot, E. (1977). Rhythm perception in early infancy. *Nature* 266, 718–719. doi: 10.1038/266718a0
- Diamond, A. (2007). Consequences of variations in genes that affect dopamine in prefrontal cortex. *Cereb. Cortex* 17, 1–19. doi: 10.1093/cercor/bhm082
- Dick, A. S., and Tremblay, P. (2012). Beyond the arcuate fasciculus: consensus and controversy in the connectional anatomy of language. *Brain* 135, 3529–3550. doi: 10.1093/brain/aws222
- Doherty, J. R., Rao, A., Mesulam, M. M., and Nobre, A. C. (2005). Synergistic effect of combined temporal and spatial expectations on visual attention. *J. Neurosci.* 25, 8259–8266. doi: 10.1523/JNEUROSCI.1821-05.2005
- Elison, J. T., Wolff, J. J., Heimer, D. C., Paterson, S. J., Gu, H., Hazlett, H. C., et al. (2013). Frontolimbic neural circuitry at 6 months predicts individual differences in joint attention at 9 months. *Dev. Sci.* 16, 186–197. doi: 10.1111/desc.12015
- Endress, A. D., Nespore, M., and Mehler, J. (2009). Perceptual and memory constraints on language acquisition. *Trends Cogn. Sci.* 13, 348–353. doi: 10.1016/j.tics.2009.05.005
- Endress, A. D., Scholl, B. J., and Mehler, J. (2005). The role of salience in the extraction of algebraic rules. *J. Exp. Psychol. Gen.* 134, 406–419. doi: 10.1037/0096-3445.134.3.406
- Fair, D. A., Cohen, A. L., Power, J. D., Dosenbach, N. U. F., Church, J. A., Miezin, F. M., et al. (2009). Functional brain networks develop from a “local to distributed” organization. *PLoS Comput. Biol.* 5:e1000381. doi: 10.1371/journal.pcbi.1000381
- Fair, D. A., Dosenbach, N. U. F., Church, J. A., Cohen, A. L., Brahmbhatt, S., Miezin, F. M., et al. (2007). Development of distinct control networks through segregation and integration. *Proc. Natl. Acad. Sci. U.S.A.* 104, 13507–13512. doi: 10.1073/pnas.0705843104
- Faraone, S. V., and Biederman, J. (1998). Neurobiology of attention-deficit hyperactivity disorder. *Biol. Psychiatry* 44, 951–958. doi: 10.1016/S0006-3223(98)00240-6
- Faraone, S. V., Sergeant, J., Gillberg, C., and Biederman, J. (2003). The worldwide prevalence of ADHD: is it an American condition? *World Psychiatry* 2, 104–113.
- Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behav. Dev.* 8, 181–195. doi: 10.1016/S0163-6383(85)80005-9
- Finneran, D. A., Francis, A. L., and Leonard, L. B. (2009). Sustained attention in children with specific language impairment (SLI). *J. Speech Lang. Hear. Res.* 52, 915–929. doi: 10.1044/1092-4388(2009/07-0053)
- Flechsig, P. (1920). *Anatomie des Menschlichen Gehirns und Rückenmarks auf Myelogenetischer Grundlage*. Leipzig: George Thomas.
- Friederici, A. D. (2011). The brain basis of language processing: from structure to function. *Physiol. Rev.* 91, 1357–1392. doi: 10.1152/physrev.00006.2011
- Friederici, A. D., Mueller, J. L., and Oberecker, R. (2011). Precursors to natural grammar learning: preliminary evidence from 4-month-old infants. *PLoS ONE* 6:e17920. doi: 10.1371/journal.pone.0017920
- Fuster, J. M. (2001). The prefrontal cortex—an update: time is of the essence. *Neuron* 30, 319–333. doi: 10.1016/S0896-6273(01)00285-9
- Gao, W., Zhu, H., Giovanello, K., Smith, K., Shen, D., Gilmore, J., et al. (2009). “Emergence of the brain’s default network: evidence from two-week-old to four-year-old healthy pediatric subjects,” in *Proceedings of the 17th Scientific Meeting of the International Society for Magnetic Resonance in Medicine*, Honolulu, HI, 3434.
- Gerken, L. (2006). Decisions, decisions: infant language learning when multiple generalizations are possible. *Cognition* 98, 67–74. doi: 10.1016/j.cognition.2005.03.003
- Gerken, L., Balcomb, F. K., and Minton, J. L. (2011). Infants avoid “labouring in vain” by attending more to learnable than unlearnable linguistic patterns. *Dev. Sci.* 14, 972–979. doi: 10.1111/j.1467-7687.2011.01046.x
- Gervain, J., Macagno, F., Cogoi, S., Peña, M., and Mehler, J. (2008). The neonate brain detects speech structure. *Proc. Natl. Acad. Sci. U.S.A.* 105, 14222–14227. doi: 10.1073/pnas.0806530105
- Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., et al. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proc. Natl. Acad. Sci. U.S.A.* 101, 8174–8179. doi: 10.1073/pnas.0402680101
- Gómez, R., and Maye, J. (2005). The developmental trajectory of nonadjacent dependency learning. *Infancy* 7, 183–206. doi: 10.1207/s15327078in0702\_4
- Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychol. Sci.* 13, 431–436. doi: 10.1111/1467-9280.00476
- Gonzalez-Gomez, N., and Nazzi, T. (2012). Acquisition of nonadjacent phonological dependencies in the native language during the first year of life. *Infancy* 17, 498–524. doi: 10.1111/j.1532-7078.2011.00104.x
- Gualtieri, C. T., Koriath, U., Van Bourgondien, M., and Saleeb, N. (1983). Language disorders in children referred for psychiatric services. *J. Am. Acad. Child Psychiatry* 22, 165–171. doi: 10.1016/S0002-7138(09)62330-X
- Haith, M. M., Hazan, C., and Goodman, G. S. (1988). Expectation and anticipation of dynamic visual events by 3.5-month-old babies. *Child Dev.* 59, 467–479. doi: 10.2307/1130325
- Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8, 393–402. doi: 10.1038/nrn2113
- Hill, J., Inder, T., Neil, J., Dierker, D., Harwell, J., and Van Essen, D. (2010). Similar patterns of cortical expansion during human development and evolution. *Proc. Natl. Acad. Sci. U.S.A.* 107, 13135–13140. doi: 10.1073/pnas.1001229107
- Hood, B. M. (1995). “Shifts of visual attention in the human infant: a neuroscientific approach,” in *Advances in Infancy Research*, Vol. 9, eds C. Rovee-Collier and L. P. Lipsitt (Norwood, NJ: Ablex), 163–216.
- Hsu, H. J., Tomblin, J. B., and Christiansen, M. H. (2014). Impaired statistical learning of non-adjacent dependencies in adolescents with specific language impairment. *Front. Psychol.* 5:175. doi: 10.3389/fpsyg.2014.00175
- Hunnius, S., and Geuze, R. H. (2004). Gaze shifting in infancy: a longitudinal study using dynamic faces and abstract stimuli. *Infant Behav. Dev.* 27, 397–416. doi: 10.1016/j.infbeh.2004.02.003
- Hutchinson, E., Bavin, E., Efron, D., and Sciberras, E. (2012). A comparison of working memory profiles in school-aged children with Specific Language Impairment, Attention Deficit/Hyperactivity Disorder, Comorbid SLI and ADHD and their typically developing peers. *Child Neuropsychol.* 18, 190–207. doi: 10.1080/09297049.2011.601288
- Ibbotson, P., Lieven, E. V. M., and Tomasello, M. (2013). The attention-grammar interface: eye-gaze cues structural choice in children and adults. *Cogn. Linguist.* 24, 457–481. doi: 10.1515/cog-2013-0020
- Jackendoff, R. (2002). *Foundations of Language*. Oxford: Oxford University Press.
- Johnson, C. J., Beitchman, J. H., Young, A., Escobar, M., Atkinson, L., Wilson, B., et al. (1999). Fourteen-year follow-up of children with and without speech/language impairments: speech/language stability and outcomes. *J. Speech Lang. Hear. Res.* 42, 744–760. doi: 10.1044/jslhr.4203.744
- Johnson, E. K. (2008). Infants use prosodically conditioned acoustic-phonetic cues to extract words from speech. *J. Acoust. Soc. Am.* 123, EL144–EL148. doi: 10.1121/1.2908407
- Johnson, K. A., Burrows, E., and Coull, J. T. (2015). Children can implicitly, but not voluntarily, direct attention in time. *PLoS ONE* 10:e0123625. doi: 10.1371/journal.pone.0123625
- Johnson, M. H. (1990). Cortical maturation and the development of visual attention in early infancy. *J. Cogn. Neurosci.* 2, 81–95. doi: 10.1162/jocn.1990.2.2.81
- Jusczyk, P. W. (2002). How infants adapt speech-processing capacities to native-language structure. *Curr. Dir. Psychol. Sci.* 11, 15–18. doi: 10.1111/1467-8721.00159
- Kanpam, A., Van Parmaksiz, G., Van De Vijver, R., and Hohle, B. (2008). “Metrical and statistical cues for word segmentation: vowel harmony and word stress as cues to word boundaries by 6-and-9-month-old Turkish learners,” in *Language Acquisition and Development*, eds A. Gavarró and M. J. Freitas (Newcastle: Cambridge Scholars Publishing), 313–324.

- Karuz, E. A., Newport, E. L., Aslin, R. N., Starling, S. J., Tivarus, M. E., and Bavelier, D. (2013). The neural correlates of statistical learning in a word segmentation task: an fMRI study. *Brain Lang.* 127, 46–54. doi: 10.1016/j.bandl.2012.11.007
- Konrad, K., Neufang, S., Hanisch, C., Fink, G. R., and Herpertz-Dahlmann, B. (2006). Dysfunctional attentional networks in children with attention deficit/hyperactivity disorder: evidence from an event-related functional magnetic resonance imaging study. *Biol. Psychiatry* 59, 643–651. doi: 10.1016/j.biopsych.2005.08.013
- Kotz, S. A., and Schwartz, M. (2010). Cortical speech processing unplugged: a timely subcortico-cortical framework. *Trends Cogn. Sci.* 14, 392–399. doi: 10.1016/j.tics.2010.06.005
- Kouider, S., Stahlhut, C., Gelskov, S. V., Barbosa, L. S., Dutat, M., de Gardelle, V., et al. (2013). A neural marker of perceptual consciousness in infants. *Science* 340, 376–380. doi: 10.1126/science.1232509
- Kristensen, L. B., Wang, L., Petersson, K. M., and Hagoort, P. (2012). The interface between language and attention: prosodic focus marking recruits a general attention network in spoken language comprehension. *Cereb. Cortex* 23, 1836–1848. doi: 10.1093/cercor/bhs164
- Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nat. Rev. Neurosci.* 5, 831–843. doi: 10.1038/nrn1533
- Lai, J., and Poletiek, F. H. (2011). The impact of adjacent-dependencies and staged-input on the learnability of center-embedded hierarchical structures. *Cognition* 118, 265–273. doi: 10.1016/j.cognition.2010.11.011
- Landry, R., and Bryson, S. E. (2004). Impaired disengagement of attention in young children with autism. *J. Child Psychol. Psychiatry* 45, 1115–1122. doi: 10.1111/j.1469-7610.2004.00304.x
- Lany, J., and Gómez, R. L. (2008). Twelve-month-old infants benefit from prior experience in statistical learning. *Psychol. Sci.* 19, 1247–1252. doi: 10.1111/j.1467-9280.2008.02233.x
- Large, E. W., and Jones, M. R. (1999). The dynamics of attending: how people track time-varying events. *Psychol. Rev.* 106, 119–159. doi: 10.1037/0033-295X.106.1.119
- Law, J., Boyle, J., Harris, F., Harkness, A., and Nye, C. (2000). Prevalence and natural history of primary speech and language delay: findings from a systematic review of the literature. *Int. J. Lang. Commun. Disord.* 35, 165–188. doi: 10.1080/136828200247133
- Leonard, L. B. (1998). *Children with Specific Language Impairment*. Cambridge, MA: MIT Press.
- Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press.
- Lindsay, G., Dockrell, J. E., and Strand, S. (2007). Longitudinal patterns of behaviour problems in children with specific speech and language difficulties: child and contextual factors. *Br. J. Educ. Psychol.* 77, 811–828. doi: 10.1348/000709906X171127
- López-Barroso, D., Catani, M., Ripollés, P., Dell'Acqua, F., Rodríguez-Fornells, A., and de Diego-Balaguer, R. (2013). Word learning is mediated by the left arcuate fasciculus. *Proc. Natl. Acad. Sci. U.S.A.* 110, 13168–13173. doi: 10.1073/pnas.1301696110
- Love, A. J., and Thompson, M. G. (1988). Language disorders and attention deficit disorders in young children referred for psychiatric services: analysis of prevalence and a conceptual synthesis. *Am. J. Orthopsychiatry* 58, 52–64. doi: 10.1111/j.1939-0025.1988.tb01566.x
- Marcus, G. F., Vijayan, S., Bandi Rao, S., and Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science* 283, 77–80. doi: 10.1126/science.283.5398.77
- Mattys, S. L., Jusczyk, P. W., Luce, P. A., and Morgan, J. L. (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cogn. Psychol.* 38, 465–494. doi: 10.1006/cogp.1999.0721
- McNealy, K., Mazziotta, J. C., and Dapretto, M. (2006). Cracking the language code: neural mechanisms underlying speech parsing. *J. Neurosci.* 26, 7629–7639. doi: 10.1523/JNEUROSCI.5501-05.2006
- Mintz, T. H. (2003). Frequent frames as a cue for grammatical categories in child directed speech. *Cognition* 90, 91–117. doi: 10.1016/S0010-0277(03)00140-9
- Mintz, T., and Walker, R. (2006). Infant's sensitivity to vowel harmony and its role in word segmentation. *Paper Presented at the Annual Meeting of the LSA*, Albuquerque, NM.
- Mueller, J. L., Bahlmann, J., and Friederici, A. D. (2010). Learnability of embedded syntactic structures depends on prosodic cues. *Cogn. Sci.* 34, 338–349. doi: 10.1111/j.1551-6709.2009.01093.x
- Mueller, J. L., Friederici, A. D., and Mannel, C. (2012). Auditory perception at the root of language learning. *Proc. Natl. Acad. Sci. U.S.A.* 109, 15953–15958. doi: 10.1073/pnas.1204319109
- Myachykov, A., Tomblin, R. S., and Posner, M. I. (2005). Attention and empirical studies of grammar. *Linguist. Rev.* 22, 347–364. doi: 10.1515/tlir.2005.22.2-4.347
- Nazzi, T., Bertoncini, J., and Bijeljac-Babic, R. (2009). A perceptual equivalent of the labial-coronal effect in the first year of life. *J. Acoust. Soc. Am.* 126, 1440–1446. doi: 10.1121/1.3158931
- Nazzi, T., Bertoncini, J., and Mehler, J. (1998). Language discrimination by newborns: toward an understanding of the role of rhythm. *J. Exp. Psychol. Hum. Percept. Perform.* 24, 756–766. doi: 10.1037/0096-1523.24.3.756
- Newport, E. L. (1990). Maturation constraints on language learning. *Cogn. Sci.* 14, 11–28. doi: 10.1207/s15516709cog1401\_2
- Nobre, A., Correa, A., and Coull, J. (2007). The hazards of time. *Curr. Opin. Neurobiol.* 17, 465–470. doi: 10.1016/j.conb.2007.07.006
- Noterdaeme, M., and Amorosa, H. (1999). Evaluation of emotional and behavioral problems in language impaired children using the Child Behavior Checklist. *Eur. Child Adolesc. Psychiatry* 8, 71–77. doi: 10.1007/s007870050087
- Onnis, L., Monaghan, P., Richmond, K., and Chater, N. (2005). Phonology impacts segmentation in online speech processing. *J. Mem. Lang.* 53, 225–237. doi: 10.1016/j.jml.2005.02.011
- Pacton, S., and Perruchet, P. (2008). An attention-based associative account of adjacent and nonadjacent dependency learning. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 80–96. doi: 10.1037/0278-7393.34.1.80
- Pacton, S., Sobaco, A., and Perruchet, P. (2015). Is an attention-based associative account of adjacent and nonadjacent dependency learning valid? *Acta Psychol. (Amst)* 157, 195–199. doi: 10.1016/j.actpsy.2015.03.002
- Parise, E., Friederici, A. D., and Striano, T. (2010). “Did you call me?” 5-month-old infants own name guides their attention. *PLoS ONE* 5:e14208. doi: 10.1371/journal.pone.0014208
- Pastor, P. N., and Reuben, C. A. (2008). Diagnosed attention deficit hyperactivity disorder and learning disability: United States, 2004–2006. *Vital Health Stat.* 10, 1–14.
- Pelucchi, B., Hay, J. F., and Saffran, J. R. (2009). Statistical learning in a natural language by 8-month-old infants. *Child Dev.* 80, 674–685. doi: 10.1111/j.1467-8624.2009.01290.x
- Peña, M., Bonatti, L. L., Nespor, M., and Mehler, J. (2002). Signal-driven computations in speech processing. *Science* 298, 604–607. doi: 10.1126/science.1072901
- Perani, D., Saccuman, M. C., Scifo, P., Anwander, A., Anwander, A., Spada, D., et al. (2011). Neural language networks at birth. *Proc. Natl. Acad. Sci. U.S.A.* 108, 16056–16061. doi: 10.1073/pnas.1102991108
- Perruchet, P., Poulin-Charronnat, B., and Pacton, S. (2012). “Learning nonadjacent dependencies,” in *Encyclopedia of the Sciences of Learning*, ed. N. M. Seel (New York, NY: Springer-Verlag), 1941–1944.
- Posner, M. I., and Cohen, Y. (1984). Components of visual orienting. *Atten. Perform.* 32, 531–556.
- Posner, M. I., and Petersen, S. E. (1990). The attention system of the human brain. *Annu. Rev. Neurosci.* 13, 25–42. doi: 10.1146/annurev.ne.13.030190.000325
- Power, J. D., Fair, D. A., Schlaggar, B. L., and Petersen, S. E. (2010). The development of human functional brain networks. *Neuron* 67, 735–748. doi: 10.1016/j.neuron.2010.08.017
- Redmond, S. M., and Rice, M. L. (1998). The socioemotional behaviors of children with SLI: Social Adaptation or Social Deviance? *J. Speech Lang. Hear. Res.* 41, 688–700. doi: 10.1044/jslhr.4103.688
- Redmond, S. M., and Rice, M. L. (2002). Stability of behavioral ratings of children with SLI. *J. Speech Lang. Hear. Res.* 45, 190–201. doi: 10.1044/1092-4388(2002/014)
- Richards, J. E., Reynolds, G. D., and Courage, M. L. (2010). The neural bases of infant attention. *Curr. Dir. Psychol. Sci.* 19, 41–46. doi: 10.1177/0963721409360003
- Roberts, T. P. L., Heiken, K., Zarnow, D., Dell, J., Nagae, L., Blaskey, L., et al. (2014). Left hemisphere diffusivity of the arcuate fasciculus: influences of autism spectrum disorder and language impairment. *AJNR Am. J. Neuroradiol.* 35, 587–592. doi: 10.3174/ajnr.A3754
- Rohenkohl, G., Gould, I. C., Pessoa, J., and Nobre, A. C. (2014). Combining spatial and temporal expectations to improve visual perception. *J. Vis.* 14, 1–13. doi: 10.1167/14.4.8

- Romberg, A. R., and Saffran, J. R. (2013). All together now: concurrent learning of multiple structures in an artificial language. *Cogn. Sci.* 37, 1290–1320. doi: 10.1111/cogs.12050
- Rueda, M. R., Pozuelos, J. P., and Cómbita, L. M. (2015). Cognitive Neuroscience of Attention From brain mechanisms to individual differences in efficiency. *AIMS Neurosci.* 2, 183–202. doi: 10.3934/Neuroscience.2015.4.183
- Ruff, H. A., and Rothbart, M. K. (1996). *Attention in Early Development: Themes and Variations*. New York: Oxford University Press.
- Sabisch, B., Hahne, C. A. A., Glass, E., von Suchodoletz, W., and Friederici, A. D. (2009). Children with specific language impairment: the role of prosodic processes in explaining difficulties in processing syntactic information. *Brain Res.* 1261, 37–44. doi: 10.1016/j.brainres.2009.01.012
- Saffran, J. R., Aslin, R. N., and Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science* 274, 1926–1928. doi: 10.1126/science.274.5294.1926
- Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., and Barrueco, S. (1997). Incidental language learning: listening (and Learning ) out of the Corner of Your Ear. *Psychol. Sci.* 8, 101–106. doi: 10.1111/j.1467-9280.1997.tb00690.x
- Sanabria, D., Capizzi, M., and Correa, A. (2011). Rhythms that speed you up. *J. Exp. Psychol. Hum. Percept. Perform.* 37, 236–244. doi: 10.1037/a0019956
- Sanders, L. D., Newport, E. L., and Neville, H. J. (2002). Segmenting nonsense: an event-related potential index of perceived onsets in continuous speech. *Nat. Neurosci.* 5, 700–703. doi: 10.1038/nn873
- Sandoval, M., and Gómez, R. L. (2013). The development of nonadjacent dependency learning in natural and artificial languages. *Wiley Interdiscip. Rev. Cogn. Sci.* 4, 511–522. doi: 10.1002/wcs.1244
- Santelmann, L. M., and Jusczyk, P. W. (1998). Sensitivity to discontinuous dependencies in language learners: evidence for limitations in processing space. *Cognition* 69, 105–134. doi: 10.1016/S0010-0277(98)00060-2
- Saur, D., Kreher, B. W., Schnell, S., Kummerer, D., Kellmeyer, P., Vry, M. S., et al. (2008). Ventral and dorsal pathways for language. *Proc. Natl. Acad. Sci. U.S.A.* 105, 18035–18040. doi: 10.1073/pnas.0805234105
- Scott-Van Zeeland, A. A., Mcnealy, K., Wang, A. T., Sigman, M., Bookheimer, S. Y., and Dapretto, M. (2010). no neural evidence of statistical learning during exposure to artificial languages in children with autism spectrum disorders ashley. *Biol. Psychiatry* 68, 345–351. doi: 10.1016/j.biopsych.2010.01.011.No
- Seidl, A., and Johnson, E. K. (2008). Boundary alignment enables 11-month-olds to segment vowel initial words from speech. *J. Child Lang.* 35, 1–24. doi: 10.1017/S0305000907008215
- Smith, L. B., Yu, C., and Pereira, A. F. (2011). Not your mother's view: the dynamics of toddler visual experience. *Dev. Sci.* 14, 9–17. doi: 10.1111/j.1467-7687.2009.00947.x.No
- Soderstrom, M., Wexler, K., and Jusczyk, P. (2002). “English-learning toddlers' sensitivity to agreement morphology in receptive grammar,” in *Proceedings of the 26th Annual Boston University Conference on Language Development*, eds B. Skarabela, S. A. Fish, and A. H. J. Do (Somerville, MA: Cascadilla Press), 643–652.
- Sowell, E. R., Thompson, P. M., Welcome, S. E., Henkenius, A. L., Toga, A. W., and Peterson, B. S. (2003). Cortical abnormalities in children and adolescents with attention-deficit hyperactivity disorder. *Lancet* 362, 1699–1707. doi: 10.1016/S0140-6736(03)14842-8
- Spaulding, T. J., Plante, E., and Vance, R. (2008). Sustained selective attention skills of preschool children with specific language impairment: evidence for separate attentional capacities. *J. Speech Lang. Hear. Res.* 51, 16–34. doi: 10.1044/1092-4388(2008/002)
- Stechler, G., and Latz, E. (1966). Some observations on attention and arousal in the human infant. *J. Am. Acad. Child Psychiatry* 5, 517–525. doi: 10.1016/S0002-7138(09)62098-7
- Stevens, C., Fanning, J., Coch, D., Sanders, L., and Neville, H. (2008). Neural mechanisms of selective auditory attention are enhanced by computerized training: electrophysiological evidence from language-impaired and typically developing children. *Brain Res.* 1205, 55–69. doi: 10.1016/j.brainres.2007.10.108
- Tannock, R., and Schachar, R. (1996). “Executive dysfunction as an underlying mechanism of behavior and language problems in attention deficit hyperactivity disorder,” in *Language, Learning, and Behavior Disorders: Developmental, Biological, and Clinical Perspectives*, eds J. H. Beitchman, N. Cohen, M. Konstantareas, and R. Tannock (New York: Cambridge University Press), 128–155.
- Teinonen, T., Fellman, V., Näätänen, R., Alku, P., and Huottilainen, M. (2009). Statistical language learning in neonates revealed by event-related brain potentials. *BMC Neurosci.* 10:21. doi: 10.1186/1471-2202-10-21
- Thiessen, E. D., Hill, E. A., and Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy* 7, 53–71. doi: 10.1207/s15327078in0701\_5
- Tirosh, E., and Cohen, A. (1998). Language deficit with attention-deficit disorder: a prevalent comorbidity. *J. Child Neurol.* 13, 493–497. doi: 10.1177/088307389801301005
- Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., and O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *J. Speech Lang. Hear. Res.* 40, 1245–1260. doi: 10.1044/jslhr.4006.1245
- Tomblin, J. B., Zhang, X., Buckwalter, P., and Catts, H. (2000). The association of reading disability, behavioral disorders, and language impairment among second-grade children. *J. Child Psychol. Psychiatry* 41, 473–482. doi: 10.1111/1469-7610.00632
- Tomblin, R. S., and Myachykov, A. (2015). “Attention and salience,” in *Mouton Handbook of Cognitive Linguistics*, eds E. Dąbrowska and D. Divjak (Berlin: de Gruyter Mouton).
- Toro, J. M., Sinnett, S., and Soto-Faraco, S. (2005). Speech segmentation by statistical learning depends on attention. *Cognition* 97, B25–B34. doi: 10.1016/j.cognition.2005.01.006
- Toro, J. M., and Trolabón, J. B. (2005). Statistical computations over a speech stream in a rodent. *Percept. Psychophys.* 67, 867–875. doi: 10.3758/BF03193539
- Trainor, L. J., and Desjardins, R. N. (2002). Pitch characteristics of infant-directed speech affect infants' ability to discriminate vowels. *Psychon. Bull. Rev.* 9, 335–340. doi: 10.3758/BF03196290
- Tummelshammer, K. S., Mareschal, D., and Kirkham, N. Z. (2014). Infants' selective attention to reliable visual cues in the presence of salient distractors. *Child Dev.* 85, 1981–1994. doi: 10.1111/cdev.12239
- Turk-Browne, N. B., Jungé, J., and Scholl, B. J. (2005). The automaticity of visual statistical learning. *J. Exp. Psychol. Gen.* 134, 552–564. doi: 10.1037/0096-3445.134.4.552
- Uylings, H. B. M. (2006). Development of the human cortex and the concept of “Critical” or “Sensitive” Periods. *Lang. Learn.* 56, 59–90. doi: 10.1111/j.1467-9922.2006.00355.x
- vanMarle, K., and Wynn, K. (2006). Six-month-old infants use analog magnitudes to represent duration. *Dev. Sci.* 9, F41–F49. doi: 10.1111/j.1467-7687.2006.00508.x
- Verhoeven, J. S., Rommel, N., Prodi, E., Leemans, A., Zink, I., Vandewalle, E., et al. (2012). Is there a common neuroanatomical substrate of language deficit between autism spectrum disorder and specific language impairment? *Cereb. Cortex* 22, 2263–2271. doi: 10.1093/cercor/bhr292
- Vydrova, R., Komarek, V., Sanda, J., Sterbova, K., Jahodova, A., Maulisova, A., et al. (2015). Structural alterations of the language connectome in children with specific language impairment. *Brain Lang.* 151, 35–41. doi: 10.1016/j.bandl.2015.10.003
- Weber, C., Hahne, A., Friedrich, M., and Friederici, A. D. (2005). Reduced stress pattern discrimination in 5-month-olds as a marker of risk for later language impairment: neurophysiological evidence. *Brain Res. Cogn. Brain Res.* 25, 180–187. doi: 10.1016/j.cogbrainres.2005.05.007
- Willcutt, E. G., and Pennington, B. F. (2000). Comorbidity of reading disability and attention-deficit/hyperactivity disorder: differences by gender and subtype. *J. Learn. Disabil.* 33, 179–191. doi: 10.1177/002221940003300206
- Wolff, P. H. (1965). THE DEVELOPMENT OF ATTENTION IN YOUNG INFANTS. *Ann. N. Y. Acad. Sci.* 118, 815–830. doi: 10.1111/j.1749-6632.1965.tb04153.x

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 de Diego-Balaguer, Martinez-Alvarez and Pons. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Can a bird brain do phonology?

Bridget D. Samuels<sup>1,2\*</sup>

<sup>1</sup> Department of Linguistics and Cognitive Science, Pomona College, Claremont, CA, USA, <sup>2</sup> Center for Craniofacial Molecular Biology, University of Southern California, Los Angeles, CA, USA

## OPEN ACCESS

### Edited by:

Cedric Boeckx,  
Catalan Institute for Research and  
Advanced Studies (ICREA)/Universitat  
de Barcelona, Spain

### Reviewed by:

Pedro Tiago Martins,  
Pompeu Fabra University, Spain  
Eric Raimy,  
University of Wisconsin-Madison, USA

### \*Correspondence:

Bridget D. Samuels,  
Department of Linguistics and  
Cognitive Science, Pomona College,  
185 E. Sixth St., Claremont,  
CA 91711, USA  
bridget.samuels@gmail.com

### Specialty section:

This article was submitted to  
Language Sciences,  
a section of the journal  
Frontiers in Psychology

**Received:** 19 June 2015

**Accepted:** 13 July 2015

**Published:** 28 July 2015

### Citation:

Samuels BD (2015) Can a bird brain  
do phonology?  
Front. Psychol. 6:1082.  
doi: 10.3389/fpsyg.2015.01082

A number of recent studies have revealed correspondences between song- and language-related neural structures, pathways, and gene expression in humans and songbirds. Analyses of vocal learning, song structure, and the distribution of song elements have similarly revealed a remarkable number of shared characteristics with human speech. This article reviews recent developments in the understanding of these issues with reference to the phonological phenomena observed in human language. This investigation suggests that birds possess a host of abilities necessary for human phonological computation, as evidenced by behavioral, neuroanatomical, and molecular genetic studies. Vocal-learning birds therefore present an excellent model for studying some areas of human phonology, though differences in the primitives of song and language as well as the absence of a human-like morphosyntax make human phonology differ from birdsong phonology in crucial ways.

**Keywords:** birdsong, phonology, language-ready brain, cognitive biology, comparative neuroscience, evolution of language, biolinguistics

## 1. Introduction

The striking similarities between how some birds learn to sing and how human infants learn to talk has been a source of fascination for researchers for generations, dating back to Darwin's (1871) *Descent of Man*. Darwin already understood that the capacity for vocal learning is a rare ability in the animal kingdom but constitutes an important component of birdsong and human language learning. For this and other reasons, Darwin called birdsong the "nearest analogy to language" and looked to birds for insight into how human language may have evolved.

Modern research has confirmed that vocal learning is indeed a rare ability, particularly among mammals. Another key component of how we process speech, namely categorical perception, was once thought to be quite rare as well, giving rise to the notion that "speech is special" because it uniquely makes use of this ability. However, an explosion of work beginning with Kuhl and Miller (1975) established that categorical perception is ubiquitous in species ranging from macaques (May et al., 1989) to crickets (Wyttenbach et al., 1996). Other animals can perceive human speech categorically and can perceive their own vocalizations categorically; moreover, humans perceive non-speech stimuli such as colors categorically.

The availability of new genetic and neuroimaging techniques has complemented these behavioral studies so that we may begin to understand birdsong and human language on the level of neural connectivity and gene expression. Interestingly, these approaches underscore the similarities between perception and production in humans and birds that are vocal learners. Here, I review some recent literature on this topic, focusing on two main areas: vocal learning and vocalization structure (phonological syntax). In each of these areas, what is used to learn, perceive, and produce birdsong appears to be highly similar to what is employed in human speech. However, human phonology is crucially different from birdsong phonology because of its connection to human



morphosyntax, which is a semantically compositional or “lexical” syntax in the sense of Marler (1998). Thus, a bird brain may not be truly language-ready, but may still provide an excellent model for understanding components of human speech and the constraints that shaped the evolution of the human language faculty.

## 2. Vocal Learning

Vertebrates all seem to have the ability for auditory learning, or committing a novel sound to memory. Vocal learners have the additional ability to imitate or mimic a learned sound. Human language relies heavily on vocal learning, since all vocabulary items and a variety of other linguistic structures must be learned in order to achieve linguistic competence. Yet, it is a well-known curiosity that our species is alone among primates in having a well-developed capacity for vocal learning, though Seyfarth and Cheney (1986) suggest that vervet monkey calls may be learned. Among the myriad species that have been studied, among mammals only humans, cetaceans, pinnipeds, elephants, and some bats are relatively strong vocal learners; oscine songbirds (passerines), parrots, and hummingbirds are among the best vocal learners in the animal kingdom (see references in Schachner et al., 2009 and Petkov and Jarvis, 2012).

Comparisons between strongly vocal-learning birds and those with a poor capacity for vocal learning can be used to shed light on how the neural plasticity and other capacities needed to support the vocal learning mechanism may have evolved. Moreover, comparing learned and innate birdsongs can provide the opportunity to probe whether or to what extent vocal learning allows more structurally complex song. Note that the capacity for complex vocal learning emerged independently in three clades of birds, which are separated by 68 million years from a common ancestor (see references in Pfenning et al., 2014). Alternatively, this capacity may only have arisen twice in birds: once in hummingbirds and once in the common ancestor of parrots and songbirds, which are closely related, with a loss of the ability in the suboscine songbirds (Suh et al., 2011; Petkov and Jarvis, 2012) and perhaps a gain in at least one suboscine species (Saranathan et al., 2007). Currently, most research on vocal learning in birds has focused on the passerines, but an intriguing recent study on suggests that one portion of the song system is similar in songbirds, hummingbirds, and parrots, while another portion evolved uniquely in parrots over 29 million years ago (Chakraborty et al., 2015). The similarity between the vocal learning systems in these avian clades is remarkable for the same reason that the similarities between the avian and human ones are: evolution has come up with nearly the same means of developing this ability time and time again. For researchers studying human language, this is fortunate since it means that birds can model the object of our study to a surprising extent.

Doupe and Kuhl (1999) provide an overview of the evidence for vocal learning in a particular species, which involves the following properties: (i) initially immature vocalizations (“babbling”) that eventually become adultlike; (ii) a relatively fixed individual-level repertoire that varies across individuals/groups; (iii) individual-level differences that depend

on experience/exposure; and (iv) the necessity of auditory feedback to maintain normal vocalizations. The behavioral evidence for vocal learning in songbirds and parallels to human first language acquisition have been reviewed widely in the literature (see e.g., Doupe and Kuhl, 1999; Bolhuis et al., 2010; Berwick et al., 2011), and I will not recap those arguments here. Schachner et al. (2009) discuss a relatively new line of research investigating the connection between vocal learning and spontaneous rhythmic motor entrainment, or the ability to align movement with auditory input (i.e., move to a beat or dance). They found support for the hypothesis that entrainment is a by-product of selection vocal mimicry that arises from a specialized connection between the auditory and motor systems (Patel, 2008): upon analyzing videos of a wide variety of animals purportedly dancing, they found that only vocal mimicking species showed any evidence of entrainment. These included the Asian elephant and 14 species of parrot. It has also been widely noted that both humans and songbirds exhibit critical or sensitive periods for native-like song/language acquisition (see e.g., Lenneberg, 1967). However, not all vocal learning species have this property; starlings, canaries, and pied flycatchers are “open-ended” learners (Brainard and Doupe, 2002; Eriksen and Lampe, 2011), and Prat et al. (2015) argue against a short critical period in Egyptian fruit bats, which are vocal learners and initially exhibit immature vocalizations akin to babbling. I therefore set this issue aside.

### 2.1. Neural and Molecular Evidence

A number of recent studies investigating the neural and molecular underpinnings of vocal learning focus on songbirds. Vocal learning is served by regions in the motor cortex and striatum in both songbirds and humans, and these regions appear to have a uniquely direct connection in both humans and vocal-learning birds, as opposed to non-vocal-learning birds and primates (Pfenning et al., 2014). The anterior forebrain pathway involved in song learning and plasticity in the adult song of vocal-learning birds links the HVC (a region formerly known as the hyperstriatum ventrale, pars caudalis) to Area X of the basal ganglia, the thalamic nucleus dorsolateralis anterior pars medialis (DLM), the lateral magnocellular nucleus of the anterior nidopallium (LMAN), and the robust nucleus of the arcopallium (RA), where it connects with the posterior motor pathway, which is also involved in song production and learning (Bolhuis et al., 2010). Pfenning et al. (2014) took a computational approach, screening gene expression databases from humans and all three clades of vocal-learning birds as well as the non-vocal-learning dove, quail, and macaque. The results of these gene expression studies confirmed that not only have human and vocal-learning bird brains evolved convergently from an anatomical perspective in ways that are not true of non-vocal-learning species, this convergence has also occurred on a molecular level. For birds and humans to arrive at the ability of vocal learning involved the convergent evolution of expression patterns of hundreds of genes in the regions of the brain that subserve this behavior. Many of these genes affect neural connectivity or function in fine motor control. Area X and VS in the songbird (finch) striatum show specialized gene expression similar to that of the

putamen and body of the caudate in the human basal ganglia. The songbird RA is the most similar in specialized gene expression to somatosensory cortex in humans, specifically the primary motor cortex and adjacent somatosensory portion of the central sulcus, as well as the ventral portion of the laryngeal motor cortex. In these areas, the number of genes with significantly shared specialized expression between finches and humans ranges from the tens to the hundreds. The expression levels of *Foxp2* in Area X have been studied extensively; see Bolhuis et al. (2010) for a recent overview of the literature on this gene in humans and other species. Levels of *FoxP2* are higher in Area X in juvenile zebra finches during the sensitive period for song learning (Haesler et al., 2004). In canaries that add new song elements to their repertoire at the end of breeding season, the level of *Foxp2* expression is higher during this period (Haesler et al., 2004). Singing downregulates *Foxp2* mRNA in Area X in both juvenile zebra finches and adult males during “undirected” singing in the absence of a female (Teramitsu and White, 2006; Teramitsu et al., 2010).

It has been suggested that the avian pallium—which contains several areas discussed above, including the HVC, RA, and LMAN—is homologous with the mammalian neocortex. Homology between these structures would be significant because computation in the laminated cortex is considered to be responsible for complex behavior. Although only mammalian brains have a cortex, birds are also capable of sophisticated behaviors including tool use, basic arithmetic, causal reasoning, and recognizing themselves in mirrors (see references in Calabrese and Woolley, 2015). Like the mammalian primary auditory cortex, the avian auditory pallium (Field L) consists of three regions that receive auditory input from the thalamus (Bolhuis et al., 2010). The auditory pallium and neocortex display highly similar patterns of connectivity (Wang et al., 2010), and gene expression analyses also highlight similarities between these two tissues (Dugas-Ford et al., 2012). Calabrese and Woolley (2015) recorded neuronal populations in different portions of Field L in zebra finches and showed that the auditory pallium exhibits the same hierarchical information-processing principles as the canonical cortical microcircuit in mammals. Their conclusion is that this microcircuit evolved in a common ancestor of birds and mammals, 300+ million years ago. As Harris (2015) notes, it may be even older; the fish brain also has a pallium, and invertebrates such as cephalopods also display striking intelligence. Harris therefore suggests that the canonical cortical microcircuit may be evolutionarily quite old, but only repurposed for intelligence in species where the benefits of doing so outweighed the costs of increased brain size, energy expenditure, and development time.

The overall picture that emerges from these studies is that the neural and molecular bases of vocal learning in humans and songbirds have strong similarities, owing in part to convergent evolution (analogy) and in part to homology. It should be noted that both analogy and homology are of potential interest to the study of language evolution. Homologies highlight our ancient heritage, the biological substrate that was adapted and/or exapted for the externalization of language. Analogies show that similar solutions may arise to similar problems (Gould,

1976). For example, the last common ancestor of the octopus and vertebrates was ca. 750 million years ago; the octopus eye emerged ca. 480 million years ago and the vertebrate eye emerged completely independently 640–490 million years ago, yet human and octopus eyes have 70% of their expressed genes in common (Ogura et al., 2004; Fernald, 2006). Of the 1052 genes expressed in the octopus eye, 1019 (97%) are evolutionarily quite old, dating back to the common ancestor of bilateria (Ogura et al., 2004). Convergent identical amino acid substitutions have been discovered in a number of areas, including the gene encoding the motor protein Prestin, which is crucial for echolocation, in bats and cetaceans (Liu et al., 2010; see Pfenning et al., 2014, for further examples). This is in part because the vertebrate brain provides a highly genetically constrained substrate upon which to build (Jarvis, 2004). Noting analogies like these helps to shed light on the physical and developmental constraints on solving the problem in question, which “may essentially force natural selection to come up with the same solution repeatedly when confronted with similar problems” (Hauser et al., 2002, p. 1572). In the context of describing the growth of language in a human child, Chomsky (2005, 2007) has dubbed properties that arise from such constraints “third factor” principles, which interact in a dynamic fashion with the genetic endowment (first factor) and experience (second factor). Studies like the ones described here highlight the fact none of these factors can be viewed in isolation, and that in particular the third factor shapes the first in a powerful fashion that we are only beginning to uncover.

### 3. Phonological Syntax

One of the properties that distinguishes vocalizations like human language and the songs of birds and whales from the calls of non-human primates is the rich structure of the former. On the other hand, primates are capable of producing distinct calls with distinguishable referents (Arnold and Zuberbühler, 2006a,b, 2008; Ouattara et al., 2009; Căsar et al., 2013), whereas the same song serves a number of expressive functions in birds. The idea that human language integrates a song-like expressive system with a lexical system like that of other primates has been recently explored by Miyagawa et al. (2013, 2014). In the sections that follow, I will review evidence suggesting that the structure of birdsong is like that of human phonology in important ways, that the elements within songs are context-sensitive like the elements of human speech, and that birds may be capable of computations as complex as those demanded by human phonology.

#### 3.1. Hierarchical Structure

The structure of birdsongs can be modeled as exhibiting hierarchy with limited depth. Each individual has a repertoire of notes, akin to phonemes in human speech, often shared with other individuals of the species. A sparrow or Bengalese finch has a repertoire of less than 8 note types, such as whistles, trills, and buzzes in the case of the sparrow, each exhibiting within-category variation (Marler, 2000). Multiple notes are produced sequentially to produce a syllable. A syllable is defined as a group of notes bordered by silence, unlike syllables in human speech, which readily follow each other without any interruption.

A typical zebra finch syllable might range from 60–180 ms in duration (Fehér et al., 2009). When interrupted by a strobe flash in the midst of a syllable, a zebra finch will complete the syllable, which suggests that these chunks are units of motor planning (Cynx, 1990). A sequence of several syllables that repeats during the course of a song is called a motif (Slater, 2000). Doupe and Kuhl (1999) liken motifs to phrases in human language, though Yip (2006) is tempted to equate them with prosodic words. An entire song bout consists of several motifs. The number of songs created by an individual bird varies greatly according to species. A winter wren may know 5–10 distinct songs, each lasting 10 s, whereas each starling may know up to 100 motifs and combine some of them in a song bout that is 30 s to a minute long (Yip, 2006). Nightingales and mockingbirds may have larger repertoires of hundreds of songs (Marler, 2000; Berwick et al., 2011), organized into less than a dozen “packages” of bouts that are typically produced together (Todt and Hultsch, 1996). It is important to note that notes and syllables do not have any meaning. This is what Marler (1998, 2000) calls “phonological syntax” or “phonocoding”; the elements of songs can be combined in different sequences, but this does not change their meaning. Similarly, human vocalizations consist of combinations of sounds (phones) into morphemes, but the phones themselves are not meaningful. Of course, this differs from human language on a word-level or sentence-level scale, which is said to have “lexical syntax” or “lexicoding”; the meaning of a word results from the meanings of its morphemes, and the meaning of a sentence arises from the meanings of its words. It is also important to consider that human speech does not bottom out at the segmental (phone) level. In all modern phonological theories, phonological processes operate over smaller units: distinctive features, elements, or articulatory gestures. There is no evidence for manipulation of any sub-note-level features in birdsong.

Analogies between birdsong syllables and human syllables, and between birdsong motifs and human prosodic words or phrases, are of limited utility. Conservatively, one can say that language and song are alike in having structure on different timescales: notes/phonemes in the tens of milliseconds, syllables around 100–200 ms, and longer timescales for larger units (Doupe and Kuhl, 1999; Yip, 2006). These elements are arranged in non-random order, as will be discussed in a later section. It has been suggested that chunking songs into motifs and syllables may serve purposes for both memorization and production, similar to breaking a ten-digit telephone number into chunks of three or four digits (Williams and Staples, 1992). I have noted in previous work (Samuels, 2011) that the maximal number of segments in a human syllable is around 5 (depending on theory-internal considerations), which is at the upper limit of the number of elements we can simultaneously hold in short-term memory (Miller, 1956; Cowan, 1998, 2001). It is also interesting to note that humpback whale songs follow the same general pattern discussed here: they typically consist of up to ten ordered elements, which are then repeated a few times as a unit (Payne, 2000). Reduplication, which is a common way of expressing pluralization, durativity, and other grammatical functions in human language and also plays a role in

many language games, resembles this order-preserving repetition (Samuels, 2011; Miyagawa et al., 2014). However, reduplication only creates a single extra copy of the elements over which it operates.

There is some experimental evidence concerning what areas of the brain control birdsong structure. Kao and Brainard (2006) found that inducing lesions in the LMAN of zebra finches reduces variability in syllable structure, which is normally greater in male birds’ undirected singing than it is in their singing to females. However, damage to the LMAN does not affect the number of motif repetitions or the sequencing of syllables. In adult finches, auditory units in the LMAN and in the HVC respond more strongly to a bird’s own song than to the songs of other conspecifics (Lewicki and Arthur, 1996; Doupe, 1997). Some neurons in the zebra finch HVC appear to integrate auditory information over a window of several 100 ms, so they are sensitive to certain sequences or combinations of syllables (Lewicki and Arthur, 1996). It has been suggested that such sequences are represented in the HVC via population coding (Nishikawa et al., 2008). Like humans, zebra finches show left-hemisphere dominance of the HVC and in the caudomedial nidopallium, which have been compared to the human Broca’s and Wernicke’s areas, respectively (Moorman et al., 2012; Pfenning et al., 2014). There is also evidence to suggest that more complex song syntax is associated with changes in gene expression and neural organization (Boeckx and Benítez-Burraco, 2014). The Bengalese finch, which is a domesticated type of white-backed munia, has a more complex song structure than its wild counterpart (Okanoya, 2004). This difference appears to be reflected in differential androgen receptor expression in the GABAergic neurons in Area X and in differential epigenetic regulation (methylation) of regions upstream of the start codon for this receptor (Wada et al., 2013). A recent vein of research into the mechanisms of human speech perception is exploring coupled theta-gamma oscillations in the auditory cortex as a means through which the different time scales of the speech stream may be integrated, perhaps via a more general mechanism of attention (Martins and Boeckx, 2014). The coupling of theta waves, which track syllabic rhythm, with gamma waves that track a shorter interval corresponding to the segment or phoneme, could enable “de-multiplexing” of the speech stream to facilitate parsing and encoding (Hyafil et al., 2015). There is evidence suggesting that coupling may be disrupted in some individuals with autism (Jochaut et al., 2015).

### 3.2. Contextual Alternations

Human speech is comprised of sounds or phones that can be categorized in terms of their membership in abstract categories known as phonemes. A phoneme may have multiple realizations, known as allophones, that are distributed in a context-sensitive manner. For example, the voiceless stop consonants /p, t, k/ in English are aspirated when they appear word-initially, unaspirated after /s/, and unreleased or glottalized word-finally. Membership in a particular phonemic category varies from language to language: the alveolar flap [ɾ] is an allophone of /t/ and /d/ that appears intervocally or trochaic foot-medially in English, as in the words *putty* and *ladder*, whereas [ɾ] is considered by some phonologists to be an allophone of /r/

in Spanish (Harris, 1969). The realization of a phoneme can also be affected by its neighbors in a phenomenon known as coarticulation, as it is attributed to anticipatory or lagging movement of the vocal apparatus. The context-dependent, rule- or constraint-governed realization of phonemes/allophones is a defining characteristic of human phonological systems.

Wohlgemuth et al. (2010) showed that the realization of a Bengalese finch syllable is significantly affected by the preceding and following syllables. A syllable is called “convergent” if it can be preceded by at least two different syllables, and is called “divergent” if it can be followed by at least two different syllables. The identity of the following syllable affected realization of its divergent predecessor 92% of the time, and the identity of the preceding syllable affected the realization of the following convergent syllable 92% of the time. These effects extended even beyond the immediately preceding/following syllable and could be detected at least two syllables away. Measurements of RA activity suggested that this region plays a role in this context-sensitive phonology, as it responds differentially to the same syllable when produced in different contexts, though RA activity is still more strongly correlated across realizations of the same syllable than across different syllables. The magnitude of differences in response to the same syllable in different contexts correlated with the magnitude of the phonological variation across those contexts.

Allophonic-style variation has also been found at the level of notes in swamp sparrows. Lachlan and Nowicki (2015) performed careful habituation/dishabituation studies showing that sparrows categorize notes differently according to their length and their position within a syllable. Among types of notes that descend rapidly in frequency, there is a clear trimodal distribution in length in the songs of male sparrows from Pennsylvania. Short notes (clustered around 8 ms in duration) typically occur syllable-initially, while long notes (clustered around 32 ms in duration) typically occur syllable-finally. Notes of intermediate length (clustered around 16 ms in duration) can occur both syllable-initially and -finally. Interestingly, these categories are learned, and male swamp sparrows from New York have a bimodal distribution of note types that is missing the cluster of intermediate-length notes. The Pennsylvania birds in Lachlan and Nowicki’s study categorized the intermediate-length notes with the short notes in syllable-initial position, but with the long notes in syllable-final position. While it is possible that the birds construct completely different categories for syllable-initial and syllable-final word types, there remains the intriguing possibility that intermediate notes serve as an “allophone” of a phoneme-like short-note category in one position but are allophones of the long-note category in another position.

### 3.3. Computational Complexity

The formal complexity of grammars can be categorized according to the type of rules sufficient to generate them (Chomsky, 1956). The following broad categories, known as the Chomsky Hierarchy, can be defined as follows (Wall, 1972):

- (1) a. Finite-state (regular):  $A \rightarrow xB$  or  $A \rightarrow x$
- b. Context-free:  $A \rightarrow \omega$ , where  $\omega \neq$  the null string

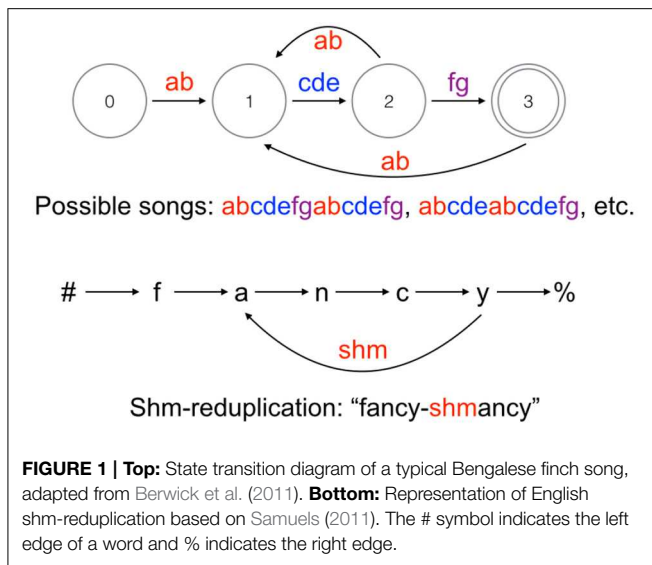
- c. Context-sensitive:  $\phi A \psi \rightarrow \phi \omega \psi$ , where  $\phi$  and  $\psi$ , but not  $\omega$ , may be the null string
- d. Unrestricted rewriting system: no restriction (Turing machine)

where  $A, B$  are nonterminals;  $x$  is a terminal;  $\phi, \omega, \psi$  are sequences of nonterminals and terminals

All known phonological alternations and phonotactics, which govern the sequential distribution of phonemes, fall into the class of regular languages and can thus be modeled with finite-state machines (Johnson, 1970; Kaplan and Kay, 1994; Karttunen, 1998). This contrasts with the domain of sentence-level syntax, which has been known since Chomsky (1956) to exhibit context-free patterns. It is now recognized that cross-serial dependencies in syntax fall outside the class of context-free languages, requiring mildly context-sensitive computations (Shieber, 1985). On the basis of this difference, Heinz and Idsardi (2011, 2013) have argued that there are likely to be multiple, distinct language learning modules that deal separately with these disparate patterns. Even within phonology, there may be more than one. Phonological patterns sometimes involve restrictions on adjacent sounds, but can also involve long-distance computations. For example, some languages including Navajo prohibit the alveolar sibilant [s] and the post-alveolar sibilant [ʃ] from co-occurring within a word, regardless of the distance between them (McDonough, 2003). Heinz and Idsardi (2013) (see also references therein) pursue the hypothesis that phonotactic constraints fall into a few distinct sub-regular classes, specifically the strictly local class when only a contiguous string of adjacent segments is involved and the strictly piecewise class for long-distance patterns like the Navajo case. Stress patterns may be of either of these types, though a few may require counting, which is measurably more complex but still falls within the class of regular languages. An intriguing question, then, is whether the phonological alternations seen in birdsongs are of these types, and/or whether birds are capable of these kinds of computations.

In nature, no known types of birdsong require more computational power than human phonological patterns: both fall within the class of regular languages. This has been shown for Bengalese finch song, which is among the more complex and variable song systems (Berwick et al., 2011). A state transition diagram of a typical Bengalese finch song (abstracting away from the probabilities of state transitions) is shown in **Figure 1** alongside a reduplication pattern found in English (see Raimy, 2000 and Samuels, 2010b, 2011 for more details on the loop formalism used to represent reduplication). Bengalese finch songs are of the simplest type recognizable by a finite-state automaton, strictly locally 2-testable languages, meaning it is possible to determine whether a sequence is licit by looking at a moving window of two-note sequences. A further interesting property of Bengalese finch songs is that they are easily learnable in a technical sense (Kakishita et al., 2009), which is not true of regular languages more broadly. As noted above, some phonotactic constraints in human languages fall into the strictly local class, though the window of observed segments must be larger than two (perhaps maximally around





five segments). Other types of birdsongs, such as those of starlings and American thrushes, are even less complex, requiring only low-order Markov models to describe the sequence of motifs (Dobson and Lemon, 1979; Gentner and Hulse, 1998). I do not know of any patterns that require strictly piecewise computation in birdsong. Attempts to determine whether starlings and finches can learn or spontaneously extract context-free patterns have generated controversy and are widely considered inconclusive at this time (Gentner et al., 2006, 2010; van Heijningen et al., 2009; ten Cate et al., 2010; Abe and Watanabe, 2011; Beckers et al., 2012; Everaert and Huybregts, 2013).

## 4. Conclusions

Although significant gaps in our knowledge remain, recent genetic, neuroanatomical, and behavioral studies have served to underscore the parallels between human language phonology and birdsong. These similarities are due in large part to convergent evolution, but some have their roots in homologies of neural structures, such as between the mammalian auditory cortex and the avian pallium. There is strong evidence that a bird brain can do some types of phonological computations, as evidenced by the patterns and relationships among elements in birdsong, which closely resemble the relationships between elements in human phonology by every measure on which they have been compared. Still, important differences remain.

One of the main differences between human and avian phonology has already been briefly mentioned in the discussion of hierarchical structure above: the primitives of birdsong are unlike those of human language. Notes seem act in a more atomic fashion than phones, which can be—and indeed must be, to provide an adequate and insightful account of human phonological systems (Jakobson et al., 1952; Halle, 2002)—decomposed into smaller phonological features (or equivalently for the present purposes, elements or gestures). It may be the case that human languages can exist without this featural

level, as has been argued for Al-Sayyid Bedouin Sign Language (Aronoff et al., 2008; Sandler et al., 2011), which lacks featural minimal pairs that are ubiquitous in all other known spoken and signed languages (cf. *bin* vs. *pin* in English, which differ in the presence or absence of a voicing feature on the first segment).

This discussion of a signed language raises another disparity between human and avian communication: unlike birdsong, human languages can be externalized in more than one modality. It is commonly held that signed and spoken language phonology are in fact identical, differing only in the (learned) content of their features (Brentari, 1998; Hale and Reiss, 2000; Mielke, 2008). Taken together, these data suggest that avian and human phonology are more comparable on a computational level than a representational one. I have argued that the underpinnings of phonological features are not unique to humans, however (Samuels, 2010a). The origins of phonological features may be attributed in part to perceptual biases known as auditory discontinuities that we inherited from the basic mammalian auditory system (see e.g., Brown and Sinnott, 2006; Kluender et al., 2006; Mesgarani et al., 2008). Some of these perceptual biases are shared with birds such as budgerigars also (Brown and Sinnott, 2006). Some birds and mammals, including non-human primates, have additionally been shown to attend spontaneously to formants (energy peaks in the acoustic signal), which are crucial correspondents of sub-segmental features in human speech (Fitch, 1994). The presence of a kinesthetic mode of language in humans also suggests that studying movement systems could also be informative. Alongside the attempts to teach primates to sign (e.g., Nim Chimpsky, Washoe the chimpanzee, Koko the gorilla, etc.), which were relatively successful relative to the prior failed attempts to teach primates to speak, some researchers have looked to “action grammars” as precursors of linguistic syntax (Greenfield et al., 1972; Greenfield, 1991, 1998; Johnson-Pynn et al., 1999; Fujita, 2007, 2009). Interestingly for the present purposes, Greenfield (1991) has suggested a parallel between action grammars and the combination of phonemes into words. Such studies suggest that moving beyond birdsong and investigating other behaviors, such as mating dances, could potentially be illuminating in this regard as well.

Birdsong also appears to be absent of non-local dependencies, which are attested in patterns such as vowel and consonant harmony in human language. Interestingly, harmony patterns provide some of the best evidence for underspecification, or the initial absence of a particular phonological feature on a certain class of segments in lexically stored morpheme forms. I have suggested elsewhere that underspecification may be a unique feature of human language, which follows if the basic elements of other vocalization systems are not composed of features like ours are (Samuels, 2015).

Another major difference is that birdsong is not fed by a recursive morphosyntactic cycle. A large number of phonological phenomena in humans are bounded by morphological or syntactic domains. For example, they may occur within words but not across them. Others are re-computed each time a new morpheme is added to the derivation, such as stress:

witness the differences between *govern* with stress on the first syllable, *governmental* with stress on the penultimate syllable, and *governmentalese* with stress on the final syllable. All this is to say that birdsong and human phonology differ substantially in the nature and structure of the input they receive. It is therefore worthwhile to consider the question of how potentially pre-existing phonological capabilities could have come to fit together with a more complex “upstream” system like that of human morphosyntax. Taken together, the evidence presented here suggests that further investigations of birds can help us to pinpoint interesting questions to ask about the cognitive abilities, neural circuitry, genetics, and epigenetics that are involved in human language, and about the nature of language evolution itself.

Of course, such studies are only one piece of the puzzle. For example, birds are not currently as amenable to genetic engineering as common laboratory species such as mice and zebrafish, which limits the availability of certain experimental approaches—but a better understanding of birds can provide the rationale for studies that may be possible in other species. Studies of *Foxp2* provide an excellent example of this kind of cross-species synergy. Initially, a heterozygous point mutation in *FOXP2* was famously identified as being associated with a language disorder, developmental verbal dyspraxia, in a British family (Lai et al., 2001). It was then established that this gene is highly conserved from reptiles to humans, but especially among mammals, with strong evidence for recent selection in the human lineage (Enard et al., 2002; Scharff and Haesler, 2005). Due to current technological limitations, RNAi-mediated knockdown using a lentiviral vector has been used to study the effect of reduced *Foxp2* expression in Area X of the zebra finch brain, rather than a transgenic approach (Haesler et al., 2007; Schulz et al., 2010). In mice, heterozygous and

homozygous *Foxp2* knockouts as well as humanized knockins have been studied, and a mouse model has been developed with a conditional null (floxed) allele, allowing crosses to transgenic lines expressing Cre drivers for tissue- and time-specific conditional knockouts (French et al., 2007). Knockdown (in finches) or haploinsufficiency (in mice) of *Foxp2* leads to altered or inaccurate vocalizations (Shu et al., 2005; Haesler et al., 2007), and in the finch this is associated with the altered density of spiny neurons in Area X (Schulz et al., 2010). Interestingly, the human version of *Foxp2* has strong effects on the plasticity of the striatum and accelerates learning when introduced into mice (Schreiweis et al., 2014). Mice with certain point mutations in one copy of *Foxp2*, including those that cause developmental verbal dyspraxia in humans, are developmentally delayed, somatically weak, and have impaired auditory-motor association learning owing to strongly altered activity in the striatal circuits, but they make the expected range of acoustically normal vocalizations (Gaub et al., 2010; French et al., 2012; Kurt et al., 2012). These studies collectively give a more robust view of this gene's role in vocalization than would be possible using a single species. In sum, looking at the communication systems of other animals as well as their cognitive abilities more generally is also necessary to achieve a better perspective on what abilities underlie human language, what species share them, and how they may have evolved.

## Acknowledgments

I am grateful to the guest editors of this Special Topic, Antonio Benítez-Burraco and Cedric Boeckx, for the opportunity to contribute to this volume. I am also indebted to the two anonymous reviewers and to Hiroki Narita for their very insightful comments and suggestions, which improved the manuscript significantly.

## References

- Abe, K., and Watanabe, D. (2011). Songbirds possess the spontaneous ability to discriminate syntactic rules. *Nat. Neurosci.* 15, 113–121. doi: 10.1038/nn.2869
- Arnold, K., and Zuberbühler, K. (2006a). The alarm-calling system of adult male putty-nosed monkeys, *cercopithecus nictitans martini*. *Anim. Behav.* 72, 643–653. doi: 10.1016/j.anbehav.2005.11.017
- Arnold, K., and Zuberbühler, K. (2006b). Language evolution: semantic combinations in primate calls. *Nature* 441:303. doi: 10.1038/441303a
- Arnold, K., and Zuberbühler, K. (2008). Meaningful call combinations in a non-human primate. *Curr. Biol.* 18, R202–R203. doi: 10.1016/j.cub.2008.01.040
- Aronoff, M., Meir, I., Padden, C. A., and Sandler, W. (2008). The roots of linguistic organization in a new language. *Interact. Stud.* 9, 133–153. doi: 10.1075/is.9.1.10aro
- Beckers, G. J., Bolhuis, J. J., Okanoya, K., and Berwick, R. C. (2012). Birdsong neurolinguistics: songbird context-free grammar claim is premature. *Neuroreport* 23, 139–145. doi: 10.1097/WNR.0b013e32834f1765
- Berwick, R. C., Okanoya, K., Beckers, G. J., and Bolhuis, J. J. (2011). Songs to syntax: the linguistics of birdsong. *Trends Cogn. Sci.* 15, 113–121. doi: 10.1016/j.tics.2011.01.002
- Boeckx, C., and Benítez-Burraco, A. (2014). The shape of the human language-ready brain. *Front. Psychol.* 5:282. doi: 10.3389/fpsyg.2014.00282
- Bolhuis, J. J., Okanoya, K., and Scharff, C. (2010). Twitter evolution: converging mechanisms in birdsong & human speech. *Nat. Rev.* 11, 747–759. doi: 10.1038/nrn2931
- Brainard, M. S., and Doupe, A. J. (2002). What songbirds teach us about learning. *Nature* 417, 351–358. doi: 10.1038/417351a
- Brentari, D. (1998). *A Prosodic Model of Sign Language Phonology*. Cambridge, MA: MIT Press.
- Brown, C. H., and Sinnott, J. M. (2006). “Cross-species comparisons of vocal perception,” in *Listening to Speech: An Auditory Perspective*, eds S. Greenberg and W. A. Ainsworth (Mahwah, NJ: Lawrence Erlbaum Associates), 183–201.
- Calabrese, A., and Woolley, S. M. (2015). Coding principles of the canonical cortical microcircuit in the avian brain. *Proc. Natl. Acad. Sci. U.S.A.* 112, 3517–3522. doi: 10.1073/pnas.1408545112
- Cäsar, C., Zuberbühler, K., Young, R. J., and Byrne, R. W. (2013). Titi monkey call sequences vary with predator location and type. *Biol. Lett.* 9:20130535. doi: 10.1098/rsbl.2013.0535
- Chakraborty, M., Walloe, S., Nedergaard, S., Fridel, E. E., Dabelsteen, T., Pakkenberg, B., et al. (2015). Core and shell song systems unique to the parrot brain. *PLoS ONE* 10:e0118496. doi: 10.1371/journal.pone.0118496
- Chomsky, N. (1956). Three models for the description of language. *IRE Trans. Inform. Theory* 2, 113–124. doi: 10.1109/TIT.1956.1056813
- Chomsky, N. (2005). Three factors in language design. *Linguist. Inq.* 35, 1–22. doi: 10.1162/0024389052993655
- Chomsky, N. (2007). “Approaching UG from below,” in *Interfaces + Recursion = Language?*, eds U. Sauerland and H. M. Gärtner (Berlin: Mouton de Gruyter), 1–29.
- Cowan, N. (1998). Visual and auditory working memory capacity. *Trends Cogn. Sci.* 2, 77–78. doi: 10.1016/S1364-6613(98)01144-9

- Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav. Brain Sci.* 24, 87–114. doi: 10.1017/S0140525X01003922
- Cynx, J. (1990). Experimental determination of a unit of song production in the zebra finch (*taeniopygia guttata*). *J. Comp. Psychol.* 104, 3–10. doi: 10.1037/0735-7036.104.1.3
- Darwin, C. (1871). *The Descent of Man and Selection in Relation to Sex*. London: John Murray.
- Dobson, C., and Lemon, R. (1979). Markov sequences in songs of American thrushes. *Behaviour* 68, 86–105.
- Doupe, A. J. (1997). Song- and order-selective neurons in the songbird anterior forebrain and their emergence during vocal development. *J. Neurosci.* 17, 1147–1167.
- Doupe, A. J., and Kuhl, P. K. (1999). Birdsong & human speech: common themes & mechanisms. *Annu. Rev. Neurosci.* 22, 567–631. doi: 10.1146/annurev.neuro.22.1.567
- Dugas-Ford, J., Rowell, J., and Ragsdale, C. (2012). Cell-type homologies and the origins of the neocortex. *Proc. Natl. Acad. Sci. U.S.A.* 109, 16974–16979. doi: 10.1073/pnas.1204773109
- Enard, W., Przeworski, M., Fisher, S. E., Lai, C. S., Wiebe, V., Kitano, T., et al. (2002). Molecular evolution of FOXP2, a gene involved in speech & language. *Nature* 418, 869–872. doi: 10.1038/nature01025
- Eriksen, A., and Lampe, H. M. (2011). Vocal plasticity – are pied flycatchers, *ficedula hypoleuca*, open-ended learners? *Ethology* 117, 188–198. doi: 10.1111/j.1439-0310.2010.01864.x
- Everaert, M., and Huybregts, R. (2013). “The design principles of natural language” in *Birdsong, Speech, and Language. Exploring the Evolution of Mind and Brain*, eds J. J. Bolhuis and M. Everaert (Cambridge, MA: MIT Press), 1–26.
- Fehér, O., Wang, H., Saar, S., Mitra, P. P., and Tchernichovski, O. (2009). *De novo* establishment of wild-type song culture in the zebra finch. *Nature* 459, 564–568. doi: 10.1038/nature07994
- Fernald, R. (2006). Casting a genetic light on the evolution of eyes. *Science* 313, 1914–1918. doi: 10.1126/science.1127889
- Fitch, W. T. (1994). *Vocal Tract Length Perception & the Evolution of Language*. Ph.D. thesis, Brown University.
- French, C., Groszer, M., Preece, C., Coupe, A., Rajewsky, K., and Fisher, S. E. (2007). Generation of mice with a conditional *foxp2* null allele. *Genesis* 45, 440–446. doi: 10.1002/dvg.20305
- French, C. A., Jin, X., Campbell, T., Gerfen, E., Groszer, M., Fisher, S. E., et al. (2012). An aetiological *foxp2* mutation causes aberrant striatal activity and alters plasticity during skill learning. *Mol. Psychiatry* 17, 1077–1085. doi: 10.1038/mp.2011.105
- Fujita, K. (2007). Facing the logical problem of language evolution. review of Jenkins (2004), *Variation & Universals in Biolinguistics*. *Engl. Linguist.* 24, 78–108. doi: 10.9793/elsj1984.24.78
- Fujita, K. (2009). A prospect for evolutionary adequacy: merge & the evolution & development of human language. *Biolinguistics* 3, 138–153.
- Gaub, S., Groszer, M., Fisher, S. E., and Ehret, G. (2010). The structure of innate vocalizations in *Foxp2*-deficient mouse pups. *Genes Brain Behav.* 9, 390–401. doi: 10.1111/j.1601-183X.2010.00570.x
- Gentner, T. Q., Fenn, K. M., Margoliash, D., and Nusbaum, H. C. (2006). Recursive syntactic pattern learning by songbirds. *Nature* 440, 1204–1207. doi: 10.1038/nature04675
- Gentner, T. Q., Fenn, K. M., Margoliash, D., and Nusbaum, H. C. (2010). Simple stimuli, simple strategies. *Proc. Natl. Acad. Sci. U.S.A.* 107, E65. doi: 10.1073/pnas.1000501107
- Gentner, T. Q., and Hulse, S. H. (1998). Perceptual mechanisms for individual vocal recognition in European starlings *sturnus vulgaris*. *Anim. Behav.* 56, 579–594. doi: 10.1006/anbe.1998.0810
- Gould, S. J. (1976). “In defense of the analog: a commentary to N. Hotton,” in *Evolution, Brain, & Behavior: Persistent Problems*, eds R. Masterson, W. Hodos, and H. Jerison (New York, NY: Wiley), 175–179.
- Greenfield, P. M. (1991). Language, tools, & brain: the ontogeny & phylogeny of hierarchically organized sequential behavior. *Behav. Brain Sci.* 14, 531–595. doi: 10.1017/S0140525X00071235
- Greenfield, P. M. (1998). Language, tools, & brain revisited. *Behav. Brain Sci.* 21, 159–163. doi: 10.1017/S0140525X98230962
- Greenfield, P. M., Nelson, K., and Saltzman, E. (1972). The development of rulebound strategies for manipulating seriated nesting cups: a parallel between action & grammar. *Cogn. Psychol.* 3, 291–310. doi: 10.1016/0010-0285(72)90009-6
- Haesler, S., Rochefort, C., Georgi, B., Licznarski, P., Osten, P., and Scharff, C. (2007). Incomplete and inaccurate vocal imitation after knockdown of *foxp2* in songbird basal ganglia nucleus Area X. *PLoS Biol.* 5:e321. doi: 10.1371/journal.pbio.0050321
- Haesler, S., Wada, K., Nshdejan, A., Morrissey, E. E., Lints, T., Jarvis, E. D., et al. (2004). FoxP2 expression in avian vocal learners and non-learners. *J. Neurosci.* 24, 3164–3175. doi: 10.1523/JNEUROSCI.4369-03.2004
- Hale, M., and Reiss, C. (2000). “Phonology as cognition,” in *Phonological Knowledge: Conceptual & Empirical Issues*, eds N. Burton-Roberts, P. Carr, and G. Docherty (Oxford: Oxford University Press), 161–184.
- Halle, M. (2002). *From Memory to Speech & Back*. Berlin: Mouton de Gruyter.
- Harris, J. (1969). *Spanish Phonology*. Cambridge, MA: MIT Press.
- Harris, K. D. (2015). Cortical computation in mammals and birds. *Proc. Natl. Acad. Sci. U.S.A.* 112, 3184–3185. doi: 10.1073/pnas.1502209112
- Hauser, M. D., Chomsky, N., and Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science* 298, 1569–1579. doi: 10.1126/science.298.5598.1569
- Heinz, J. N., and Idsardi, W. J. (2011). Sentence and word complexity. *Science* 333, 295–297. doi: 10.1126/science.1210358
- Heinz, J. N., and Idsardi, W. J. (2013). What complexity differences reveal about domains in language. *Top. Cogn. Sci.* 5, 111–131. doi: 10.1111/tops.12000
- Hyafil, A., Fontolan, L., Kabdebon, C., Gutkin, B., and Giraud, A.-L. (2015). Speech encoding by coupled cortical theta and gamma oscillations. *eLife* 4:e06213. doi: 10.7554/eLife.06213
- Jakobson, R., Fant, C. G. M., and Halle, M. (1952). *Preliminaries to Speech Analysis: The Distinctive Features & their Correlates*. Cambridge, MA: MIT Press.
- Jarvis, E. D. (2004). Learned birdsong and the neurobiology of human language. *Ann. N.Y. Acad. Sci.* 1016, 749–777. doi: 10.1196/annals.1298.038
- Jochaut, D., Lehongre, K., Saitovitch, A., Devauchelle, A.-D., Ollasagasti, I., Chabane, N., et al. (2015). Atypical coordination of cortical oscillations in response to speech in autism. *Front. Hum. Neurosci.* 9:171. doi: 10.3389/fnhum.2015.00171
- Johnson, C. D. (1970). *Formal Aspects of Phonological Representation*. Ph.D. thesis, University of California, Berkeley, CA.
- Johnson-Pynn, J., Frigaszy, D. M., Hirsh, E. M., Brakke, K. E., and Greenfield, P. M. (1999). Strategies used to combine seriated cups by chimpanzees (*Pan troglodytes*), bonobos (*Pan paniscus*), and capuchins (*Cebus apella*). *J. Comp. Psychol.* 113, 137–148. doi: 10.1037/0735-7036.113.2.137
- Kakishita, Y., Sasahara, K., Nishino, T., Takahashi, M., and Okanoya, K. (2009). Ethological data mining: an automata-based approach to extract behavioral units and rules. *Data Mining Knowl. Discov.* 18, 446–471. doi: 10.1007/s10618-008-0122-1
- Kao, M. H., and Brainard, M. S. (2006). Lesions of an avian basal ganglia circuit prevent context-dependent changes to song variability. *J. Neurophysiol.* 96, 1441–1455. doi: 10.1152/jn.01138.2005
- Kaplan, R., and Kay, M. (1994). Regular models of phonological rule systems. *Comput. Linguist.* 20, 331–378.
- Karttunen, L. (1998). “The proper treatment of optimality in computational phonology,” in *Proceedings of the International Workshop on Finite-State Methods in Natural Language Processing* (Ankara: Bilkent University), 1–12.
- Kluender, K. R., Lotto, A. J., and Holt, L. L. (2006). “Contributions of nonhuman animal models to understanding human speech perception,” in *Listening to Speech: An Auditory Perspective*, eds S. Greenberg and W. A. Ainsworth (Mahwah, NJ: Lawrence Erlbaum Associates), 203–220.
- Kuhl, P. K., and Miller, J. D. (1975). Speech perception by the chinchilla. *Science* 190, 69–72. doi: 10.1126/science.1166301
- Kurt, S., Fisher, S. E., and Ehret, G. (2012). Foxp2 mutations impair auditory-motor association learning. *PLoS ONE* 7:e33130. doi: 10.1371/journal.pone.0033130
- Lachlan, R. F., and Nowicki, S. (2015). Context-dependent categorical perception in a songbird. *Proc. Natl. Acad. Sci. U.S.A.* 112, 1892–1897. doi: 10.1073/pnas.1410844112



- Lai, C. S., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F., and Monaco, A. P. (2001). A forkhead-domain gene is mutated in a severe speech & language disorder. *Nature* 413, 519–523. doi: 10.1038/35097076
- Lenneberg, E. H. (1967). *Biological Foundations of Language*. New York, NY: Wiley.
- Lewicki, M., and Arthur, B. (1996). Hierarchical organization of auditory temporal context sensitivity. *J. Neurosci.* 16, 6987–6998.
- Liu, Y., Cotton, J., Shen, B., Han, X., Rossiter, S., and Zhang, S. (2010). Convergent sequence evolution between echolocating bats and dolphins. *Curr. Biol.* 20, R53–R54. doi: 10.1016/j.cub.2009.11.058
- Marler, P. (1998). “Animal communication and human language,” in *The Origin and Diversification of Language*, eds N. Jablonski and L. Aiello (San Francisco, CA: California Academy of Sciences), 1–19.
- Marler, P. (2000). “Origins of music and speech: insights from animals,” in *The Origins of Music*, eds N. L. Wallin, B. Merker and S. Brown (Cambridge, MA: MIT Press), 31–48.
- Martins, P. T., and Boeckx, C. (2014). Attention mechanisms and the mosaic evolution of speech. *Front. Psychol.* 5:1463. doi: 10.3389/fpsyg.2014.01463
- May, B., Moody, D., and Stebbins, W. (1989). Categorical perception of conspecific communication sounds by Japanese macaques, *macaca fuscata*. *J. Acoust. Soc. Am.* 85, 837–847. doi: 10.1121/1.397555
- McDonough, J. M. (2003). *The Navajo Sound System*. Dordrecht: Kluwer.
- Mesgarani, N., David, S. V., Fritz, J. B., and Shamma, S. A. (2008). Phoneme representation & classification in primary auditory cortex. *J. Acoust. Soc. Am.* 123, 899–909. doi: 10.1121/1.2816572
- Mielke, J. (2008). *The Emergence of Distinctive Features*. Oxford: Oxford University Press.
- Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.* 63, 81–97. doi: 10.1037/h0043158
- Miyagawa, S., Berwick, R. C., and Okanoya, K. (2013). The emergence of hierarchical structure in human language. *Front. Psychol.* 4:71. doi: 10.3389/fpsyg.2013.00071
- Miyagawa, S., Ojima, S., Berwick, R. C., and Okanoya, K. (2014). The integration hypothesis of human language evolution and the nature of contemporary languages. *Front. Psychol.* 5:564. doi: 10.3389/fpsyg.2014.00564
- Moorman, S., Gobes, S., Kuijpers, M., Kerkhofs, A., Zandbergen, M., and Bolhuis, J. J. (2012). Human-like brain hemispheric dominance in birdsong learning. *Proc. Natl. Acad. Sci. U.S.A.* 109, 12782–12787. doi: 10.1073/pnas.1207207109
- Nishikawa, J., Okada, M., and Okanoya, K. (2008). Population coding of song element sequence in the Bengalese finch HVC. *Eur. J. Neurosci.* 27, 3273–3283. doi: 10.1111/j.1460-9568.2008.06291.x
- Ogura, A., Ikeo, K., and Gajabori, T. (2004). Comparative analysis of gene expression for convergent evolution of camera eye between octopus and human. *Genome Res.* 14, 1555–1561. doi: 10.1101/gr.2268104
- Okanoya, K. (2004). The Bengalese finch: a window on the behavioral neurobiology of birdsong syntax. *Ann. N.Y. Acad. Sci.* 1016, 724–735. doi: 10.1196/annals.1298.026
- Ouattara, K., Lemasson, A., and Zuberbühler, K. (2009). Campbell's monkeys concatenate vocalizations into context-specific call sequences. *Proc. Natl. Acad. Sci. U.S.A.* 106, 22026–22031. doi: 10.1073/pnas.0908118106
- Patel, A. (2008). *Language & Music*. Oxford: Oxford University Press.
- Payne, K. (2000). “The progressively changing songs of humpback whales: a window on the creative process in a wild animal” in *The Origins of Music*, eds N. Wallin, B. Merker, and S. Brown (Cambridge, MA: MIT Press), 135–150.
- Petkov, C. I., and Jarvis, E. D. (2012). Birds, primates, and spoken language origins: behavioral phenotypes and neurobiological substrates. *Front. Evol. Neurosci.* 4:12. doi: 10.3389/fnevo.2012.00012
- Pfenning, A. R., Hara, E., Whitney, O., Rivas, M. V., Wang, R., Roulhac, P. L., et al. (2014). Convergent transcriptional specializations in the brains of humans and song-learning birds. *Science* 346:1256846. doi: 10.1126/science.1256846
- Prat, Y., Taub, M., and Yovel, Y. (2015). Vocal learning in a social animal: demonstrated by isolation and playback experiments in bats. *Sci. Adv.* 1:e1500019. doi: 10.1126/sciadv.1500019
- Raimy, E. (2000). *The Phonology & Morphology of Reduplication*. Berlin: Mouton de Gruyter.
- Samuels, B. (2010a). “Phonological forms: from ferrets to fingers,” in *Paper Presented at the Language Design Conference* (Montréal).
- Samuels, B. (2010b). The topology of infixation & reduplication. *Linguist. Rev.* 27, 131–176. doi: 10.1515/tlir.2010.006
- Samuels, B. (2011). *Phonological Architecture: A Biolinguistic Perspective*. Oxford: Oxford University Press.
- Samuels, B. (2015). Biolinguistics in phonology: a prospectus. *Phonol. Stud.* 18, 161–171.
- Sandler, W., Aronoff, M., Meir, I., and Padden, C. A. (2011). The gradual emergence of phonological form in a new language. *Nat. Lang. Linguist. Theory* 29, 503–543. doi: 10.1007/s11049-011-9128-2
- Saranathan, V., Hamilton, D., Powell, G., Kroodsmas, D., and Prum, R. (2007). Genetic evidence supports song learning in the three-wattled bellbird *procnias tricarunculata* (Cotingidae). *Mol. Ecol.* 16, 3689–3702. doi: 10.1111/j.1365-294X.2007.03415.x
- Schachner, A., Brady, T. F., Pepperberg, I. M., and Hauser, M. D. (2009). Spontaneous motor entrainment to music in multiple vocal mimicking species. *Curr. Biol.* 19, 1–6. doi: 10.1016/j.cub.2009.03.061
- Scharff, C., and Haesler, S. (2005). An evolutionary perspective on *foxp2*: strictly for the birds? *Curr. Opin. Neurobiol.* 15, 694–703. doi: 10.1016/j.conb.2005.10.004
- Schreiweis, C., Bornschein, U., Burguière, E., Kerimoglu, C., Schreier, S., Dannemann, M., et al. (2014). Humanized *foxp2* accelerates learning by enhancing transitions from declarative to procedural performance. *Proc. Natl. Acad. Sci. U.S.A.* 111, 14253–14258. doi: 10.1073/pnas.1414542111
- Schulz, S., Haesler, S., Scharff, C., and Rochefort, C. (2010). Knockdown of *foxp2* alters spine density in Area X of the zebra finch. *Genes Brain Behav.* 9, 732–740. doi: 10.1111/j.1601-183X.2010.00607.x
- Seyfarth, R. M., and Cheney, D. L. (1986). Vocal development in vervet monkeys. *Anim. Behav.* 34, 1640–1658. doi: 10.1016/S0003-3472(86)80252-4
- Shieber, S. (1985). Evidence against the context-freeness of natural language. *Linguist. Philos.* 8, 333–343. doi: 10.1007/BF00630917
- Shu, W., Cho, J. Y., Jiang, Y., Zhang, M., Weisz, D., Elder, G. A., et al. (2005). Altered ultrasonic vocalization in mice with a disruption in the *Foxp2* gene. *Proc. Natl. Acad. Sci. U.S.A.* 102, 9643–9648. doi: 10.1073/pnas.0503739102
- Slater, P. J. (2000). “Birdsong repertoires: their origins & use,” in *The Origins of Music*, eds N. Wallin, B. Merker, and S. Brown (Cambridge, MA: MIT Press), 49–64.
- Suh, A., Paus, M., Kieffmann, M., Churakov, G., Franke, F., Brosius, J., et al. (2011). Mesozoic retroposons reveal parrots as the closest living relatives of passerine birds. *Nat. Commun.* 2:443. doi: 10.1038/ncomms1448
- ten Cate, C., van Heijningen, C. A., and Zuidema, W. (2010). Reply to Gentner et al.: as simple as possible, but not simpler. *Proc. Natl. Acad. Sci. U.S.A.* 107, E66–E67. doi: 10.1073/pnas.1002174107
- Teramitsu, I., Poopatanapong, A., Torrisi, S., and White, S. (2010). Striatal *foxp2* is actively regulated during songbird sensorimotor learning. *PLoS ONE* 5:e8548. doi: 10.1371/journal.pone.0008548
- Teramitsu, I., and White, S. (2006). *FoxP2* regulation during undirected singing in adult songbirds. *J. Neurosci.* 26, 7390–7394. doi: 10.1523/JNEUROSCI.1662-06.2006
- Todt, D., and Hultsch, H. (1996). “Acquisition and performance of repertoires: ways of coping with diversity and versatility,” in *Ecology and Evolution of Communication*, eds D. Kroodsmas and E. Miller (Ithaca, NY: Cornell University Press), 79–96.
- van Heijningen, C. A., de Visser, J., Zuidema, W., and ten Cate, C. (2009). Simple rules can explain discrimination of putative recursive syntactic structures by a songbird species. *Proc. Natl. Acad. Sci. U.S.A.* 106, 20538–20543. doi: 10.1073/pnas.0908113106
- Wada, K., Hayase, S., Imai, R., Mori, C., Kobayashi, M., Liu, W.-C., et al. (2013). Differential androgen receptor expression and DNA methylation state in striatum song nucleus Area X between wild and domesticated songbird strains. *Eur. J. Neurosci.* 38, 2600–2610. doi: 10.1111/ejn.12258
- Wall, R. (1972). *Introduction to Mathematical Linguistics*. Englewood Cliffs, NJ: Prentice-Hall.
- Wang, Y., Brzozowska-Prechtl, A., and Karten, H. (2010). Laminar and columnar auditory cortex in avian brain. *Proc. Natl. Acad. Sci. U.S.A.* 107, 12676–12681. doi: 10.1073/pnas.1006645107
- Williams, H., and Staples, K. (1992). Syllable chunking in zebra finch (*Taeniopygia guttata*) song. *J. Comp. Psychol.* 106, 278–286. doi: 10.1037/0735-7036.106.3.278



- Wohlgemuth, M. J., Sober, S. J., and Brainard, M. S. (2010). Linked control of syllable sequence & phonology in birdsong. *J. Neurosci.* 30, 12936–12949. doi: 10.1523/JNEUROSCI.2690-10.2010
- Wytenbach, R., May, M., and Hoy, R. (1996). Categorical perception of sound frequency by crickets. *Science* 273, 1542–1544. doi: 10.1126/science.273.5281.1542
- Yip, M. (2006). “Is there such a thing as animal phonology?,” in *Wondering at the Natural Fecundity of Things: Studies in Honor of Alan Prince*. Available online at: <http://repositories.cdlib.org/lrc/prince/15>

**Conflict of Interest Statement:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Samuels. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Evolution of speech-specific cognitive adaptations

**Bart de Boer\***

*Artificial Intelligence Laboratory, Vrije Universiteit Brussel, Brussels, Belgium*

## OPEN ACCESS

### Edited by:

Cedric Boeckx,  
Catalan Institute for Research and  
Advanced Studies – Universitat de  
Barcelona, Spain

### Reviewed by:

Monica Tamariz,  
The University of Edinburgh, UK  
Gareth Roberts,  
University of Pennsylvania, USA

### \*Correspondence:

Bart de Boer,  
Artificial Intelligence Laboratory, Vrije  
Universiteit Brussel,  
Pleinlaan 2, 1050 Brussels, Belgium  
bart@ai.vub.ac.be

### Specialty section:

This article was submitted to  
Language Sciences,  
a section of the journal  
Frontiers in Psychology

**Received:** 29 July 2015

**Accepted:** 17 September 2015

**Published:** 29 September 2015

### Citation:

de Boer B (2015) Evolution of  
speech-specific cognitive  
adaptations.  
Front. Psychol. 6:1505.  
doi: 10.3389/fpsyg.2015.01505

This paper argues that an evolutionary perspective is natural when investigating cognitive adaptations related to language. This is because there appears to be correspondence between traits that linguists consider interesting and traits that have undergone selective pressure related to language. The paper briefly reviews theoretical results that shed light on what kind of adaptations we can expect to have evolved and then reviews concrete work related to the evolution of adaptations for combinatorial speech. It turns out that there is as yet no strong direct evidence for cognitive traits that have undergone selection related to speech, but there is indirect evidence that indicates selection. However, the traits that may have undergone selection are expected to be continuously variable ones, rather than the discrete ones that linguists have focused on traditionally.

**Keywords:** evolution of speech, combinatorial structure, language evolution, biology-culture co-evolution, language-specific selection

## Introduction

What properties of the brain make it language-ready? Many properties of the brain are needed, including “obvious” ones such as a supply of oxygen and nutrients. However, when cognitive scientists and linguists consider this question, they focus on properties that are to at least some extent unique to language and/or unique to humans (Hockett, 1960; Hauser et al., 2002). This is implicitly an evolutionary point of view, because what is investigated is defined in relation to what is found in related species. Here it is argued that even though the language-ready brain can be studied fruitfully without reference to its evolutionary history or without reference to comparable abilities in other species, keeping in mind the evolutionary perspective is important. After all, the behaviors and brain regions that are involved must either be similar to those of other apes, and if they are more different than would be expected from random drift, there must be an evolutionary reason, either related to language or not.

An evolutionary perspective may also help to resolve the debate about whether behaviors or mechanisms related to language are “language-specific” or “domain-general.” The problem is that one researcher’s “language-specific” is another researcher’s “domain-general,” as it is essentially arbitrary where one draws the line. From the evolutionary perspective this is even clearer as any cognitive mechanism involved in language must be based on an earlier one that was not. However, the evolutionary perspective may provide a way out, as the question of whether a trait has undergone selective pressure related to language is in principle amenable to empirical investigation (even though this may be very hard). Hence the question of whether a trait is domain-general or language-specific can be operationalized by asking whether it has undergone selective pressure related to language. In this paper, certain aspects of the language-ready brain related to speech will be considered from an evolutionary perspective. Speech is here defined as the physical signal that is used to convey language, and although this paper will focus on signals in the acoustic modality, most of what is said is true for sign language as well. Researchers with a naïve view of biology sometimes consider speech as a somewhat uninteresting process of externalization unrelated to the core properties of language

(e.g., Bolhuis et al., 2014). However, from an evolutionary perspective it is one of the most interesting aspects of language. There are three reasons for this. Firstly, speech is the aspect of language that is closest to the physical world and therefore the most likely to leave traces in the fossil record (de Boer, 2012; reviewed in, e.g., Fitch, 2010, section 2). Secondly, and related to this, speech has close analogies in other animals' behaviors. Thirdly, speech has very interesting cognitive properties (defined more precisely below) that have been proposed by some researchers as direct precursors to syntax (Carstairs-McCarthy, 1999; Studdert-Kennedy, 2005).

Two cognitive properties that allow speech but that are not found in closely related primates are precise voluntary control over the larynx and extensive vocal imitation (Ackermann et al., 2014). This paper will focus on a third aspect: *combinatorial speech*, the ability to use a small set of learned building blocks that can be recombined into an unlimited number of utterances using learned rules. This ability to deal with combinatorial structure is the basis of the phonology and phonotactics of modern human languages. Before looking at evidence for language-specific selective pressure in cognitive traits for dealing with combinatorial structure, a brief theoretical discussion is necessary about what kinds of traits can evolve, and what can therefore be expected.

## Constraints on Evolution

An important constraint on evolution is that it needs to work with what is already there: selection works on variations in the population, and this variation is caused by randomness in transmission. However, transmission in complex organisms must be relatively high-fidelity and variation must therefore be small. Evolution will consequently be gradual. However, this appears to pose no important constraints on language evolution. Precursors of many of the prerequisites for language have been inferred for the latest common ancestor with the other apes (Fitch, 2010, chapter 6). In addition processes of analogous evolution observed in other groups of species show that traits required for language that are missing in the latest common ancestor can evolve relatively quickly, for instance vocal mimicry<sup>1</sup> or (song) structure (Honda and Okanoya, 1999).

A more subtle constraint arises because language itself evolves culturally while humans evolve biologically. It has been argued that because culture changes much more quickly than biology, language provides an insufficiently stable target, and therefore arbitrary adaptations to it cannot evolve (Chater et al., 2009). Mathematical analysis shows that only the smallest stable learning biases need to evolve (Kirby et al., 2007; Smith, 2011; Thompson et al., 2012) because once a learning bias is in place cultural evolution will tend to amplify the effect of the bias, therefore masking the distinction between strong and weak biases, and thus

eliminating any selective advantage of a stronger bias. If only small learning biases can evolve, it may be that these are too small to detect experimentally.

Nevertheless larger adaptations to culturally changing language can evolve through co-evolution between language and cognition (e.g., Deacon, 1997). This can happen when cultural evolution pushes the language to become more challenging for the learners (through expanding vocabulary, or through expanding the sound system, for instance). Biological evolution can then make a small adaptation (in the sense mentioned above). This will allow for cultural evolution to make the language even more complex than before, and through continuous co-evolution a large adaptation to language can eventually evolve. Candidate for such adaptations can be the ability to produce and perceive a large range of signals (de Boer, 2015) or the ability to learn large lexicons (de Boer, 2014). Such traits are by necessity continuously variable, whereas in general traits that are considered by linguists are discrete in nature, e.g., the ability to use recursion (Bolhuis et al., 2014), or the universals considered by Evans and Levinson (2009).

## Experimental Investigation

What evidence exists for adaptations dealing with combinatorial structure? The fact that languages can be analyzed as having combinatorial structure does not necessarily mean that this structure is also represented in the brain (Zuidema and de Boer, 2009). However, evidence from for instance speech errors (Meyer, 1992), treatment of loanwords (e.g., Vendelin and Peperkamp, 2006) or poetry (Maddieson, 2008) indicate that speakers are aware of the building blocks, even if these building blocks do not necessarily correspond to phonemes. Moreover, evidence from acquisition indicates that infants learn the building blocks and the structure of their language from a very young age, both in production of intonation (Mampe et al., 2009) or phonemes (e.g., Kuhl and Meltzoff, 1996) and in perception of phonemes (Maye et al., 2002; Kuhl, 2004). This indicates that there must be cognitive mechanisms that help in learning building blocks of speech, whereas there is no evidence that these mechanisms are present in other apes. On the other hand, evidence from the emerging sign languages ABSL (Sandler et al., 2011) and CTSL (Caselli et al., 2014) indicate that combinatorial structure emerges gradually in new human languages, and that full languages can exist without much combinatorial structure.

One way to operationalize the search for traits that have undergone selection related to language is to look for brain regions that react preferentially to language. There is good evidence that there are regions specialized for processing speech and phonetic cues (e.g., Leaver and Rauschecker, 2010) and that there are even regions specialized for phonotactics (Raettig and Kotz, 2008; Rossi et al., 2011). However, there is also evidence that the precise processing of phonotactic structure is influenced by literacy (Castro-Caldas et al., 1998). Incidentally, Vendelin and Peperkamp (2006) also found that orthography influences how loanwords are treated. This raises the question of how much of the observed specialization and behavior is due to acquisition, and how much of it is indicative of evolutionary selection due to speech. DNA studies may provide insight, but although our

<sup>1</sup>For example, whereas the Black-browed Reed Warbler (*Acrocephalus bistrigiceps*) mimics 2–5 species, the closely related Marsh Warbler (*Acrocephalus palustris*) mimics more than 100 (Hamao and Eda-Fujiwara, 2004). Nevertheless their cytochrome b (mtDNA) distance is only 10–11% (Leisler et al., 1997) whereas that between humans and chimpanzees is 15–16% (Castresana, 2001).

knowledge is expanding rapidly (Dediu, 2015), we are still far from being able to relate genetic evidence with speech, the vocal tract or the brain.

Another way to operationalize the search for language-related selection is to search for behaviors that behave differently for linguistic than for non-linguistic signals. For this one needs to conduct experiments using artificial signals or to have participants devise their own signals. This allows for the possibility to include the degree of resemblance to language as a condition in the experiments and therefore to detect specialization for language. In the context of language evolution, the first such experiments were done by Galantucci (2005), but these were mostly meant to investigate emergence of signals and their structure. Since then many experiments have been done to investigate language evolution in a laboratory setting (for reviews: Galantucci, 2009; Scott-Phillips and Kirby, 2010; Kirby et al., 2014). However, few of these experiments look at speech and signals, and those that do mainly focus on cultural processes of emergence of structure (e.g., Garrod et al., 2010; Roberts et al., 2015).

Verhoef et al. (2014) however have compared two different accounts of the emergence of combinatorial structure, one based on the communication-relevant needs for distinct signals, the other on cognitive principles of processing efficiency, and found that the way human participants create structure can best be explained by the latter account. Nonetheless, this study could not determine whether these cognitive processes were language-specific or not.

A study by van der Ham and de Boer (2015a) has looked at behavior of human participants in a distributional learning task of language-like stimuli and has explicitly tested whether reproduction behavior was as predicted by a domain-general learning mechanism or by a learning mechanism specialized for language. It was found that in this case, behavior could

be explained by the domain-general mechanism. Another way to detect cognitive mechanisms that have undergone selective pressure related to speech is to look for mechanisms that behave differently for speech-like stimuli than for less speech-like stimuli. An experiment along these lines has compared category learning and reproduction in the acoustic, visual and tactile modalities (van der Ham and de Boer, 2015b) and found that humans are somewhat better in the tactile and acoustic modalities, but that there is no indication of strong specialization. Results so far therefore do not show unambiguous evidence that point to selective pressure related to language.

## Discussion

Although so far no cognitive traits that have undergone selective pressure related to speech have been identified, and although identifying the selective pressures that have shaped any trait is very difficult, nevertheless the evolutionary perspective can help structure research into the cognition of speech and language. After all, the intuitive notion of what cognitive traits are linguistically interesting corresponds to what traits have evolved under selective pressure for language. In addition the evolutionary perspective may help determine what kind of traits can have evolved and those may be rather different than the kind of traits linguists have traditionally focused on—less discrete and formal, more continuous and related to the function of language. Finally, the interdisciplinary approach that the evolutionary perspective entails has led to a number of promising new tools to investigate cognitive adaptations related to language.

## Acknowledgment

The research was funded by ERC starting grant 283435, ABACUS.

## References

- Ackermann, H., Hage, S. R., and Ziegler, W. (2014). Brain mechanisms of acoustic communication in humans and nonhuman primates: an evolutionary perspective. *Behav. Brain Sci.* 37, 529–546. doi: 10.1017/s0140525x13003099
- Bolhuis, J., Tattersall, I., Chomsky, N., and Berwick, R. C. (2014). How could language have evolved? *PLOS Biol.* 12:e1001934. doi: 10.1371/journal.pbio.1001934
- Carstairs-McCarthy, A. (1999). *The Origins of Complex Language: An Inquiry in the Evolutionary Beginnings of Sentences, Syllables, and Truth*. Oxford: Oxford University Press.
- Caselli, N., Ergin, R., Jackendoff, R., and Cohen-Goldberg, A. (2014). *The Emergence of Phonological Structure in Central Taurus Sign Language*. Padua: From Sound to Gesture.
- Castresana, J. (2001). Cytochrome b phylogeny and the taxonomy of great apes and mammals. *Mol. Biol. Evol.* 18, 465–471. doi: 10.1093/oxfordjournals.molbev.a003825
- Castro-Caldas, A., Petersson, K. M., Reis, A., Stone-Elander, S., and Ingvar, M. (1998). The illiterate brain: learning to read and write during childhood influences the functional organization of the adult brain. *Brain* 121, 1053–1063. doi: 10.1093/brain/121.6.1053
- Chater, N., Real, F., and Christiansen, M. H. (2009). Restrictions on biological adaptation in language evolution. *Proc. Natl. Acad. Sci. U.S.A.* 106, 1015–1020. doi: 10.1073/pnas.0807191106
- Deacon, T. W. (1997). *Symbolic Species, the Co-Evolution of Language and the Human Brain*. London: The Penguin Press.
- de Boer, B. (2012). Loss of air sacs improved hominin speech abilities. *J. Hum. Evol.* 62, 1–6. doi: 10.1016/j.jhevol.2011.07.007
- de Boer, B. (2014). “Biological adaptation to cultural traits,” in *The Evolution of Language: Proceedings of the 10th International Conference (EVOLANG10)*, eds E. A. Cartmill, S. Roberts, H. Lyn, and H. Cornish (Hackensack, NJ: World Scientific), 419–420.
- de Boer, B. (2015). Biology, culture, evolution and the cognitive nature of sound systems. *J. Phon.* (in press). doi: 10.1016/j.wocn.2015.07.001
- Dediu, D. (2015). *An Introduction to Genetics for Language Scientists: Current Concepts, Methods and Findings*. Cambridge: Cambridge University Press.
- Evans, N., and Levinson, S. C. (2009). The myth of language universals: language diversity and its importance for cognitive science. *Behav. Brain Sci.* 32, 429–492. doi: 10.1017/S0140525X0999094X
- Fitch, W. T. (2010). *The Evolution of Language*. Cambridge: Cambridge University Press.
- Galantucci, B. (2005). An experimental study of the emergence of human communication. *Cogn. Sci.* 29, 737–767. doi: 10.1207/s15516709cog0000\_34
- Galantucci, B. (2009). Experimental semiotics: a new approach for studying communication as a form of joint action. *Top. Cogn. Sci.* 1, 393–410. doi: 10.1111/j.1756-8765.2009.01027.x
- Garrod, S., Fay, N., Rogers, S., Walker, B., and Swoboda, N. (2010). Can iterated learning explain the emergence of graphical symbols? *Interact. Stud.* 11, 33–50. doi: 10.1075/is.11.1.04gar
- Hamao, S., and Eda-Fujiwara, H. (2004). Vocal mimicry by the Black-browed Reed Warbler *Acrocephalus bistrigiceps*: objective identification of mimetic sounds. *Ibis* 146, 61–68. doi: 10.1111/j.1474-919X.2004.00226.x



- Hauser, M. D., Chomsky, N., and Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science* 298, 1569–1579. doi: 10.1126/science.298.5598.1569
- Hockett, C. (1960). The origin of speech. *Sci. Am.* 203, 88–111. doi: 10.1038/scientificamerican0960-88
- Honda, E., and Okanoya, K. (1999). Acoustical and syntactical comparisons between songs of the white-backed munia (*Lonchura striata*) and its domesticated strain, the Bengalese finch (*Lonchura striata* var. *domestica*). *Zool. Sci.* 16, 319–326. doi: 10.2108/zsj.16.319
- Kirby, S., Dowman, M., and Griffiths, T. L. (2007). Innateness and culture in the evolution of language. *Proc. Natl. Acad. Sci. U.S.A.* 104, 5241–5245. doi: 10.1073/pnas.0608222104
- Kirby, S., Griffiths, T. L., and Smith, K. (2014). Iterated learning and the evolution of language. *Curr. Opin. Neurobiol.* 28, 108–114. doi: 10.1016/j.conb.2014.07.014
- Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nat. Rev. Neurosci.* 5, 831–843. doi: 10.1038/nrn1533
- Kuhl, P. K., and Meltzoff, A. N. (1996). Infant vocalization in response to speech: vocal imitation and developmental change. *J. Acoust. Soc. Am.* 100, 2425–2438. doi: 10.1121/1.417951
- Leaver, A. M., and Rauschecker, J. P. (2010). Cortical representation of natural complex sounds: effects of acoustic features and auditory object category. *J. Neurosci.* 30, 7604–7612. doi: 10.1523/JNEUROSCI.0296-10.2010
- Leisler, B., Heidrich, P., Schulze-Hagen, K., and Wink, M. (1997). Taxonomy and phylogeny of reed warblers (genus *Acrocephalus*) based on mtDNA sequences and morphology. *J. Ornithol.* 138, 469–496. doi: 10.1007/BF01651381
- Maddieson, I. (2008). “Segmentation, not just similarity: alliteration as a natural experiment,” in *LabPhon 11 Abstracts*, ed. P. Warren (Wellington), 89–90.
- Mampe, B., Friederici, A. D., Christophe, A., and Wermke, K. (2009). Newborns’ cry melody is shaped by their native language. *Curr. Biol.* 19, 1994–1997. doi: 10.1016/j.cub.2009.09.064
- Maye, J., Werker, J. F., and Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 82, B101–B111. doi: 10.1016/S0010-0277(01)00157-3
- Meyer, A. S. (1992). Investigation of phonological encoding through speech error analyses: achievements, limitations, and alternatives. *Cognition* 42, 181–211. doi: 10.1016/0010-0277(92)90043-H
- Raettig, T., and Kotz, S. A. (2008). Auditory processing of different types of pseudo-words: an event-related fMRI study. *Neuroimage* 39, 1420–1428. doi: 10.1016/j.neuroimage.2007.09.030
- Roberts, G., Lewandowski, J., and Galantucci, B. (2015). How communication changes when we cannot mime the world: experimental evidence for the effect of iconicity on combinatoriality. *Cognition* 141, 52–66. doi: 10.1016/j.cognition.2015.04.001
- Rossi, S., Jürgenson, I. B., Hanulíková, A., Telkemeyer, S., Wartenburger, I., and Obrig, H. (2011). Implicit processing of phonotactic cues: evidence from electrophysiological and vascular responses. *J. Cogn. Neurosci.* 23, 1752–1764. doi: 10.1162/jocn.2010.21547
- Sandler, W., Aronoff, M., Meir, I., and Padden, C. (2011). The gradual emergence of phonological form in a new language. *Nat. Lang. Linguist. Theory* 29, 503–543. doi: 10.1007/s11049-011-9128-2
- Scott-Phillips, T. C., and Kirby, S. (2010). Language evolution in the laboratory. *Trends Cogn. Sci.* 14, 411–417. doi: 10.1016/j.tics.2010.06.006
- Smith, K. (2011). Learning bias, cultural evolution of language, and the biological evolution of the language faculty. *Hum. Biol.* 83, 261–278. doi: 10.3378/027.083.0207
- Studdert-Kennedy, M. (2005). “How did language go discrete?,” in *Language Origins: Perspectives on Evolution*, ed. M. Tallermann (Oxford: Oxford University Press), 48–67.
- Thompson, B., Smith, K., and Kirby, S. (2012). “Cultural evolution renders linguistic nativism implausible,” in *The Evolution of Language: Proceedings of the 8th International Conference (EVOLANG9)*, eds T. C. Scott-Phillips, M. Tamariz, E. A. Cartmill, and J. R. Hurford (Hackensack, NJ: World Scientific), 557–558.
- van der Ham, S., and de Boer, B. (2015a). Cognitive bias for learning speech sounds from a continuous signal space seems non-linguistic. *iPerception*.
- van der Ham, S., and de Boer, B. (2015b). *Statistical Learning of Categories is Not Specialized for Speech*. San Sebastian: Interdisciplinary Advances in Statistical Learning, 153.
- Vendelin, I., and Peperkamp, S. (2006). The influence of orthography on loanword adaptations. *Lingua* 116, 996–1007. doi: 10.1016/j.lingua.2005.07.005
- Verhoef, T., Kirby, S., and de Boer, B. (2014). Emergence of combinatorial structure and economy through iterated learning. *J. Phon.* 43, 57–68. doi: 10.1016/j.wocn.2014.02.005
- Zuidema, W., and de Boer, B. (2009). The evolution of combinatorial phonology. *J. Phon.* 37, 125–144. doi: 10.1016/j.wocn.2008.10.003

**Conflict of Interest Statement:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 de Boer. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Why language really is not a communication system: a cognitive view of language evolution

Anne C. Reboul\*

CNRS UMR 5304, Laboratory on Language, Brain and Cognition (L2C2), Institute for Cognitive Sciences-Marc Jeannerod, Bron, France

## OPEN ACCESS

### Edited by:

Cedric Boeckx,  
Catalan Institute for Research  
and Advanced Studies/Universitat  
de Barcelona, Spain

### Reviewed by:

Norbert Hornstein,  
University of Maryland, USA  
David Adger,  
Queen Mary University of London, UK

### \*Correspondence:

Anne C. Reboul,  
CNRS UMR 5304, Laboratory on  
Language, Brain and Cognition  
(L2C2), Institute for Cognitive  
Sciences-Marc Jeannerod,  
67 boulevard Pinel,  
69675 Bron cedex, France  
reboul@isc.cnrs.fr

### Specialty section:

This article was submitted to  
Language Sciences,  
a section of the journal  
Frontiers in Psychology

**Received:** 28 July 2015

**Accepted:** 08 September 2015

**Published:** 24 September 2015

### Citation:

Reboul AC (2015) Why language  
really is not a communication system:  
a cognitive view of language  
evolution. *Front. Psychol.* 6:1434.  
doi: 10.3389/fpsyg.2015.01434

While most evolutionary scenarios for language see it as a communication system with consequences on the language-ready brain, there are major difficulties for such a view. First, language has a core combination of features—semanticity, discrete infinity, and decoupling—that makes it unique among communication systems and that raise deep problems for the view that it evolved for communication. Second, extant models of communication systems—the code model of communication (Millikan, 2005) and the ostensive model of communication (Scott-Phillips, 2015) cannot account for language evolution. I propose an alternative view, according to which language first evolved as a cognitive tool, following Fodor's (1975, 2008) Language of Thought Hypothesis, and was then exapted (externalized) for communication. On this view, a language-ready brain is a brain profoundly reorganized in terms of connectivity, allowing the human conceptual system to emerge, triggering the emergence of syntax. Language as used in communication inherited its core combination of features from the Language of Thought.

**Keywords:** language evolution, language-ready brain, communication, code model, ostensive model, Language of Thought, globularity

## Introduction

Language evolution has been mainly approached through the evolutionary notion of *function*. As language is routinely used in human communication, the natural assumption is that the function of language is communication. As a consequence, theories of language evolution have centered on scenarios that try to explain the kinds of selection pressures that could have triggered the emergence of this rather remarkable communication system. Inevitably given that communication is the epitome of a social phenomenon, these scenarios have been “social”<sup>1</sup>. However, seeing language as a system of communication and proposing that it has evolved as a system of communication (i.e., seeing language as being a system of communication in the *strong* sense) rather than *being merely used in communication* (i.e., seeing it as being a system of communication in the *weak* sense) raises a host of difficult issues which have to do with the very nature of language. The question of whether language is or is not a communication system in the strong sense that it evolved *for* communication is far from anecdotal as its answer strongly constrains what a language-ready brain would comprise in terms of necessary preliminary cognitive abilities.

<sup>1</sup>Számado and Szathmáry (2006) list eleven different scenarios (gossip, grooming, group bonding/ritual, hunting, language as a mental tool, pair bonding, motherese, sexual selection, song, status for information, and tool making), only one of which—language as a mental tool—is clearly and unquestionably non-social.

That language is eccentric among animal communication systems cannot be seriously disputed. It has a core combination of features—semanticity, discrete infinity, and decoupling—that is found nowhere else in nature to our present knowledge (Chomsky, 1966/2009). Relative to the evolution of language as a system of communication, this core combination of features raises two major difficulties:

- (a) Given that it gives rise to linguistic creativity (the potential production of an infinite number of different sentences with different contents), it must be explained why humans—and only humans—need to be able to communicate a potential infinity of different contents;
- (b) We are also owed an explanation of how, given that decoupling facilitates cheating and deceiving, while the evolution of communication systems is subject to stringent constraints of honesty (see Animal Communication Systems), a communication system that incorporates decoupling could ever get off the ground.

As Számado and Szathmáry (2006) have noted, none of the extant scenarios can satisfactorily answer these two questions<sup>2</sup>. Given these possibly intractable difficulties, it makes sense to reexamine the evidence in favor of the conclusion that language is a communication system and that it has evolved as a communication system.

Thus, the main goal of this paper is to assess the notion that language is a communication system in the strong sense. Here, a few words (for a complete presentation, see Animal Communication Systems) about what a communication system is are in order. The traditional view of communication systems is the code model<sup>3</sup>: the communicator encodes the message she wants to communicate, this encoded message is relayed along a channel to the receiver who decodes it and recovers the intended message. Though it is generally considered that this applies fairly well to animal communication systems<sup>4</sup> (see Animal Communication Systems), there are serious doubts that it can apply to the use of language in human communication. This is because, as has been abundantly argued (Sperber and Wilson, 1995; Carston, 2002; Recanati, 2004, 2010, following in the steps of Grice, 1989), on the whole, the semantic meaning of an utterance (the *sentence meaning*) fails to correspond exactly to what the speaker intended to communicate (the *speaker's meaning*). In other words, encoding–decoding processes are not sufficient to recover the message. While this *contextualist* position is by now largely acknowledged in both philosophy of language and linguistics, it did not penetrate the field of language evolution until very recently, when Scott-Phillips (2015) proposed a new view of language evolution. According to him, there are two main roads to the evolution of a system of communication:

- (a) Almost all communication systems (i.e., animal communication systems) are congruent with a description in terms of the code model of communication and have evolved, independently of any social abilities, either biologically or culturally, signal by signal;
- (b) A single other communication system (i.e., language) is an *ostensive* communication system<sup>5</sup>, and developed as a system on top of an *ostensive-inferential communicative mechanism* (based on the social ability of mind reading) through the establishment of a set of linguistic conventions.

In other words, while, on the code model view of language, it is *continuous with all other animal communication systems*, on the ostensive view of language, it is *discontinuous with all other animal communication systems*.

Obviously, arguments against the code model view of language as a communication system may well be inoperative against the ostensive view of language as a communication system. Thus, both theoretical frameworks will have to be examined, and we will begin with the most popular one, i.e., the code model.

## Language as a Communication System under the Code Model

As we have just seen, under the code model of communication, language is continuous with animal communication systems, and here it is useful to make a brief incursion into animal communication systems.

### Animal Communication Systems

Though whole books have been written on the subject of the evolution of communication in animals (e.g., Hauser, 1996; Oller and Griebel, 2004), their authors have often been content to use the word without giving it a precise definition. They rely on its vernacular meaning and on a rather vague notion of *information transfer*<sup>6</sup>, waving at Shannon and Weaver's (1949) quantitative definition of *information*. As pointed out by Owren et al. (2010), this is usually accompanied by the idea that this transfer of information is based on an *encoding* (on the signaler's side) and a *decoding* (on the receiver's side) process<sup>7</sup>. It is this view of communication as information transfer that makes honesty central to the evolution of communication systems.

Another line of thought was opened by Krebs and Dawkins (1984), who claim that the root of the evolution of animal communication lies in manipulation, linking the sending of a signal (the unit of animal communication systems) to a response (by the recipient) advantageous to the signaler. This

<sup>2</sup>Basically, the two questions above subsume the four questions proposed in Számado and Szathmáry (2006).

<sup>3</sup>Generally considered to have its origins in Shannon and Weaver's (1949) theory of information.

<sup>4</sup>Though for a dissenting view, see Owren et al. (2010) and Section "Animal Communication Systems".

<sup>5</sup>I leave a more complete presentation of an ostensive system of communication to Section "Language as a Communication System under the Ostensive Model" below, where Scott-Phillips' proposal will be discussed.

<sup>6</sup>Cheney and Seyfarth's (1990) classical analysis of vervet monkeys' alarm calls is an example of that strategy.

<sup>7</sup>This is where the investigation into animal communication systems meets with the code model of communication.

view of communication was clearly influential, as shown by Maynard Maynard Smith and Harper (2003, p. 3) definition of a signal:

“We define a ‘signal’ as any act or structure which alters the behaviour of other organisms, which evolved because of that effect, and which is effective because the receiver’s response has also evolved.”

In other words, the evolution of communication is not the evolution of the signal in isolation, but rather of pairs of signal–responses. This might be thought to go all the way toward a manipulation account of communication, but this is not the case. In the comments that follow, Maynard Smith and Harper outline some consequences of their definition that put them squarely on the information transfer side. First, if the signal affects the receiver’s behavior, it must do so in a way that is not, on the whole, detrimental to the receiver (otherwise selection would rapidly eliminate receptivity to it). Second, this means, on Maynard Smith and Harper’s view, that the signal must *reliably and honestly* (truthfully) convey information about the environment or about the signaler’s present state and/or future behavior. In other words, the signal evolved for its behavior-altering effects, but that does not mean that it does not carry information.

A stronger challenge to the information-based studies of communication has developed, however, through a series of papers by Owren et al. (2010, for a synthetic presentation), initially inspired by Krebs and Dawkins (1984), but presenting an alternative, rather than a mere addition to the information-based view. Owren et al.’s (2010) most convincing examples are mating signals. While mating signals have generally been analyzed in the information-based literature as transmitting information to females about males’ genetic worth<sup>8</sup>, Owren et al. (2010) propose an alternative view. Mating signals, whether visual, auditory, etc., are (in general) not informing the receiver of the signaler’s genetic quality, nor is it their function to do so. Rather, mating signals exploit pre-existing sensory preferences of females. These preferences usually have evolved in entirely different contexts (e.g., foraging for food), but once evolved they are ripe for exploitation. Thus, mating signs directly impinge on females’ sensory systems, and did not evolve for the purpose of transmitting (reliable) information about the signaler’s genetic value. It is important to note that Owren et al. (2010) do not exclude the possibility that mating signals may occasionally carry (reliable) information about the signaler’s genetic worth. Rather, if they do so, this is incidental. Their main function, which explains why they evolved, is not to signal fitness, but to attract females. This, basically, is Owren et al.’s (2010) alternative view of animal communication: its main function is not to transfer information between organisms, but to induce behaviors in the receiver that are advantageous to the communicator. Eschewing the negatively loaded term *manipulation*, they propose an *influence*-based view of animal communication.

<sup>8</sup>Given their greater biological investment in reproduction in the vertebrate and even more in the mammal species, females are generally the “choosy” sex.

This is clearly not the place to settle that debate (the interested reader is directed to the papers in Stegmann, 2013), but there one thing worth pointing out. While Owren et al. (2010) rightly deplore the detrimental effect on the animal communication literature of the (language-inspired) information-based approach, one may equally deplore the effects on the language evolution literature of an approach based on animal communication<sup>9</sup>, however, tainted by (mis-)conceptions of human language.

One of the best examples of a view of language evolution that sees language as continuous with animal communication systems, in keeping with the code model, is Millikan’s account of language and its evolution. I will mainly discuss her most recent book centering on language (Millikan, 2005).

### Millikan’s Account

Millikan’s approach to language belongs to the presently influential philosophical program aiming at “naturalizing” the mental<sup>10</sup>, concerning both mental representations and their communicative counterparts. In a move that has become classical in such programs, she aims at establishing a continuity between *natural* signs or meaning and *non-natural* signs or meaning.

The distinction between some form of natural signification (based on correlations that are, more often than not, grounded in causality) and linguistic signification is far from new, but it was given a paramount importance in Grice’s (1989) classical analysis of meaning, which is also Millikan’s main target. Grice’s strategy was to look at two uses of the verb *to mean*. Thus, he began by comparing the following examples:

- (1) These spots mean (meant) measles.
- (2) These three rings on the bus bell mean (meant) that the bus is full.

While in the first example, the verb *to mean* is used in its *natural* sense, in the second, it is used in its *non-natural* sense. Grice noted that these two uses of the verb are distinguished by the implications that one is entitled to draw from each of them. While natural meaning is *factive*, in the sense that *x means (meant) p* entails *p*, non-natural meaning (henceforth *meaning<sub>nn</sub>*) is *non-factive* in the sense that *x means<sub>nn</sub> p* does not entail *p*. On the other hand, *meaning<sub>nn</sub>* is *under voluntary control* in the sense that from *x means<sub>nn</sub> p* one can deduce that *Someone meant<sub>nn</sub> p by x*. However, natural meaning is *not under voluntary control* (it does not license the corresponding inference). So, in short, natural meaning is *factive* and *not under voluntary control* while *meaning<sub>nn</sub>* is *non-factive* and *under voluntary control*.

Grice (1989, p. 219) went further, however, and added the following definition of *meaning<sub>nn</sub>*:

“A meant<sub>nn</sub> something by *x*” is roughly equivalent to “A intended the utterance of *x* to produce some effect in an audience by means of the recognition of that intention.”

<sup>9</sup>Interestingly, in their paper, Owren et al. (2010) strongly suggest that language and animal communication are entirely disjoint phenomena, a view with which I concur.

<sup>10</sup>Initiated by Dretske (1981).



In other words, meaning<sub>nn</sub> is not only under voluntary control: additionally, the speaker has a double intention:

- The *primary* intention to produce a given effect in her audience;
- The *secondary* intention to produce that effect *via* the audience's recognition of her (the speaker's) primary intention.

Grice was at pains to emphasize that the primary intention is crucial to the definition: cases where the audience recognizes the meaning without recognizing the primary intention are not cases of meaning<sub>nn</sub>. Additionally, Grice insisted that, though meaning<sub>nn</sub> could be conventional, it did not have to be conventional. In other words, on Grice's view, normal linguistic communication is not a matter of encoding and decoding as such, but rather of recognizing the speaker's primary intention.

Grice's account of meaning<sub>nn</sub> has been Millikan's target all along her philosophical career (the first instance was Millikan, 1984). Her goal has been to show that the psychological side of Gricean meaning<sub>nn</sub> is not necessary, that linguistic communication is indeed a matter of encoding-decoding and that meaning is conventional in a utterly non-psychological sense<sup>11</sup>. In other words, what distinguishes natural signs or meaning from non-natural signs or meaning is only factivity, not volition: natural signs are factive, non-natural signs are not (the signaler may be mistaken or deceptive). Thus, Millikan's distinction between natural and non-natural meaning is wholly non-psychological.

Millikan's account of meaning centrally uses the notion of *function*, explicitly borrowed from evolutionary biology. Millikan (1984) introduces the notion of *proper function*, which is fundamentally *historical* in the following sense: it does not refer to what an entity (be it an organ or a behavior) actually does, but rather to why that entity not only exists now, but has persisted (possibly with modifications) since its emergence, in other words, why it has been selected for. So whatever the state of your heart, and regardless of whether it actually reliably pumps blood throughout your body, its proper function is to pump blood, because this is the reason why hearts have evolved, been preserved (and improved) throughout vertebrate history. Note that proper functions are not limited to biological organisms: they can also characterize artifacts of all kinds, from institutions to tools. In other words, they can be the product of either biological or cultural evolution. The essential thing is that the entity considered has a history which explains why it persisted throughout time by the function it normally performs.

On Millikan's view, language is a communication system, on a par with the other animal communication systems, as far as its evolution is concerned. She shares with Maynard Smith and Harper's (2003) definition of a signal the idea that signals evolve in tandem with responses (indeed, she views language as the solution to coordination problems in

humans<sup>12</sup>). Her idea is that the proper function of a signal is to evoke a specific response in the receiver, and that it does so through information transfer<sup>13</sup>. While clearly the notion of information transfer involved applies to natural language as well as to animal communication systems, Millikan acknowledges that linguistic signals and animal signals are different up to a point. This can be seen through her analysis (Millikan, 2004) of vervet alarm calls. In linguistic terms, such calls (e.g., the leopard call) have a double direction of fit: both world-to-signal (i.e., the call reflects the current state of the environment, e.g., the presence of a leopard in it) and signal-to-world (i.e., the signal simultaneously enjoins the recipient to give a specific response, e.g., flying to the top of the canopy). Millikan proposes to call such double-directed signals *pushmi-pullyu representations*. As Millikan (2013) herself concedes, it does not make any sense to "translate" animal signals into language. For instance, the vervet leopard alarm call is in no way equivalent to the complex sentence "There is a leopard here and you must climb to the top of the nearest tree". Though this might reflect fairly faithfully the meaning of the call, it is not a translation, because animal signals are, on the whole, holistic<sup>14</sup>: *the signal means something as a whole, not as a combination of its parts*. Indeed, as Millikan acknowledges, it is only with language that the two directions of fit (*indication* = world-to-signal and *direction* = signal-to-world) become differentiated.

Thus, Millikan acknowledges that animal signals are bidirectional, but linguistic utterances are not. I will now turn to a criticism of Millikan's position, using two kinds of arguments: general arguments regarding the very notion of a linguistic signal in signal-information/response pairings, and pragmatic arguments regarding signal-information/response pairings.

### Some Difficulties with Millikan's Position

The very structure of Millikan's theory raises major difficulties and those difficulties are all linked, in one way or another, to the essential historicity of Millikan's notion of signal, inherited from her notion of proper function. Basically, for pairings such as those that Millikan proposes as the origin of signals to occur, the signal-type, the information-type, and the response-type each have to be perennial and the repeated couplings between signals of that type, information of that type and responses of that type also have to be perennial.

<sup>12</sup>Again, though she borrows the term *coordination problem* from Lewis, 1969/2002, there is very little left of Lewis' account of convention in Millikan's theory. I will not discuss this here, as it is hardly central to my main purpose.

<sup>13</sup>Subject to the same strictures as mentioned by Maynard Smith and Harper (2003): while the response must be advantageous for the sender, it must not be generally detrimental to the receiver, otherwise selection would eliminate receptivity to it. This led Maynard Smith and Harper to the (correct) conclusion that, on such an account, the evolution of signals is bound by honesty constraints.

<sup>14</sup>There is evidence that some monkey species occasionally combine two calls to produce a modified meaning (Zuberbühler, 2002). These fairly limited phenomena are still poorly understood (for an intriguing pragmatic account, see Schlenker et al., 2014), but they hardly challenge the huge difference in compositionality between animal communication system and language (the only example of fairly sophisticated combinatoriality is birdsong, which, however, is not semantically compositional).

<sup>11</sup>Which is where she parts ways with Lewis, 1969/2002 account of convention.

This raises difficulties for the three main components of the pairings:

- Signals;
- Information;
- Responses.

I will examine them one after the other.

## Signals

A first and major question is what a linguistic signal should be. Under Millikan's broad definition of a signal, something is a signal if its proper function is to trigger a specific response in an audience, through information transfer.

Signals have to be units of communication, i.e., they have to transfer the information/produce the response in their own right. Basically, this means that they have not only to be semantic units, but also have to be communicative units (though the two normally coincide in holistic animal signals, as we shall see, they do not in language). Traditionally, it has been considered that language is *doubly articulated*<sup>15</sup>: on a rough and ready description, at the phonological level, phonemes are combined into meaningful words; at the syntactic level, words are combined into meaningful sentences. Clearly, phonemes, being semantically vacant, are not semantic units, and hence not signals. So, the first candidates for signals are words. On the face of it, they seem to be good candidates: they are perennial enough both in their forms and in their meanings<sup>16</sup>. The main problem with words is that, while they are semantic units, they are not communicative units. Though shouting "Fire!" may be a perfectly well-formed communicative act in some circumstances, most linguistic communicative acts do not correspond to isolated words. This leaves us with the sentence, understood as a utterance-type.

There is, however, a major problem with the notion that sentences are linguistic signals in the required sense. Couched as an argument:

**Lack of History Argument (Syntactic):** Given linguistic creativity, sentences are fairly often one-off, that is, they lack the history necessary to the establishment (through signal-information pairing due to repeated correlations of signal and information) of a proper function.

To show why this is the case, I will now examine (and reject) an objection to the notion that language is characterized by linguistic creativity. This objection targets one of the core properties of language, i.e., discrete infinity.

It is to the effect that humans being finite cannot be said to produce an infinity of different sentences. This *Finitude Argument* has been formulated as follows by Li and Hombert (2002, p. 196): "Theoretically the number of possible sentences in English is indefinitely large because theoretically 'the longest

English sentence' does not exist. If one chooses to describe English syntax or certain aspect of English syntax in terms of rewriting rules, one can claim that a recursive function is needed. However, one never conjoins or embeds an indefinitely large number of sentences in either spoken or written language. 'Indefinitely large number of sentences' or 'infinitely long sentences' are theoretical properties." This seems to rests on a profound misunderstanding of both discrete infinity and recursion. To see it, an analogy with another system providing discrete infinity, i.e., mathematics, is useful. Saying that, because we do not (and could not, as finite beings) produce infinitely long sentences, discrete infinity and recursion are not relevant features of language is on a par with saying that, because we do not (and could not) count to infinity, discrete infinity and recursion are not relevant features of mathematics. The argument is, to say the least, mystifying. Arguably, recursion is needed to count up to any number greater than one, just as it is needed to produce any sentence with an embedding. Once you have the relevant recursive ability, you have the theoretical possibility of counting to infinity or to producing infinitely long sentences, and whether you do it or not is utterly irrelevant. Discrete infinity is a structural, not a behavioral property. Thus, human finitude is no argument against linguistic creativity.

More crucially, the argument is no answer to our worry regarding the absence of history for sentences. Even though each human, being a finite organism, cannot produce an infinity of different sentences with different contents, linguistic creativity *as a structural property of language* allows each human to produce sentences different from all those produced before, with contents different from all of those produced before. This being so, the fact that sentences may not have the necessary history to function as signals in pairs of signal-information/response remains a central problem. In sum, human finiteness is not an argument against linguistic creativity and is no answer to the absence of history for sentences.

This, then, is the first major problem for Millikan's theory and it is, obviously, a syntactic argument. There are, however, further objections to her proposal and we will now turn to information.

## Information

Regarding information, Millikan has concentrated on two main pragmatic phenomena, illocutionary force (Millikan, 1984, 2004, 2005) and implicatures of the scalar variety (Millikan, 2005). Beginning with the former, from 1984 on, her argument has been mainly based on the pairing between sentence forms (affirmative, interrogative, imperative, etc.) and the corresponding speech acts, covering both information and response. Leaving responses aside for further discussion later on, let us concentrate on information<sup>17</sup>. The "information" pairing is between sentence

<sup>15</sup>Anderson (2013) rightly points out that this is not, strictly speaking, correct, given that there is a third articulation at the morphological level. I will ignore this complication here.

<sup>16</sup>Obviously, words change both in acoustic form and meaning with time. But while this may be a relatively quick process (taking at most decades rather than centuries or millennia), words still are stable enough to qualify as signals.

<sup>17</sup>Setting aside both the evolutionary side of Millikan's proposal as well as the pairings between signal and response, it is clear that Millikan's view of the pairings between signals and information has much in common with contemporary constructivist approaches to language acquisition in linguistics (Goldberg, 2006). I will not discuss constructivism as such here for reasons of space. Note however that *mutatis mutandis*, the pragmatic arguments against Millikan's account also apply to constructivism.

form and illocutionary act (or illocutionary force) and, as Millikan herself acknowledges (following Strawson, 1964), fairly often, an utterance can be linked to widely different illocutionary forces. Consider (3):

- (3) Peter will come tomorrow.

Depending on the circumstances, this can indeed be interpreted as a promise, a menace, a warning or a prediction. Millikan proposes to get around this problem through a multiplicity of (proper) functions. As said above, the proper function of an entity is not what it actually does but why it has persisted through time. And even if it is not *always* reliably associated with that function, it is sufficient that it is associated with it *often enough*. Thus, the existence of occasional functions different from the proper function of a sentence is not a problem. Here, it is interesting to look at Millikan's view of language change (which concerns the emergence of implicature readings). According to Millikan, if a linguistic form with a given proper function becomes associated often enough with another different function, this second function will become its new or additional proper function. In other words, the proper function of a linguistic item depends on the frequency with which this item is associated with this function and a linguistic signal can have several functions, proper or otherwise.

Let us look at an example:

- (4) The pianist played *some* Mozart sonatas.

Notoriously, this utterance can be given two interpretations:

- (5) The pianist played *at least some* (= *some and maybe all*) Mozart sonatas. [*semantic* interpretation].  
 (6) The pianist played *only some* (= *some and not all*) Mozart sonatas. [*pragmatic* interpretation].

According to Millikan, the initial proper function of (4) is to communicate (5). However, (4) is sometimes used to communicate (6) and, in time, this gives rise to a new function for (4). In addition to (5), (4) has also the function of communicating (6).

There is something mysterious about the process, however. How is it, if the proper function of (4) is to communicate (5), that, on the first occasion of its being used to communicate (6), the hearer will recognize that this is the case? Here, we turn to a first pragmatic argument:

**First Occasion Argument:** If meaning is established through repeated pairings, for such a pairing to take off, the meaning of a linguistic signal (or construction) has to be established on the occasion of its first production. A pragmatic inference will more often than not be necessary.

Note that the same argument applies to (3) above. Suppose that the initial function of (3) is to convey the illocutionary force of prediction. How does (3) acquire the additional functions of conveying the illocutionary forces of warning, menace of promise?

A final problem to do with first occasion arises for those signals who are associated with a given speaker meaning on a single occasion (one-off), as is clearly the case for some creative metaphors, such as<sup>18</sup>:

- (7) "She smiled herself to an upgrade" (Adams, 1979).  
 (8) "We laughed our conversation to an end" (Hart, 1992).

In such cases, there is no way to recover the intended meaning through semantic compositionality, and pragmatic inferences to the speaker's intentions are obviously necessary.

This is not the only difficulty, however. If a single linguistic signal can have several (proper) functions, this approach leads to widespread ambiguity in linguistic signals. And this suggests a second pragmatic argument:

**Ambiguity Argument<sup>19</sup>:** This approach supposes widespread ambiguity in linguistic signals. The resolution of that ambiguity will have to be done through pragmatic inferences.

Note, however, that what is central to Millikan's view is not the absence of context-based pragmatic inference *per se*, but rather the absence of the *Gricean* kind of pragmatic inferences. Specifically what this means is that Millikan does not reject contextualism as such but that she rejects any brand of contextualism in which either the context includes psychological representations (e.g., speaker's intentions or beliefs) or the interpretation process leads to psychological representations (e.g., *By X, the speaker meant Y*).

Here, it is interesting to go back to Millikan's analysis of natural signs. As she notes, while natural signs do not have proper functions, they are nevertheless paired with types of information: smoke and the presence of a fire, clouds and future rain, etc. However, while natural signs are factive, they are not necessarily paired bi-univocally with the information they convey. Sometimes, two different natural signs with identical forms will be associated with two different informations depending on which environment each of them occurs in. Let us take an example. It so happens that identical tracks can be left by, e.g., a small bird and a small rodent. However, in wood A, there are only birds and no rodents, while in wood B, there are only rodents and no birds. Thus, natural signs with the same form will be read (factively) as corresponding to birds in wood A and to rodents in wood B. In other words, even natural signs can be context-dependent relative to the information they convey. If this is the case, why not apply the same solution (context-dependency) to sentences? Sentences would always be associated with context types, and utterance types would correspond not to sentences, but to couples of sentences and context-types. It is these composite utterance types that would be paired with proper functions, rather than sentences in isolation. And, obviously, such composite utterance types would make perfect sense as signals in signal-information/response pairs. Note that on such

<sup>18</sup>Strangely, these are given by Goldberg (2006, p. 6) as examples of constructions, that is as examples of repeated pairings between forms and functions.

<sup>19</sup>This argument was first opposed to Millikan's view (though not under that name) by Origg and Sperber (2000).

a view (which reflects Millikan's see Millikan, 2004, Chap. 10), nothing like a Gricean "psychological" account is needed. The type of contexts concerned do not include any representation of the speaker's intentions or beliefs, or indeed, of anyone's mental states.

Let us now come back to example (3) above. As said before, a sentence such as *Peter will come tomorrow* may be understood as a promise, a menace, a warning or a prediction. Can we make sense of this in terms of utterance type, i.e., in terms of couples of sentences and (non-psychological) context types? In this specific case, it seems rather difficult to distinguish between these different illocutionary forces without appealing to mental states in both the speaker and the hearer. Presumably, leaving aside the fairly neutral speech act of prediction, what illocutionary force such an utterance will have will very much depend, not only on the speaker's intention but also on what she knows, or believes she knows, about her hearer's mental attitudes to Peter's coming. The same reasoning applies to (4): whether it will be interpreted as (5) or (6) will depend at least in part on the intention the hearer attributes to the speaker.

In other words, the requirement that the context be non-psychological seems a gratuitous complication as far as linguistic communication is concerned, as distinguishing between different illocutionary forces will, more often than not, depend on the representation of the relevant attitudes in the speaker, the hearer or both. There is yet another worry, which again, goes back to the first occasion argument. Given that utterance types are themselves composite, being couples of sentences and context types, one can also ask how such couples come into existence, leading to a higher order first occasion problem. This problem is especially acute for linguistic communication, given decoupling, which allows speakers to speak of absent or non-existent objects, introducing a further difficulty as both the signal and its referent have to be present for any association process to operate.

Hence, neither the assumption of widespread ambiguity for sentences, nor the assumption of composite utterance types, leading to semantic inflation, can work given psychological parsimony. Basically, exchanging *semantic parsimony* + *psychological inflation*, as proposed by Grice, for *semantic inflation* + *psychological parsimony*, as proposed by Millikan, is not tenable. Whether one goes for semantic parsimony or for semantic inflation, one cannot escape psychological inflation. Thus, it does not seem that composite utterance types can play the role of signals in signal-information/response pairs either.

## Responses

Let me now come to my third objection to Millikan, relative to the response type associated with the signal. Going back to Millikan's central example, speech acts, the "information" pairing is between sentence form and illocutionary act, but the "response" pairing is between sentence form and perlocutionary act. Here, it is important to see why Millikan shares with Maynard Smith and Harper the view that it is not signals that have evolved, but rather signal-response pairs. This makes sense on an evolutionary view (be it biological or cultural) because, while conveying information does not as such make sense in

evolutionary terms (information is a precious commodity, so why share it?), triggering responses in others, as long as these responses are advantageous to the signaler, makes perfect sense. So, on a view such as Millikan's, according to which language is a communication system, it seems reasonable to see linguistic signals (whatever they are) as paired with responses rather than only with information.

Millikan's main example is assertion, which, on the response side, is, according to her, paired with receiver's belief. Obviously, not all assertions lead to receiver's belief, but, as indicated above, for the pairing between assertion and receiver's belief to be established (or, in other words, for receiver's belief to be the proper function of assertion), it is sufficient that assertion be paired with belief often enough. Here, I want to discuss the appropriateness of belief as a receiver's response in an evolutionary perspective.

On the face of it, it would seem that any receiver's response in signal-response pairs should be detectable if the pairing is to have evolved<sup>20</sup>:

**Detectability of Response Argument:** for signal-response pairings to get off the ground, both the signal and the response must be detectable (respectively, by the receiver and by the signaler).

The problem with belief is not only that it is a mental state (and as such less easy to detect than a behavior or an action); it is in addition especially difficult to detect among mental states. While intentions are fairly often obvious from bodily preparation for action<sup>21</sup>, and emotions or feelings are detectable through facial expressions, belief seems to be wholly internal and not linked to any specific exteriorization<sup>22</sup>. One could argue of course that, given a belief with a certain content in her hearer, the speaker can detect its presence through his behavior interpreted *via* Theory of Mind, i.e., *via* the attribution of mental states. This, however, not only seems uncertain (see below), it also is not clear whether Millikan would agree with such a development, which is tantamount to re-introducing a rather Gricean (psychological) factor in the evolution of communication. Thus, belief appears to be a fairly strange candidate for a response in signal-response pairings.

This, however, is only a first objection. A second, and potentially more decisive objection is that responses, on such a view, have to be advantageous to the signaler (or, in the case of language, to the speaker). But belief as such is not advantageous to the speaker. Rather it is the behavioral consequences of the receiver's belief (his deciding "to act on his belief", so to speak) that may be advantageous to her. But, how exactly a hearer will act on his belief will depend on a host of other things, including his other beliefs and his desires, which strongly underdetermines the behavioral consequences of his (speaker induced) belief. Let

<sup>20</sup>No association is possible otherwise.

<sup>21</sup>It seems indeed to be the mental state that most animals or young children detect fairly easily, though perhaps in less mental terms such as *goal* or *purpose*.

<sup>22</sup>This may be because belief is phenomenologically vacant: there is nothing it is like to believe something outside of religious or quasi-religious (e.g., political, esthetic, and ethical) belief.



us suppose, for instance, that John wants to go, while Mary wants him to stay. Mary could say:

(9) It is raining.

While the belief that it is raining might indeed induce John to stay, it might equally well make him take his umbrella, phone for a taxi or do a number of other things, none of which is staying, and none of which is what Mary wishes him to do. In other words, even in such simple cases, hearer's behavioral responses are far from being obvious and there is certainly no way to predict them with any degree of certainty. And linguistic communication is of course far from being limited to such simple circumstances. In other words:

#### **Underdetermination of Behavioral Response Argument:**

In humans at least, the automaticity or even the frequency of a given response to a given linguistic signal is largely underdetermined, undermining the pairing of signals and responses.

So Millikan's choice of example, associating a linguistic signal (assertion) with a response that is a mental state (belief) can be explained through the fact that human action is not so automatic that it can be reliably associated with signals, barring imperatives in such strongly authoritative circumstances that the hearer has no choice but to comply. This, however, has two fairly negative consequences for her view of the evolution of linguistic communication: first, mental states are not the most detectable of responses, which raises a major difficulty for a signal-response pairing account such as hers (Detectability of Response Argument); second, mental states are additionally only indirectly advantageous to the speaker: they can only be advantageous to her if they lead her hearer to a behavior that she wants him to perform, but this is uncertain in most cases (Underdetermination of Response Argument).

Thus, Millikan's endeavor to "de-psychologize" language and range it among all other animal communication systems fails. We will now turn to Scott-Phillips's (2015) highly different view of language as a communication system.

## **Language as a Communication System under the Ostensive Model**

*Ostensive communication*, a notion that Scott-Phillips borrows from Sperber and Wilson (1995), corresponds to the view that human communication is intimately linked to the crucial notion of *relevance*. Relevance is a minimax notion and the communicative version of relevance goes as follows:

**Relevance:** An utterance is relevant to the extent that:

- It is less costly to interpret;
- It produces cognitive effects.

The cognitive effects produced by the interpretation of an utterance can be of three sorts: strengthening or weakening the conviction with which previous assumptions are entertained; deleting a previous assumption that is contradicted by the

new information obtained (depending on the confidence the hearer places in the speaker); producing new assumptions. The *Communicative Principle of Relevance*<sup>23</sup> says:

Every utterance carries the guarantee of its own optimal relevance.

*Optimal relevance* is achieved when the cognitive effects of an utterance balance its interpretive costs. The reason why utterances carry the guarantee of their own optimal relevance is because any utterance is an instance of *ostensive-inferential communication*. A behavior is an act of ostensive-inferential communication in as much as it makes it obvious to the receiver that the signaler has produced it with a communicative intention—this is the *ostension* part—and it is produced as evidence to be used in the inferential process through which the receiver will recover the signaler's informative intention (i.e., the content she intended to communicate)—this is the *inference* part. Thus, an act of ostensive communication guarantees that it is worthwhile for the hearer to pay attention to it. Hence, by putting ostensive-inferential communication at the heart, not only of linguistic communication, but, as we shall now see, of language evolution, Scott-Phillips is taking a position which is the opposite of Millikan's relative to language. Millikan's rejection of inferential pragmatics and insistence on signal-response pairings makes her analysis unable to deal with the semantic underdetermination that is characteristic of linguistic communication. Scott-Phillips' proposal can deal with it. But, as we shall see, it does more than that: his proposal basically reverses the problem.

At the center of Scott-Phillips' view is a distinction between *natural codes* (which correspond to what Millikan describes) and *conventional codes* (which do not). The originality of Scott-Phillips' proposal is to see ostensive communication (a shorthand for ostensive-inferential communication) not as a way of solving the problem of the semantic underdetermination of the conventional linguistic code (which would thus still be the basic root of linguistic communication), but as itself the root of human, including linguistic, communication, the conventional codes constituting language as a system being added to give human communication more expressive power. In other words (Scott-Phillips, 2015, p. 577), "there is a qualitative difference between the codes used in the code model, and the linguistic code. Put simply, one makes a type of communication possible, the other makes a different type of communication expressively powerful." Conventional codes are ubiquitous in language, being found at the phonological, lexical, syntactic and even pragmatic (e.g., politeness conventions) levels. Scott-Phillips (2015, pp. 628–629) concludes: "This view of a language as a set of conventional codes that augments ostensive communication recognizes both the pragmatic foundations of linguistic behavior, and the importance and nature of the conventions that make languages different to other, simpler cases of ostensive-inferential communication, such as points, non-linguistic vocalizations, nods of the head, and so on."

So, to sum up, on Scott-Phillips' view, language is indeed a communication system, but it is a communication system entirely

<sup>23</sup>There is also a Cognitive Principle of Relevance, which we will ignore here.

discontinuous with most if not all animal communication systems as it has evolved in the wake of abilities for ostensive communication that themselves depend on the previous evolution of a sophisticated Theory of Mind, developed on the basis of pre-existing primate abilities in social cognition, but outstripping them by far. Language itself is a collection of conventional codes, which greatly enhance the expressive power of ostensive communication, but which, nevertheless, are still in need of pragmatic inferencing, as they are, more often than not, semantically underdetermined relative to speaker's meaning.

There is no doubt that Scott-Phillips' proposal differs in many ways from Millikan's. There is, however, one point on which they seem to meet. It is highly difficult, from Scott-Phillips' presentation to see where exactly his conventional codes would differ from constructions, and, as we have seen, Millikan is also something of a constructivist. What is more, Scott-Phillips adopts a few other constructivist tenets. For instance in his fifth chapter, he rejects the Chomskyan notion of Universal Grammar<sup>24</sup>, which he sees as unnecessary. He also rejects the idea that recursion is a central factor in syntax and in linguistic creativity, though he seems to accept linguistic creativity in as much as he claims that linguistic communication is unlimited in the number of different contents language may be used to communicate.

This is not the only aspect in which Scott-Phillips' theory meets Millikan's. Another important meeting point between the two accounts is the notion of a signal-response pair as the basic communicative unit. Basically, Scott-Phillips distinguishes between *signals*, *cues*, *coercion*, *accidents*, by whether or not the behavior is designed to give rise to (designed) responses. In the case of a signal (the only communicative unit), the signal is designed to trigger the designed response (very much in keeping with Maynard Smith and Harper's definition, see Language as a Communication System under the Code Model). The cue is not designed to trigger the response, though the response is designed as a response to that type of cue. In coercion the action is designed to trigger the response, but the response is not designed as a response to that type of action. And finally in an accident, neither the accident nor the response are designed relative to one another.

Given these two important points of agreement between Millikan's and Scott-Phillips' views, it makes sense to ask whether Scott-Phillips' proposal falls foul of the objections raised above against Millikan's. Obviously, the pragmatic objections (First Occasion Argument and Ambiguity Argument) do not apply. But, as we shall see, both the Lack of History Argument and the Underdetermination of Behavioral Response Argument do apply to Scott-Phillips' theory.

As discussed above, any theory that defines communicative units as the result of pairings between signals and information/responses *ipso facto* supposes perennality in signal types, in information types, in response types and

in the pairings that link them. Scott-Phillips differs from Millikan in acknowledging from the start that the information communicated by different utterances of a given sentence will differ from occasion to occasion, and he does not explain this through widespread ambiguity. He explains it through the deep semantic underdetermination of linguistic (conventional) codes. This deep underdetermination affects speaker's meaning, and makes it necessary for the conventional codes to be supplemented by pragmatic inference. While on Scott-Phillips' model, pragmatic inference is available, this nevertheless means that different utterances of the same sentence will not be repeatedly paired with the same information. This leads us to a *pragmatic* version of the Lack of History Argument:

**Lack of History Argument (Pragmatic):** Given semantic underdetermination, the speaker meaning attributed to one utterance of a given sentence will often be one-off, that is, it will not necessarily be attributed to any other utterance of the same sentence. In other words, utterances lack the semantic stability necessary to the establishment of a conventional code.

Let us now turn to responses. The example Scott-Phillips gives of a signal is of a man pushing a woman down under the eyes of another colleague, who laughs in response<sup>25</sup>. The pushing was intended to be seen by the laughing colleague and thus it is a communicative signal designed to trigger as its designed response the laughter. While this example is certainly not susceptible to the Detectability of Response Argument (laughter being detectable), it nevertheless is susceptible to the Underdetermination of Behavioral Responses Argument. Rather obviously, the intended receiver might have remonstrated instead of laughing.

Thus, while Scott-Phillips offers an original and attractive theory, it falls foul of some of the same difficulties that plague Millikan's. My diagnosis is that this can basically be explained by the fact that these difficulties come from what the basic proposition shared by the two views is: that language is a communication system.

## The Language-ready Brain

The proposition that language is a communication system imposes obvious constraints on the abilities that have to pre-exist for language to get off the ground. Unsurprisingly, given that communication is the epitome of a social phenomenon, these abilities are social. On the code model, the main constraint is *honesty* (see Animal Communication Systems)—and this is all the more important in language, given the opportunities for cheating that decoupling offers. This has led to the view that *altruism*, as a phylogenetic pro-social tendency, is a prerequisite for human linguistic communication and for language evolution. On the ostensive model, linguistic communication and language

<sup>24</sup>Though Scott-Phillips acknowledges that there may well be linguistic universals, he proposes to explain them through Cultural Attraction Theory, not Universal Grammar.

<sup>25</sup>He rightly notes that under his view one and the same behavioral token could be at one a signal, a cue, coercion and an accident depending on who observes it.

evolution basically depend on the preexistence of a Theory of Mind of some kind. However, as we have just seen, the notion that language is a communication system in the strong sense that it evolved *for* communication is implausible in view of the difficulties it meets with. One fairly obvious suggestion to account for its use in human communication is that it originally evolved for entirely different purposes and was then exapted (Gould and Vrba, 1982) for communication. Determining what those purposes were is a prerequisite for determining which pre-existing abilities should comprise the language-ready brain.

Here, recall the two questions listed in Section “Introduction”, and more specifically the question of why humans—and only humans—need a system of communication that allows them to communicate a potential infinity of different contents. Communication is rife in nature, but language is unique. This immediately raises a further question: where does this infinity of different contents come from? As Millikan (2013) rightly notes, human cognitive sophistication is also unique. Thus, one potential answer to the question above is that a cognitively sophisticated species needs an appropriately sophisticated system of communication. What this means basically is that human intelligence, rather than human sociability, is the key to language. We can go one step further, however, following Fodor and Pylyshyn (2015), and note that thoughts and sentences share the same structural organization: just as sentences structurally compose words in a creative way, thoughts structurally compose concepts in a creative way. Language is creative, because thought is creative. Or, in Fodor and Pylyshyn (2015, p. 89) words, “That thoughts and sentences match up so nicely is part of why you can sometimes say what you think and vice versa.”

Hinzen (2013) goes farther and proposes that language is primarily an internal tool for thought and that syntax is the root of the semantic and propositional organization of thought in humans, and hence of the specificity of human thought, compared with non-human animal thought. This, it should be clear, also answers the second question raised in Section “Introduction”, i.e., why does linguistic communication allows decoupling which clearly facilitates cheating and deceiving? On a view in which language evolved for thought, discrete infinity, semanticity and decoupling are not structural features specific to linguistic communication, they are structural features specific to thought *and in no way dependent on whether language is externalized for communication* or not. Note that discrete infinity and decoupling, which are obvious embarrassments for a theory of language as a communicative system in the strong sense, raise no problem for a theory of thought: obviously, discrete infinity and decoupling are ways of exponentially increase thought production, while the question of honesty does not arise for thought. So, basically, all of this comes to the suggestion that language did not evolve *for* communication, it evolved *for* thought (as advocated by Chomsky: see, Chomsky, 2014). It allows us to construct what medieval philosophers (Panaccio, 1999) called *complex concepts*, propositions, judgments, etc. This is essentially Fodor’s *Language of Thought Hypothesis* (Fodor, 1975, 2008). Language was then externalized for communication,

and its externalized version inherited its core combination of properties.

While this explains why language is such an exotic communication system<sup>26</sup>, it does not, in and off itself, explain why such a sophisticated system of thought is unique to humans: why did humans—and only humans—need such a sophisticated system of thought? Another human specificity is the richness of the human conceptual system. While some core conceptual mechanisms (Carey, 2009) may be shared with other species (notably with great apes, see Gómez, 2004), the extent of the human conceptual system is unique. This difference is obvious from the very limited size of the vocabulary acquired by animals engaged in language research programs ( $\approx 300$  words) as compared to the size of human vocabularies (300 words at 3-year-old, 6,000 at 6-year-old, and around 200,000 at 18-year-old; for animals’ lexicon, see Anderson, 2004; for human lexicons at different ages, Bloom, 2000). While some of the difference may be due to externalized language itself, it is highly unlikely that this is the only explanation. Indeed, other considerations militate against an identity between the human conceptual system and non-human animal conceptual systems, including those of other primates. While monkeys can learn to visually categorize images (Fabre-Thorpe, 2003), they usually do so after intensive training (involving thousands of trials), by contrast with young children who learn new concepts (and the corresponding words) instantaneously (Bloom, 2000; Waxman, 2004). Apart from any reservation, one might have to consider visual categorization as a proof of concept possession, this hints at highly different mechanisms of conceptualization. Finally, though Orangutans may be an exception (Vonk and MacDonald, 2004), other great ape species, though able of categorical discrimination at different levels of abstraction, present a highly different profile from what is found in humans: the intermediate or basic level (roughly corresponding to the level of the species), which is by far the most easily accessed in humans, is the most difficult for them (it is the level at which they fail to transfer learned categories: see Vonk and MacDonald, 2002). Thus, all in all, there are good reasons to doubt that conceptualization follows the same path in humans and in non-human animals.

Here, the hypothesis is that different mechanisms operate in human conceptualization, explaining why humans have conceptual repertoires so much wider than other species do. Having a huge conceptual system, however, can only be useful if the concepts can be assembled into complex concepts or propositions (thoughts). While association can bind concepts together (and does in both human and non-human animals), its power is limited: at most, it could lead to sequences of concepts. On the other hand, syntax allows structured and compositional mental representations to emerge (see Hinzen, 2013; Fodor and Pylyshyn, 2015 for more detailed arguments). The suggestion is thus that syntax emerged at the mental level to organize concepts into (propositional) thoughts.

<sup>26</sup>Because it is discontinuous from all other animal communication systems not merely by being the only ostensive communication system (Scott-Phillips, 2015, and see Language as a Communication System under the Ostensive Model), but also by *not* having evolved *for* communication in the first place.

The obvious question is what led to the emergence of such different mechanisms of conceptualization in humans. Here, it is hard to avoid speculation, but, as argued by Boeckx and Benítez-Burraco (2014a,b) and Benítez-Burraco and Boeckx, 2015, there are important differences between modern humans and the Neanderthal/Denisovan branch, the main one being the globularity of the modern human skull compared to the elongated shape of the Neanderthal/Denisovan skull. Boeckx and Benítez-Burraco hypothesize that this change in shape corresponds to major brain reorganization leading to greater cerebral connectivity. Additionally, they point out that this change is not so much due to the enlargement of the frontal lobes as to the expansion and reorganization of parietal areas. Now, one of the peculiarities of the Neanderthal/Denisovan skull is the so-called Neanderthal bun, a bump on the occipital part of the skull, corresponding to the primary visual cortex (V5, Brodmann area 17). While there is clearly more to conceptualization than perception, it is hard not to link the change in human conceptualization to the specificity of human perceptual preference for global processing of visual scenes (Navon, 1977; Kimschi et al., 2005) as opposed to non-human primate perceptual preference for local processing (Fagot and Deruelle, 1997; Fagot and Tomonaga, 1999; Fagot et al., 1999, 2001). It is not impossible that the reorganization assumed by Boeckx and Benítez-Burraco also concerned the occipital area with capital consequences on visual preferences in modern humans, leading to improved conceptualization.

## Conclusion

While most of this paper has been dedicated to show that language is not a communication system in the strong sense (i.e., it did not evolve *for* communication) and to outline an alternative cognitive account and its consequences for the language-ready brain, I do not want to close it without saying a word about the externalization of a pre-existing language for communication. While I criticized Scott-Phillips' (2015) account above, I nonetheless think that it makes a lot of sense as an account of the *externalization* of language. As noted

above (see Language as a Communication System under the Ostensive Model), Scott-Phillips, following Sperber and Wilson (1995), sees language as a sophisticated brand of ostensive communication and proposes that a less sophisticated and wholly unconventionalized brand of ostensive communication preceded the formation of linguistic conventions. As he notes, all ostensive communication rests on mind-reading abilities. While he is content to suppose that such abilities somehow derived from previous primate social abilities, this is unlikely for a number of reasons, the main one being that primates seem pretty restricted in that area. At best, chimpanzees may be able of recognizing intentions (and even that is in dispute: for a general presentation, see Lurz, 2011). Scott-Phillips gives no reason why mind-reading abilities would make such a jump in humans. The hallmark of human mind reading is that, in Dennett's (1987) words, it involves higher-order intentions (e.g., *Peter believes that Mary believes that p*), in other words, metarepresentations. Now metarepresentation crucially depends on recursion as the representations involved are structurally recursive. Under the scenario I propose, the development of recursive syntax in the Language of Thought allowed humans to develop mind-reading abilities far in excess of anything to be found in non-human species. This allowed humans to indulge in ostensive communication, leading to linguistic conventions, roughly along the lines indicated by Scott-Phillips. Note, in addition, that under this revised scenario, acquiring words means matching words to pre-existing concepts (as largely recognized in the lexical acquisition literature, Bloom, 2000). This largely dispels the problem described in the semantic version of the Lack of History Argument I opposed to Scott-Phillips' view (see Language as a Communication System under the Ostensive Model). While speaker's meaning has to be stable on the view that language evolved as a communication system (and clearly is not), sentence meaning stability is quite enough to ensure the establishment and learning of lexical conventions on the view that language is a communication system only in a weak sense. This is because language as a communication system in the weak sense can piggyback on the pre-existing conceptual system and Language of Thought, that, as argued by Hinzen (2013), fixes referential and propositional meaning.

## References

- Adams, D. (1979). *Hitchhiker's Guide to the Galaxy*. New York: Harmony Books.
- Anderson, S. R. (2004). *Doctor Dolittle's Delusion: Animals and the Uniqueness of Language*. London: Yale University Press.
- Anderson, S. R. (2013). "What is special about the human language faculty and how did it get that way?," in *The Evolutionary Emergence of Language: Evidence and Inference*, eds R. Botha and M. Everaert (Oxford: Oxford University Press), 18–41.
- Benítez-Burraco, A., and Boeckx, C. (2015). Possible functional links among brain- and skull-related genes selected in modern humans. *Front. Psychol.* 6:794. doi: 10.3389/fpsyg.2015.00794
- Bloom, P. (2000). *How Children Acquire the Meaning of Words*. Cambridge, MA: The MIT Press.
- Boeckx, C., and Benítez-Burraco, A. (2014a). The shape of the human language-ready brain. *Front. Psychol.* 5:282. doi: 10.3389/fpsyg.2014.00282
- Boeckx, C., and Benítez-Burraco, A. (2014b). Globularity and language-readiness: generating new predictions by expanding the set of genes of interest. *Front. Psychol.* 5:1324. doi: 10.3389/fpsyg.2014.01324
- Carey, S. (2009). *The Origin of Concepts*. New York: Oxford University Press.
- Carston, R. (2002). *Thoughts and Utterances: The Pragmatics of Explicit Communication*. Oxford: Blackwell Publishing.
- Cheney, D., and Seyfarth, R. M. (1990). *How Monkeys See the World*. Chicago: University of Chicago Press.
- Chomsky, N. (1966/2009). *Cartesian Linguistics: A Chapter in the History of Rationalist Thought*. Cambridge: Cambridge University Press.
- Chomsky, N. (2014). *Language Architecture and its Import for Evolution*. Available at: <http://www.reddit.com/r/generativelinguistics/comments/2szxp8/language-architectureanditsimportfor/> [accessed August 27, 2015].
- Dennett, D. C. (1987). *The Intentional Stance*. Cambridge, MA: The MIT Press.
- Dretske, F. I. (1981). *Knowledge and the Flow of Information*. Cambridge, MA: The MIT Press.
- Fabre-Thorpe, M. (2003). Visual categorization: accessing abstraction in non-human primates. *Philos. Trans. R. Soc. Lond. B Sci.* 358, 1215–1223. doi: 10.1098/rstb.2003.1310
- Fagot, J., and Deruelle, C. (1997). Processing of global and local visual information and hemispheric specialization in humans (*Homo sapiens*) and baboons (*Papio papio*). *J. Exp. Psychol.* 23, 429–442.



- Fagot, J., Deruelle, C., and Tomonaga, M. (1999). Perception des dimensions globale et locale de stimuli visuels chez le primate. *Primatologie* 2, 61–77.
- Fagot, J., and Tomonaga, M. (1999). Comparative assessment of global-local processing in humans (*Homo sapiens*) and chimpanzees (*Pan troglodytes*): Use of a visual search task with compound stimuli. *J. Comp. Psychol.* 113, 3–12. doi: 10.1037/0735-7036.113.1.3
- Fagot, J., Tomonaga, M., and Deruelle, C. (2001). "Processing of the global and local dimensions of visual hierarchical stimuli by humans (*Homo sapiens*), chimpanzees (*Pan troglodytes*), and baboons (*Papio papio*)," in *Primate Origins of Human Cognition and Behavior*, ed. T. Matsuzawa (Berlin: Springer), 87–103.
- Fodor, J. A. (1975). *The Language of Thought*. New York: Thomas Y. Crowell Company.
- Fodor, J. A. (2008). *LOT2: The Language of Thought Revisited*. Oxford: Clarendon Press.
- Fodor, J. A., and Pylyshyn, Z. W. (2015). *Minds Without Meaning: An Essay on the Content of Concepts*. Cambridge, MA: The MIT Press.
- Goldberg, A. E. (2006). *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.
- Gómez, J. C. (2004). *Apes, Monkeys, Children and the Growth of Mind*. Cambridge, MA: Harvard University Press.
- Gould, S. J., and Vrba, E. S. (1982). Exaptation: a missing term in the science of form. *Paleobiology* 8, 4–15.
- Grice, H. P. (1989). *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.
- Hart, J. (1992). *Sin*. New York: Ivory Books.
- Hauser, M. D. (1996). *The Evolution of Communication*. Cambridge, MA: The MIT Press.
- Hinzen, W. (2013). Narrow syntax and the language of thought. *Philos. Psychol.* 26, 1–23. doi: 10.1080/09515089.2011.627537
- Kimschi, R., Hadad, B., Behrmann, M., and Palmer, S. E. (2005). Microgenesis and ontogenesis of perceptual organization : evidence from global and local processing of hierarchical patterns. *Psychol. Sci.* 16, 282–290. doi: 10.1111/j.0956-7976.2005.01529.x
- Krebs, J., and Dawkins, R. (1984). "Animal signals: mind-reading and manipulation," in *Behavioural Ecology: An Evolutionary Approach*, 2nd Edn, eds J. R. Krebs and N. B. Davies (Oxford: Blackwell), 380–402.
- Lewis, D. (1969/2002). *Convention: A Philosophical Study*. Oxford: Backwell Publishers.
- Li, C. N., and Hombert, J. M. (2002). "On the evolutionary origin of language," in *Mirror Neurons and the Evolution of Brain and Language*, eds M. I. Stamenov and V. Gallese (Philadelphia: John Benjamins Publishing Company), 175–206.
- Lurz, R. W. (2011). *Mindreading Animals: The Debate over What Animals Know about Other Minds*. Cambridge, MA: The MIT Press.
- Maynard Smith, J., and Harper, D. (2003). *Animal Signals*. Oxford: Oxford University Press.
- Millikan, R. G. (1984). *Language, Thought, and Other Biological Categories*. Cambridge, MA: Bradford Books.
- Millikan, R. G. (2004). *Varieties of Meaning: The 2002 Jean Nicod Lectures*. Cambridge, MA: The MIT Press.
- Millikan, R. G. (2005). *Language: A Biological Model*. Oxford: Clarendon Press.
- Millikan, R. G. (2013). "Natural information, intentional signs and animal communication," in *Animal Communication Theory: Information and Influence*, ed. U. E. Stegmann (Cambridge: Cambridge University Press), 3316–3603. (Kindle edition).
- Navon, D. (1977). Forest before trees: the precedence of global features in visual perception. *Cogn. Psychol.* 9, 353–383. doi: 10.1016/0010-0285(77)90012-3
- Oller, D. H., and Griebel, U. (eds). (2004). *Evolution of Communication Systems: A Comparative Approach*. Cambridge, MA: The MIT Press.
- Origg, G., and Sperber, D. (2000). "Evolution, communication and the proper function of language," in *Evolution and the Human Mind: Language, Modularity and Social Cognition*, eds P. Carruthers and A. Chamberlain (Cambridge: Cambridge University Press), 140–169.
- Owren, M. J., Rendall, D., and Ryan, M. J. (2010). Redefining animal signaling: influence versus information in communication. *Biol. Philos.* 25, 755–780. doi: 10.1007/s10539-010-9224-4
- Panaccio, C. (1999). *Le Discours Intérieur: de Platon à Guillaume d'Ockham*. Paris: Le Seuil.
- Recanati, F. (2004). *Literal Meaning*. New York: Cambridge University Press.
- Recanati, F. (2010). *Truth-Conditional Pragmatics*. Oxford: Oxford University Press.
- Schlenker, P., Chemla, E., Arnold, K., Lemasson, A., Ouattara, K., Keenan, S., et al. (2014). Monkey semantics: Two 'dialects' of Campbell's monkey alarm calls. *Linguist. Philos.* 37, 439–501. doi: 10.1007/s10988-014-9155-7
- Scott-Phillips, T. (2015). *Speaking Our Minds: Why Human Communication is Different, and How Language Evolved to Make it Special*. Basingstoke: Palgrave Macmillan.
- Shannon, C. E., and Weaver, W. (1949). *A Mathematical Model of Communication*. Urbana: University of Illinois Press.
- Sperber, D., and Wilson, D. (1995). *Relevance: Communication and Cognition*. Oxford: Basil Blackwell.
- Stegmann, U. E. (ed.). (2013). *Animal Communication Theory: Information and Influence*. Cambridge: Cambridge University Press. (Kindle edition).
- Strawson, P. F. (1964). Intention and convention in speech acts. *Philos. Rev.* 73, 439–460. doi: 10.2307/2183301
- Számado, S., and Szathmáry, E. (2006). Selective scenarios for the emergence of natural language. *Trends Ecol. Evol.* 21, 555–561. doi: 10.1016/j.tree.2006.06.021
- Vonk, J., and MacDonald, S. (2002). Natural concepts in a juvenile gorilla (*Gorilla Gorilla Gorilla*) at three levels of abstraction. *J. Exp. Anal. Behav.* 78, 315–332. doi: 10.1901/jeab.2002.78-315
- Vonk, J., and MacDonald, S. E. (2004). Levels of abstraction in orangutan (*Pongo abelii*) categorization. *J. Comp. Psychol.* 118, 3–13. doi: 10.1037/0735-7036.118.1.3
- Waxman, S. R. (2004). "Everything had a name, and each name gave birth to a new thought: links between early word learning and conceptual organization," in *Weaving a Lexicon*, eds D. G. Hall and S. Waxman (Cambridge, MA: The MIT Press), 295–335.
- Zuberbühler, K. (2002). A syntactic rule in forest monkey communication. *Anim. Behav.* 63, 293–299. doi: 10.1006/anbe.2001.1914

**Conflict of Interest Statement:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Reboul. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Brain readiness and the nature of language

**Denis Bouchard \***

*Département de linguistique, Université du Québec à Montréal, Montreal, QC, Canada*

## OPEN ACCESS

### Edited by:

Cedric Boeckx,  
*Universitat de Barcelona, Spain*

### Reviewed by:

Christopher I. Petkov,  
*Newcastle University, UK*  
Wendy Sandler,  
*University of Haifa, Israel*

### \*Correspondence:

Denis Bouchard,  
*Département de linguistique,  
Université du Québec à Montréal, C.P.  
8888, Succursale Centre-Ville,  
Montreal, QC H3C 3P8, Canada  
bouchard.denis@uqam.ca*

### Specialty section:

*This article was submitted to  
Language Sciences,  
a section of the journal  
Frontiers in Psychology*

**Received:** 26 June 2015

**Accepted:** 26 August 2015

**Published:** 09 September 2015

### Citation:

Bouchard D (2015) Brain readiness  
and the nature of language.  
*Front. Psychol.* 6:1376.  
doi: 10.3389/fpsyg.2015.01376

To identify the neural components that make a brain ready for language, it is important to have well defined linguistic phenotypes, to know precisely what language is. There are two central features to language: the capacity to form signs (words), and the capacity to combine them into complex structures. We must determine how the human brain enables these capacities. A sign is a link between a perceptual form and a conceptual meaning. Acoustic elements and content elements, are already brain-internal in non-human animals, but as categorical systems linked with brain-external elements. Being indexically tied to objects of the world, they cannot freely link to form signs. A crucial property of a language-ready brain is the capacity to process perceptual forms and contents offline, detached from any brain-external phenomena, so their “representations” may be linked into signs. These brain systems appear to have pleiotropic effects on a variety of phenotypic traits and not to be specifically designed for language. Syntax combines signs, so the combination of two signs operates simultaneously on their meaning and form. The operation combining the meanings long antedates its function in language: the primitive mode of predication operative in representing some information about an object. The combination of the forms is enabled by the capacity of the brain to segment vocal and visual information into discrete elements. Discrete temporal units have order and juxtaposition, and vocal units have intonation, length, and stress. These are primitive combinatorial processes. So the prior properties of the physical and conceptual elements of the sign introduce combinatoriality into the linguistic system, and from these primitive combinatorial systems derive concatenation in phonology and combination in morphosyntax. Given the nature of language, a key feature to our understanding of the language-ready brain is to be found in the mechanisms in human brains that enable the unique means of representation that allow perceptual forms and contents to be linked into signs.

**Keywords:** language evolution, evolvability, linguistic signs, brain readiness, self-organization

## Introduction

The main point of this paper is that the central trait of human language is the capacity to form signs by linking perceptual forms and meanings<sup>1</sup>. This predicts that the core mechanisms that make a brain ready for language are those that enable this capacity. Moreover, switching from a computational view of language to a sign-based theory provides a unified approach to the

<sup>1</sup>This paper presents the main hypotheses exposed in Bouchard (2013). I have therefore borrowed substantially from that text without indicating it by quotation marks or references in order not to overburden the readers.

functioning of the main subsystems of language. The perceptual and conceptual substances of signs create a system that reaches a level of such complexity that it triggers self-organization, deriving specific properties of signs, as well as the basic structuring of language in its phonology, semantics, and syntax.

## The Core Competence for Language

A language-ready brain raises two evolutionary puzzles: a puzzle of emergence and a puzzle of design (Hoefler, 2009, p. 1). The puzzle of emergence addresses the problem of bridging the gap from a stage where our ancestors had no language to a stage where they had language as we know it today. How and why did language emerge in humans and not in other species?

Lewontin (1998) raises strong doubts about the possibility of reconstructing the evolutionary history and the causal mechanisms of the acquisition of linguistic competence (and cognition in general). He emphasizes the near impossibility to come up with evidence “that there was heritable variation for, say, linguistic ability, in our remote ancestors when the human species was still evolving into its present form and that those who possessed this ability, in the remote past, left more offspring by virtue of that ability” (p. 111). So it is extremely difficult for the standard theory of evolution by natural selection to inform us on how language, and more generally, cognition arose and spread and changed. As he points out, humans had an ancestor in common with the chimpanzee and the gorilla about 10 million years ago. So 20 million years of evolution separate us from our closest relatives. During that period, “a major difference in the consequences of cognitive power has taken place during human evolution that makes the cognitive difference between gorillas and chimpanzees trivial compared to our cognitive distance from them” (p. 116). Evolved forms may diverge very dramatically in a relatively short period of time. Lewontin gives the example of cows, goats, and deer that differentiated 10 million years ago. Therefore, it is unlikely that we can determine—even approximately—when our linguistic capacity emerged in our ancestry. In addition, a trait may derive from analogy just as well as from homology. Moreover, we cannot measure the actual reproductive advantages of cognition or language. Fossils, furthermore, are of very little help concerning cognition, and often we cannot even be sure whether a fossil is from an ancestor or some relative on another branch of the bush-like relations between species. So we cannot tell what our immediate non-linguistic ancestors were like cognitively. Almost two decades after the publication of his paper, the problem still appears to be substantial, though advances in our knowledge of genes open some research avenues concerning heritable variation, even for remote ancestors.

Nevertheless, there is room for testable theories about what language is, what brain mechanisms this requires, and whether some of these brain mechanisms are unique to humans at least compared to other current species. As we progress in our understanding of the human brain, we can compare it with the neuro-anatomy of related species and see how they differ in form and function. We can pinpoint some current neurological distinctive trait(s) that enable(s) language, and hence determine

WHAT made language emerge. Regarding WHEN and HOW the organism evolved to get that change, we can only speculate. But at least we can elaborate a theory that passes the test of evolvability: if a theory can show how some actual neuro-anatomical element enables language as we know it, then that theory is in accord with the fact that an organism with a language-ready brain is an evolvable organism, because this neuro-anatomical element can indeed develop according to the laws and principles of biological evolution, since it exists in human brains. Moreover, the nature of the neuro-anatomical trait can give us an indication of what it could have come from. This is particularly the case if language is a side effect of the neuro-anatomical trait, as I argue below: the other functional effects of the trait can further restrict the possible scenarios.

This brings us to the second evolutionary puzzle, the question of design: how and why did language evolve with the properties that we observe rather than some other set? To identify the components that make a brain ready for language, neuroscientists must know precisely what such a brain must do, hence ultimately, what language is. Not that the brain mechanisms will somehow be analogical to the functional aspects of language: examples abound where it has been shown that the neural substrates or the mechanisms supporting behavior, are not predicted by psychological models. However, we must understand precisely what language is and have well-defined linguistic phenotypes to search for the neural substrates that enable these phenotypes.

There are numerous properties that have been attributed to language. Many have been recently proposed and many are not widely accepted because they depend on narrow theoretical assumptions. It would be a formidable task to look at hundreds of properties in exploring the language-readiness of the brain, and probably futile in many instances since the properties are probably ephemeral. It is more productive to investigate two properties of language for which there is a long-standing and broad consensus among scholars—the capacity to form signs (words, morphemes), and the capacity to combine them into complex structures:

“at least two basic problems arise when we consider the origins of the faculty of language [...]: first, the core semantics of minimal meaning-bearing elements, including the simplest of them; and second, the principles that allow infinite combinations of symbols, hierarchically organized, which provide the means for use of language in its many aspects” (Chomsky, 2005, p. 4).

If we can explain how the brain is ready for these two basic properties, how it enables them, we are heading in the right direction. However, if we consider what the founder of the most prominent theoretical model in linguistics says about the evolution of these two properties, the prospects look rather dim. Concerning the capacity to form signs, Chomsky (2010) says that it is “of totally mysterious origin.” Moreover, though he has contributed to a very influential paper on the origin of linguistic combinatoriality (Hauser et al., 2002), Chomsky and some of his colleagues now believe that the origin of combinatoriality is also a mystery, as indicated in the very title of their paper: “The mystery of language evolution” (Hauser et al., 2014).

The problem is further amplified by the fact that, despite recent attempts to limit it, the current model still relies on a large set of innate, language-specific conditions—Universal Grammar (UG)—which is a repertory of unexplained properties (Chomsky, 2007, p. 19)<sup>2</sup>. UG is therefore a highly problematic component from an explanatory point of view, since the richer the set of language-specific brain features, the harder it will be to account for it: “Aspects of the computational system that do not yield to principled explanation fall under UG, to be explained **somehow in other terms** [my emphasis, DB], questions that may lie beyond the reach of contemporary inquiry, Lewontin (1998) has argued” (Chomsky, 2007, p. 24). This is as close as one can get to saying that UG is also an unsolved mystery, maybe even an unsolvable one<sup>3</sup>.

The three mysteries are not simply subcases of the difficulty to reconstruct evolutionary history and the causal mechanisms of the acquisition of linguistic competence: they are also problems of evolvability. The UG model appears incapable of providing a principled explanation based on some neuro-anatomical elements that would account for the numerous language-specific components it postulates. Brain readiness and evolvability are closely linked, so evolvability is an important test for linguistic theories: the traits that a linguistic theory requires of the human brain must be highly plausible according to the known laws and principles of biological evolution. We may not be able to trace the evolutionary path of how language emerged, but we can evaluate the degree of evolvability of a linguistic model, its plausibility given known laws of evolution.

<sup>2</sup>For instance, here is an illustrative sample of UG elements taken mostly from Hornstein and Boeckx (2009) and Narita and Fujita (2010):

- endocentricity labeling;
- c-command;
- uninterpretable features and specifications about which elements they may attach to and when;
- numerous functional categories;
- an intractable number of micro or macro parameters distributed over different modules;
- binding conditions for pronouns;
- displacement;
- agreement;
- constituency;
- cycle/phase bounding nodes;
- Phase Impenetrability;
- Transfer;
- locality conditions (Ross's Problem: why locality holds for Move but not pronominalization);
- condition on theta assignment: arguments must be initially merged in theta-positions;
- Linearize: there has to be a procedure Linearize, with something like the Linear Correspondence Axiom (Kayne, 1994) to constrain it.

<sup>3</sup>Moreover, the UG problem has actually increased, since analyses in that model have drifted toward a constant increase in functional categories (to wit, the cartography approach, Cinque, 1994, 1999, 2002; Belletti, 2004; Rizzi, 2004 and nano-syntax Kayne, 2010). Most of these functional categories are redundant system-internal correlates (there are functional categories of SIZE, COLOR, ORIGIN, etc. because there are adjectives of those categories): they add nothing to our understanding of the facts. They are not even discovered correlations but invented correlations, elements added to the theory solely to correspond to some phenomenon (much in the behaviorist way so fiercely criticized in Chomsky's, 1959 review of Skinner; see the discussion in Bouchard, 2001).

In the face of the triple mystery assessment, we might judge that the evolvability of the language-ready brain is too hard a problem and decide to simply drop it. But scientists don't like to give up. If the problem appears insurmountable from the perspective a theory, however widely scholars adhere to it, its apparent incapacity to deal with such core issues as signs, combinatoriality and language-specific conditions in general, can be a motive to scrutinize that theory to figure out why it fails in this respect, and to use this assessment to elaborate an alternative model that can adequately address the core issues. Proponents of UG, and those who share the mystery assessment about language such as Lewontin (1998), all put a high emphasis on the property of discrete infinity found in language, which is assumed to be the core property of the language phenotype: “the core competence for language is a biological capacity shared by all humans and distinguished by the central feature of discrete infinity—the capacity for unbounded composition of various linguistic objects into complex structures” (Hauser et al., 2014, p. 2). This is understandable from a historical background. Generative grammar was born in the context of emerging tools in mathematical logic. For the first time, these tools provided the means to formalize recursion, which had been informally recognized as a property of language for some time (cf. Humboldt's infinite use of finite means). In this context, the most striking characteristic of human language is its discrete infinity. It is tempting to see discrete infinity as an essential property of language, and to put the corresponding technical tools of recursion at the heart of the model. It is then natural to assume that recursion is the crucial distinctive property of human language.

But this core assumption leads to a triple mystery. We should therefore question that assumption. The language phenotype, like all “facts,” is a set of observational propositions which are part of the theory: they are not external to the theory and independent (Lakatos, 1970), and their status can be questioned like any other proposition, particularly in the face of an overwhelming problem such as when a theory leads to a shroud of mysteries. It turns out that the assumption of the centrality of recursion and discrete infinity, though shared by many language scientists, is incorrect. Although it is an observable trait of language, it is not the core phenotype it is assumed to be, but a side effect. The core competence for language is the capacity to take elements from two substances with no logical or natural connection between their elements—perceptual forms and meanings—and to link them into signs (words, morphemes). This capacity to form Saussurean signs is the sole distinctive trait of human language. The fact that only human language has discrete infinity does not imply that recursion is a distinguishing mechanism. This mechanism is uniquely human; however, it is not original: it actually arises from prior elements of the two substances of signs that contain primitive combinatorial processes and produce the effects of recursion<sup>4</sup>.

<sup>4</sup>In this paper, I compare my view with that of Chomsky, since it is the most influential one. There are many other theories of language and its origin, some of which relate to the brain and machinery prior to language. Because of space limitations, I cannot do them justice here, so I refer the interested readers to the extended discussion of other approaches in Bouchard (2013).

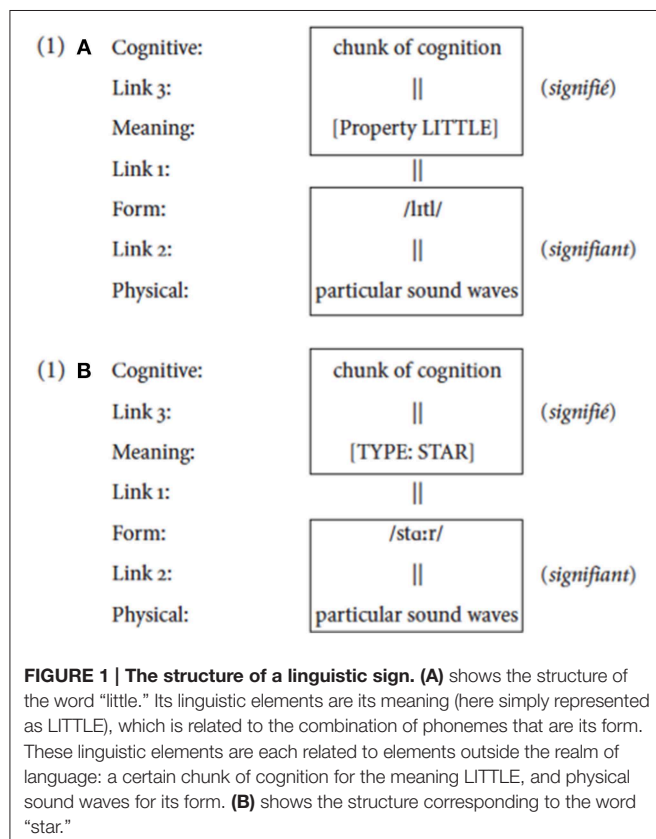


To see this, let us now turn to the detailed properties of linguistic signs.

## The Sign Theory of Language

A linguistic sign is generally presented as involving two elements—a meaning and a form—and a link between the two. Saussure (1916) introduced the terms signified and signifier to emphasize that this linking is purely mental, established by speakers. I use the terms “concept” and “percept” in this spirit: they are dynamical mental creations, cognitive structures (see Jackendoff, 2002, ch. 10). This is an oversimplification, however. A linguistic form (signifier/percept) is a mental state linked to an acoustic/visual material element: this element is not linguistic but in the domain of the sciences that deal with the physical and mental properties of acoustic perception and production (Henceforth, I will only discuss acoustic material of the oral modality, but the ideas carry over to the gestural modality). Similarly, a linguistic meaning (signified/concept) is a mental state linked to a psychological element, a chunk of cognition that the mental state evokes: this element also is not linguistic but in the domain of the sciences that deal with psychological phenomena related to thought. It is only when a language establishes Link 1 between a representation of a perceptual element and a representation of a conceptual element that these are linguistically relevant and become a signifier and a signified.

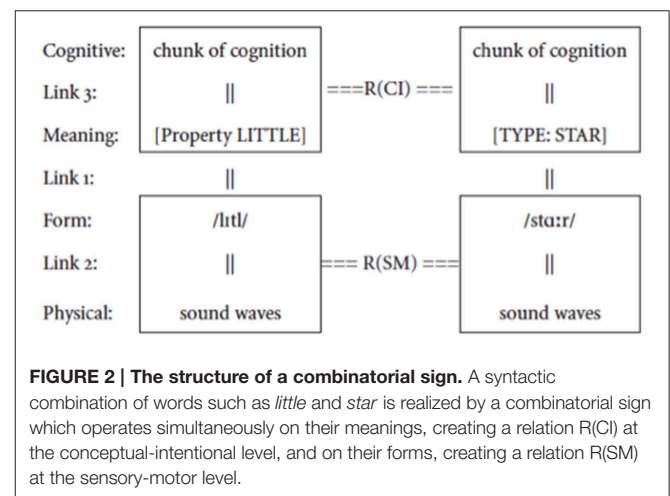
(1) Figure 1



The linguistically crucial part of a sign is a reciprocal predication: it is the systematic attribution of a vocal form and a meaning to each other. The link between signifier and signified is not determined by logic or by intrinsic properties in the nature of the phonic-acoustic or conceptual substances: it is purely linguistic. The properties of the substances to which the signifiers and signifieds are linked cannot explain why a particular phonetic entity is tagged as the signifier of a certain meaning or why a particular conceptual entity is tagged as the signified of a certain form. These links are not due to natural causes, but rather are arbitrary because the nature of the sounds that our phonatory articulators produce and the nature of the concepts that our conceptual system constructs are so different that they cannot entertain a meaningful, logical, or iconic relation (Saussure, 1916, pp. 155–156).

Now consider syntax. If we look at it in terms as neutral as possible, syntax is minimally defined as the processes by which signs are combined. Consider a simple example of the syntactic combination of the two signs *little* and *star*. Each sign is complex by definition—a form resulting from the union of a signified and a signifier. Syntax does not combine just signifiers or just signifieds, it combines relations between signifiers and signifieds, i.e., signs. Since signified and signifier are irreducibly united, any operation applying to one is reflected on the other. So when two signs are combined by a relation R, R operates simultaneously on both their signifieds and their signifiers, as shown in the combination of *little* and *star* in (2).

(2) Figure 2



Since R operates simultaneously on both the signifieds and the signifiers of the signs in (2), it is itself a sign. I will refer to this set of signs that combine syntactic elements as combinatorial signs (C-signs), to distinguish them from the more familiar unit signs (U-signs), namely words/morphemes. This immediately raises two questions: What is the signifier of a C-sign? What is the signified of a C-sign? As already indicated in Bouchard (1996, 2002), the signifier of a C-sign will take whatever form a language arbitrarily selects from among those that our physiology provides as a combinatorial percept in the modality of that language. These

forms are drawn from physical traits of the forms of words. For instance, a first trait in an oral language is that vocal units appear linearly ordered. So signifiers made up of these vocal units can share a temporal edge—they can be temporally juxtaposed: two signifiers can be ordered next to one another, and this can be grammatically significant in the system of a language. For instance, in (3), it is grammatically significant that *saw* and *John* are juxtaposed, but not that *John* and *yesterday* are juxtaposed: the juxtaposition of *yesterday* is grammatically relevant only with respect to the phrase *saw John* (or *Mary saw John* under different assumptions).

(3) Mary saw John yesterday.

The order of juxtaposition is also frequently significant, as in the pairs in (4):

- (4) a. John saw Mary—Mary saw John  
b. John is sick—Is John sick?

A second trait is that the two signifiers can share a temporal space, as when a modulation is superimposed on the phonemes of a constituent: one signifier is the intonation placed on the other signifier, such as an intonation expressing a question (4b). Other possible superimposed elements are stress and length<sup>5</sup>. All these combinatorial percepts depend on the physiological traits of the modality, so they vary across modalities. For instance, the visual-gestural channel of sign languages has more types of combinatorial percepts because it uses more articulators and more dimensions than the auditory-oral channel (Bouchard, 1996).

The set of possible signifiers for a C-sign is extremely restricted because the set of physiological relational vocal percepts is small. So arbitrariness is limited by what are ultimately principles of physical science, as Thompson (1917) anticipated for biological systems in general. Languages vary in their choices of signifiers among these combinatorial percepts, as expected in the light of arbitrariness. For instance, the syntactic relation “direct object” can be expressed by any of these combinatorial signifiers: juxtaposition in the order V-NP or NP-V, a Case affix or a Case tone on the complement, an object affix or an object tone on the verb. This follows from Saussure’s general principle of arbitrariness. There is no “reason of nature” for a language—let alone all languages—to choose any particular combinatorial signifier among those enabled by our physiology: any signifier is a possible candidate, because each one can optimally satisfy the requirement to encode meanings in a form. Indeed, each possibility is instantiated in some language or other. Languages choose from among the various possibilities of combinatorial signs, just as they arbitrarily choose from among the various possibilities of unit signs. Which combinatorial percepts are possible signifiers is not stipulated in some universal list, but is determined by prior properties of the perceptual substance of the modality of the particular language. Under this view, if there was no variation in the way languages express a relation such as

“direct object,” if they all had the same signifier for it, this would be a most improbable accident, just as it would be if the signifier of a unit sign happened to be the same in all languages. Since Saussurean arbitrariness extends to C-signs, variation in syntax is a virtual necessity. Consequently, which particular combinatorial signifier is used in any specific case in a language must be learned just as much as any signifier at the word level. The numerous instances in which each language must choose a C-sign create the impression that languages can be amazingly different. But this is just an impression due to the cumulative effect of the choices; in fact, each choice of C-sign involves only one of the very few percepts that human physiology allows as the signifier of a C-sign. Though each combination is very simple, these combinatorial means cumulatively allow syntax to create organized groups of signs which can attain a very high degree of complexity overall.

Consider now the nature of the meaning of a C-sign, that is, the relation R at the conceptual-intentional level. The signified of R is a relation of predication. Predication, namely the capacity to attribute properties/information to objects, is a universal trait of human cognition. As Hurford (2007a, p. 527) indicates, “In the very earliest mental processes, long antedating language, binary structure can be found, with components that one can associate with the functions of identifying or locating an object and representing some information about it.”

In a combination of signs as in (2), the semantic part of the C-sign links two elements so that one adds its denotation as a restriction on the other, either in the usual sense for subject-predicate and topic-comment relations, as in (5), or in the sense of saturation, as in (6).

- (5) a. John is sick/left early (the property of the VP is attributed to the subject).  
b. that book, I really liked (the property of the comment is attributed to the topic).  
(6) a. liked that book (the property of the direct object is attributed to the V, it saturates the verb).  
b. in the kitchen (the property of the Noun Phrase is attributed to the locative preposition).

In summary, syntax is a set of combinatorial signs that allow the formation of complex signs. The perceptual form of a C-sign can only be either a juxtaposition or a superimposition of a vocal (or gestural) percept; this limitation on the combinatorial signifiers is due to properties of the human sensorimotor systems. The signified of a C-sign is predication, which was exapted from the pre-linguistic cognitive system of humans. Like other signs, combinatorial signs are subject to arbitrariness due to the nature of the two substances that they link. Therefore, which combinatorial signifier a language chooses for any particular predicative relation (i.e., “construction”) is arbitrarily selected from among those permitted by its modality. These are the main tenets of the Sign Theory of Language.

## The Kind of Brain Mechanisms Required for the Formation of Signs

A sign is a link between elements from domains of very different natures—a physical/perceptual form and a

<sup>5</sup>In addition to these very direct ways of indicating that there is a relation between two signs, we can also indicate that a relation is being established between two percepts by physically shaping one in a conventionalized way, in a paradigm that indicates what relation is being established with the other, as with Case marking or agreement.

psychological/conceptual meaning. The core problem is to identify the brain mechanisms that enable links between these two kinds of elements. This neurological property (or set of properties) must be unique to the human brain since only humans have words: no other animal comes close to having equivalent signs detached from the immediate environment and as productively created. In Bouchard (2013), I suggest that these neuronal systems must have properties similar to the uniquely human systems of neurons discussed by Hurley (2008). These systems have the capacity to operate offline for input as well as output: they can be triggered not only by external events stimulating our perceptual systems but also by brain-internal events (including counterfactuals); they can also be activated while inhibiting output to any external (motoric) system. These Offline Brain Systems (OBS) are not specifically designed for language but they provide the crucial trait.

As early as 1891, Saussure understood that the fundamental duality of language is not in the linking of sound and meaning, but “resides in the duality of the vocal phenomenon AS SUCH, and of the vocal phenomenon AS A SIGN—of the physical fact (objective) and of the physical-mental fact (subjective)” (quoted in Bouquet and Engler, 2002, p. 20). The question is in what way, exclusive to humans, the vocal phenomenon enters into the mental domain, into the brain. Non-human animals can correctly classify and appropriately respond to stimuli, so acoustic elements, as well as informative content elements, are already brain-internal, but as categorical systems linked with brain-external elements. Being indexically tied to objects of the world, they are restricted in their mental activations and they cannot freely undergo linkings, they cannot form signs. Something different must be present in human brains. The brain mechanisms we are looking for must enable a vocal sound to be represented in the brain in a way detached from any brain-external phenomenon, as a purely brain-dependent entity, an activation of an OBS or something similar. Consequently, these representations of percepts can be arbitrarily linked to concepts: they can function as signifiers. I refer to these neural systems as Detached Representation systems (DR systems).

In addition to the physical element of a sign becoming a purely mental representation, the informational content of a sign is also different from that of an animal communication system unit. The content of an ACS unit is a category, i.e., a neural linking of similar results from sensory input, a class of input stimuli. This level is still linked to perceptual input—to the outside world. Even the signals that apes learn through intensive training remain at the level of action observation and embodied simulation of action triggered by external events. The content of a linguistic sign is at a more abstract level. It comes from human-specific cognemes that are abstracted from any sensory input or immediacy. This is the level at which detachment is attained. The concepts/meanings of signs do not represent or stand in for outer objects, but are brain activations that take internal events as inputs. This notion of “concept” is similar to the “amodal symbols” of Barsalou (1999) and the “types” of Penn et al. (2008):

“[...] only humans form general categories based on structural rather than perceptual criteria, find analogies between perceptually disparate relations, draw inferences based on the

hierarchical or logical relation between relations, cognize the abstract functional role played by constituents in a relation as distinct from the constituents’ perceptual characteristics, or postulate relations involving unobservable causes such as mental states and hypothetical physical forces. There is not simply a consistent absence of evidence for any of these higher-order relational operations in nonhuman animals; there is compelling evidence of an absence” (Penn et al., 2008, p. 110).

In order to be able to form linguistic signs, humans had to evolve brain systems that enable a more abstract representational level, so that concepts and percepts can be linked. It is not a percept *per se* that is linked with a concept *per se* in a linguistic sign, but a representation of the percept and a representation of the concept, i.e., a mental state corresponding to each of them, as we saw in figure (1). The crucial innovation is in the way some human neuronal systems function. Language did not emerge because there was environmental pressure for better communication or thought organization (though it brought leverage for both). It is not a system with a function of communication that emerged, nor with the function of organizing thought. It is a system of signs that emerged because elements from two very different substances met in the brain via their representations by new neuronal systems.

If the known laws of biology are extrapolated, we expect these brain systems to be in continuity with neuronal systems that are part of the machinery of the pre-linguistic brain, i.e., the brain of a prehuman species that has not yet achieved the capacity for detachment of the sort discussed above. Given biological continuity, it is likely that these are not radically different systems, but rather that they are offline activations of systems involving neurons in essentially the same parts of the brain.

In Bouchard (2013), I conjecture that these systems developed this novel kind of activation due to an increase in synaptic interactions that was triggered by several compounding factors. A large brain with a huge cortex offers a greatly increased potential for synaptic interactions. In addition, the more globular shape of the brain, with the thalamus in the middle, affords more cross-modular interactions (Boeckx, 2012). Moreover, alleles such as ApoE4 significantly improve synaptic repair; hence, they dramatically increase synaptic interactions. In addition, the long dependency during infancy feeds more cultural material into these additional brain capacities. With such a massive increase in synaptic interactions and complexity of circuitry due to biological changes and extensive cultural stimulation, a critical level was reached in hominid brains; some neuronal systems started being triggered by strictly internal brain events, introducing a new form of offline activation with no link to external events related to sensory inputs or motoric outputs. These strictly internal (offline) activations of some micro-anatomical structures represent a small evolutionary step: like the latching discussed by Russo and Treves (2011), they occur without altering the make-up of the neuronal network or any of its constituent properties. But DR systems have gigantic consequences: they enable brain activity of a novel kind and complexity, a unique representational capability that leads to higher level mentalizing.

The dramatic increases in both the number of neurons and the number connections between neuronal networks are instances

where quantity produces quality, the brain activity becoming less input-driven and less rigid. It is not obvious that there is an immediate functional behavioral advantage for an individual to have this kind of detached brain activity. It can slow down reactions to the immediate environment, creating a sort of framing problem. From our current perspective, we see a quality in the innovation; but it may have come only in the long run—part of the pleiotropy of the innovation in brain activation that occurs due to material design, with no teleological push for an improvement of the individual's immediate well-being. Enhancements in the number of neurons and of connections lead to an increase in computational abilities and internal activity, but have little effect on the link between the brain and the perceptual systems interacting with the outer world. This kind of system does not evolve due to functional pressures: it takes on functions after its emergence. As Gould and Lewontin (1979) remark, a trait is not necessarily for something: it can just be a consequence.

This considerable upgrade in the quantity and quality of brain activity is like duplication in genes: other areas/systems can take over (Deacon, 2006), particularly given that the novel functional property of these micro-anatomical structures is less specialized, not tied to particular systems related to perception, but has a general representational capacity. Consequently, the various brain operations related to these systems are expected to exhibit great plasticity, with their anatomical location being diffuse. This is another feature that neuroscientists should be looking for.

Though, it may not be possible to reconstruct the evolutionary history of the causal factors for the brain systems that enable the formation of linguistic signs, we can nevertheless test whether such systems actually exist and whether they exhibit some of the predicted properties, such as plasticity and pleiotropy. There is already evidence in support of the hypothesis.

Concerning the existence of these offline systems, we can see them at work in language once we isolate their effects from those of other activities concurrent with language at the motoric and conceptual levels. For instance, Meister and Iacoboni (2007) report on an experiment in which they compare the processing of visual stimuli while performing an action perception task and two linguistic tasks. They did not find any area specifically activated or with higher activity during the two linguistic tasks: “when visual stimuli concerning object-oriented actions are processed perceptually, they activate a large bilateral fronto-parietal network. When the same stimuli are processed linguistically, they activate only a subset of this network and no additional areas” (p. 6). They argue that these results support “the evolutionary hypothesis that neural mechanisms for language in humans co-opted phylogenetically older fronto-parietal neurons concerned with action perception” (p. 6). The identification of neural systems involved in language, and their role, is extremely difficult. As Dehaene and Cohen (2007) point out, module sharing may involve all levels of brain hierarchic organizations: micro-maps (millimeter-size columns), meso-maps (centimeter-size circuits), and macro-maps (larger-size networks). But with the rapid progress in technology to probe the brain, scientists can refine the testing of linguistic properties relating to neural systems, and eventually put the hypothesis to a test.

Regarding plasticity, Hein and Knight (2008) provide evidence that the same brain region can support different cognitive operations (theory of mind, audiovisual integration, motion processing, speech processing, and face processing) depending on task-dependent network connections (see also Bookheimer, 2002, p. 153). There is no fixed macro-anatomical structure that is exclusively dedicated to language: linguistic processing is a widespread property of the neural networks (Fedor et al., 2009). Language exhibits extensive plasticity for the localization of its components between and within individuals (Neville and Bavelier, 1998), during its development (Karmiloff-Smith, 2006), in its repair (Hagoort, 2009), and depending on its modality (Neville, 1993; Mayberry, 2002). The often-noted association between human praxis and language also points in the same direction. There is a genetic linkage between handedness and language dominance, and clinical correlations between aphasia and apraxia (Donald, 1998).

Regarding pleiotropy, if human brains have systems of neurons that are functionally less specialized, systems that can be activated *in absentia*, triggered by representations of events instead of the events themselves, and produce representations of events with no brain-external realization, then we should find evidence for this capacity in other functional traits unique to humans. There is compelling evidence that several interrelated traits are uniquely human, and absent or in very rudimentary forms in other animals (e.g., Premack, 2004; Penn et al., 2008; Fedor et al., 2009). This Human-specific Adaptive Suite extends across many domains and involves qualitatively huge differences from species that are closely related to us. Here is an indicative list, with a few of the relevant references.

#### *Human-specific cognitive traits*

1. Language: signs and syntactic combinations
2. Imitation (Meltzoff and Moore, 1997; Rizzolatti and Craighero, 2004; Karmiloff-Smith, 2006)
3. Advanced Theory of Mind (Flavell, 1992; Povinelli, 2000)
4. Detachment from immediate situation, episodic memory (of noncurrent scenes and events) (Gärdenfors and Osvath, 2005)
5. Object permanence (Hurford, 2007b)

#### *Human-specific neurological traits*

6. Brain with large amount of neurons and increased connectivity (Russo and Treves, 2011, Deacon, unpublished)
7. ApoE4 (apolipoprotein E4) (provides better synaptic interactions) (Bufill and Carbonell, 2004) and other proteins with effects on language (Fitch et al., 2010)
8. Plasticity of the brain for several functions (Fedor et al., 2009; Hagoort, 2009)
9. Offline Brain Systems (offline activations, inhibiting input or output) (Hurley, 2008)

The human-specific cognitive and neurological traits are so closely linked that several scholars assume that at least a good part of them coevolved synergistically from a common factor underlying these various cognitive modules (Szathmáry, 2008; Fedor et al., 2009). Some assume that the underlying supermodule is one of the functional modules, the two most



popular being Theory of Mind and language. However, Penn et al. (2008) argue compellingly that the suite of discontinuities between human and non-human minds cannot be explained by relating an *explanans* directly to the functioning of these cognitive domains. (See Bouchard, 2013, pp. 113–114) for arguments against the language-first and ToM-first hypotheses). The Human-specific Adaptive Suite provides initial evidence for a neurobiological innovation with general representational potential. Given the limitations in the current techniques available, the specifics of many of these traits are still unclear, but they may ultimately help us resolve the problem of the neural basis of sign formation.

Though this is difficult, the hypothesis can nevertheless be tested. One useful line of inquiry can be found in recent experiments by Stanislas Dehaene and Laurent Cohen (Dehaene, 2005; Dehaene and Cohen, 2007). They show that some adaptations can occur much faster than is expected on a genetic scale, due to a process that they call “neuronal recycling” that operates during cultural acquisitions such as reading and arithmetic. They observe that part of the human cortex is specialized for these two cultural domains. Since invention of these cultural activities is too recent to have influenced the evolution of our species, they hypothesize that this specialization results from neuronal recycling: reading and writing are not genetically encoded, but they nevertheless find their niche in a well-suited set of neural circuits.

Note that under the hypothesis that the novel brain systems coincidentally allowed mental states corresponding to elements of the perceptual and conceptual substances to meet in our brains to form linguistic signs, this does not raise what Chomsky (2005, 2007, 2010, 2011) refers to as the Jacob-Luria problem. Though he accepts that pressures to communicate may have played a role in the gradual fine-tuning of language, Chomsky has repeatedly claimed that, at its origin, language could not have evolved due to communicative pressures because this raises a problem:

(7) *Luria/Jacob problem*: How can a mutation that brings about a better communication system provide any survival advantage to the first single individual who gets it?

A mutation occurs in a single individual, whereas communication takes place between individuals<sup>6</sup>. Under my hypothesis, the offline systems with general representational capacity took on this other function of linking percepts and concepts after they were in place due to a suite of evolutionary pressures. The Luria/Jacob problem does not arise in this approach because the change was not for language or any of its functions like communication or organizing thought. The change produced offline systems. Linguistic signs are a side effect of this neurobiological property. Even if it depended on a mutation (but I doubt this to be the case as indicated above), the new trait could spread in a population because it has evolvability of its own, and all the members of that group are then brain-ready for the innovative side effect when it occurs:

by the time words come around, they can be understood by conspecifics.

The advent of some kind of DR system is the crucial small change that made a big difference. This provided the core biological mechanism of the language phenotype—the capacity to link percepts and concepts into signs. Given this capacity and the prior properties of the two substances of the elements linked by signs, the rest of the linguistic properties follow without the need of any additional language-specific rules or conditions. In the next sections, I show how this happens in the three core components of grammar: phonology, semantics, and syntax.

Before turning to these issues, an important question remains to be addressed. Once humans had the capacity to form a limitless number of signs, they developed a capacity to learn and remember a vast set of such signs. How exactly this additional capacity depends on the mechanisms of the first capacity is a question that can now be asked, given my hypothesis. If I am correct in supposing that the DR system is likely to be (part of) what provides humans with a more advanced Theory of Mind (such as a shared attention mechanism and a meta-representation of others' mental states, Baron-Cohen, 1995), and if it also turns out to be correct that word learning strongly depends on an advanced ToM (Bloom, 2000), then the DR system would be crucial for both the capacity to form signs and the capacity to learn and remember them. See the discussion in section 9.5 of Bouchard (2013) and references therein.

## Contrastive Dispersion of Percepts and Combinatorial Phonology

As is the case in other biological systems, DR systems are complemented by epigenetic self-organizing constraints that emerge from interactions among properties of building materials that limit adaptive scope and channel evolutionary patterns (Jacob, 1982; Erwin, 2003). Since the linguistic linking between a percept and a concept is arbitrary—that is, it is not hard-wired but made possible by their representations in DR systems—the representation of any percept can potentially be linked to the representation of any concept, and the links can change very rapidly. So there are innumerable possible links. This is compounded by the fact that there are infinitely many incrementally different vocal forms that we can produce and perceive, and an untold number of possible concepts/signifieds because DR systems introduce a detachment from the immediate situation that opens the door to any imaginable situation, presented from a multitude of perspectives. Moreover, there is the logical possibility that individuals will choose different linkings: in the extreme case, each individual would have its own system. Therefore, DR systems introduce an unprecedented sort of chaotic system in the brain. This creates randomness that is confronted with material constraints. As in other situations far from equilibrium, small chance disturbances are progressively amplified by material properties and result in clusterings, in order out of chaos (Prigogine and Stengers, 1984).

In this kind of self-organization, local interactions of components of a system generate complex organized structures

<sup>6</sup>Pinker (1994: 365), foreseeing the objection, counters that the initial grammar mutant could talk to the 50% of brothers and sisters, and sons and daughters, who shared the new gene.

on a global scale. In language, the potential chaotic dispersions of arbitrary signs are constrained by the physical and cognitive properties with which the signs are confronted. These constraints restrict the linguistic sign system in a way that maximizes contrastive dispersion and creates clusterings that result in the various properties of language that we observe in phonology, semantics, and morphosyntax.

## Phonological Segments

The signifier/percept of a sign is the part most noticeably influenced by material properties. Though the representation of any percept at all could in theory become a signifier, the possibilities of the chaotic system are considerably narrowed by material properties of our production systems and perception systems.

A salient property of human vocalizations is that they are perceived as segments: discrete elements. This is a general design feature of human neurophysiology: information that unfolds over time is chunked in the acoustic domain, as well as in other domains such as vision. This is a bilateral stimulus-neutral system of temporal segmentation that operates before feeding specialized lateralized systems such as the processing of speech or music (Poeppel, 2001). Sensory input is analyzed on different timescales by the two hemispheres. High-pass (global) information from short 20- to 50-ms temporal integration windows is passed to left hemisphere areas, whereas low-pass (local) information from long 150- to 250-ms integration windows is passed to the right hemisphere (Poeppel, 2001) (However, the issue is still unclear and recent work shows that lateralization in this case may be weak; Giraud and Poeppel, 2012). These oscillations arise naturally in our perception of vocalizations (Poeppel, 2003; Sanders and Poeppel, 2007), and the temporal integration of vocalizations is reflected as oscillatory neuronal activity. The timings correspond to typical segments and syllables.

Similar, bilateral segmentation systems appear to be shared by other species; they are the basis of the auditory processing of species-specific vocalizations in macaque monkeys, and the ability of squirrel monkeys to discriminate between conspecific and non-conspecific vocalizations (according to studies reported in Poeppel, 2001). This timing ability is the basis of a system with an important adaptive benefit: a strong change in rhythm signals danger. In sum, we perceive sound as segments, in a digital, not analog way. Segments are perceived as being produced concatenated. An important question is what determines the particular repertoire of possible phonemes. Why do digitized vocal percepts cluster in a few particular hot spots among the innumerable, chaotic possibilities we can produce and perceive? As in other chaotic systems, the clusterings depend on frequency and accumulation: chance vocalizations are progressively amplified by material properties pertaining to ease of production and distinctness of perception. On the production side, vocalizations involve the displacement of organs, hence muscular energy. Certain vocalizations are easier to pronounce and require less energy; this is likely to favor their use and increase their frequency (Lindblom, 1992).

The human perceptual systems also set upper bounds on the distinctions that we can perceive or produce as signifiers.

Distinctness of expression is particularly important in the case of acoustic information since it is only physically available for a very short length of time and cannot be recovered in the case of an erroneous perception. Nowak et al. (2002) found that the demands of discriminability (as well as memory and time to learn) constrain the system to a fairly small set of signals, an observation already made by Wang (1976, p. 61). The actual repertoire is very small: a few dozen discrete perceptual elements. This observation extends to sign languages that use the gestural modality: there are very few gestural minimal elements, and like phonemes, they are made up of articulatory features (see, for instance, Brentari, 2002). This small set of percepts is a result of self-organization. Vocalizations that are easier to produce and can be more distinctly perceived have a higher frequency of use. As frequencies increase, accumulations occur at certain points in the articulatory-acoustic continuum. Percepts cluster in particular hot spots as a result of this contrastive dispersion. As Lindblom (1992) (following Liljencrants and Lindblom, 1972) indicates, a compromise between perceptual distinctiveness and articulatory cost brings about quasi-optimal perceptual distinctiveness. But this is not sufficient, because the search space is too large for convergence on a structure as complex as the human phonological system. However, if we take into account the properties of building materials, self-organization derives the phonemic clusters. Thus, Carré and Mrayati (1990) and Oudeyer (2005, 2006, 2007) show that canalization by the vocal tract and general acoustic theory define “eight discrete regions of such a tube where deformations, or constrictions, afford greatest acoustic contrast for least articulator effort” (Studdert-Kennedy, 2005, p. 64), and these correspond to places of articulation in natural languages. Thus, vocalic systems most frequently have peripheral vowels, which are the most contrasted (Ménard, 2013).

## Phonological Combinations

This severe limitation on the number of usable percepts is the source of the clash between the possibilities of the perceptual and conceptual substances. There are innumerable meanings and ways to partition meaning (more on this below), but discriminable speech sounds are limited by the material properties of sound production and perception. The combinatorial formation of signifiers is usually attributed to this clash between the possibilities of the two systems. “If the symbols were holistic vocalizations like primate calls, even a thousand symbols would be impossible to keep distinct in perception and memory” (Jackendoff, 2002, p. 242). In simulations like Oudeyer's, the small number of clusters “automatically brings it about that targets are systematically re-used to build the complex sounds that agents produce: their vocalizations are now compositional” (Oudeyer, 2005, p. 444). How could that be? Where do the compositional processes come from? The answer is again found in the material properties already present in the forms. Vocal units have the following universal material properties:

- they occur in time, so they can be ordered and juxtaposed;
- they can have various intonations;

- \_ they can be shortened or lengthened;
- \_ they can be stressed or unstressed.

These acoustic and auditory properties are also distinguishing elements in the signals of other mammals (Lieberman, 1968).

Vocalizations occur in time, and the material properties of vocal articulators are such that we cannot produce more than one vocal unit at a time. This is a contingent property of language production. Since vocal units are aligned in time, our perceptual system captures the linear properties of vocalizations when they are produced, in particular the linear relationship between two vocal units, the most salient one being linear adjacency. The linear adjacency of two vocal percepts is itself a percept and can be represented by a DR system, like any other percept. The relational percept of juxtaposition is already in the stock of our perceptual system; hence, it is available for DR systems that link concepts and percepts. Another material property of vocalizations is intonation; therefore, another perceptual element represented by DR systems is the tone superimposed on a vocal unit, of which there are a few distinctive values due to contrastive dispersion. Similarly, the length and stress of a segment are percepts that can be represented by DR systems, within the limits of distinctive values. Crucially, in an arbitrary system, the percept represented by a DR system and linked to a concept can be any element among those recognized by the perceptual system: a vocal unit, a juxtaposition of units, an intonation, a length, or a stress. Because the system is arbitrary, it makes no difference whether the represented element is simple or complex. The acoustic image can be a single phoneme or the relational percept of juxtaposition applying any number of times to phonemes, as well as any of the available distinctive intonations, lengths, and stresses on these elements. These complex elements remain within the limits of what humans can distinctively perceive or produce because their parts have the appropriate qualities. Phonological combinatoriality comes from a material property of the articulatory and perceptual systems, namely the fact that vocalizations are temporally linearized, which entails the percept of juxtaposition. The phonetic data provide information on the source of more abstract principles: segmenting into phonemes, as well into as words/morphemes, already contains computational properties (see DeWitt and Rauschecker, 2011 for combinatorial properties in basic perception). This simple concatenation-recursion of phonemes allows an unlimited derivation of signifiers: any combination of distinguishable percepts can be a signifier. This system is subject to a general law of nature whereby the frequency of an element is inversely correlated with its complexity: the simpler an element is, the more likely it is to appear in nature (cf. Zipf, 1965/1949). Though concatenation-recursion of phonemes can derive infinitely complex signifiers, the simpler ones are much more likely to be formed, produced, or heard. This higher frequency creates accumulations that make the system relatively conservative in terms of the number and complexity of elements that form its signifiers. In addition, production ease and auditory salience influence not only the dispersion of vowels and consonants, but also syllabic templates, or sequences of segments: the combinations of phonemes are subject to phonotactic constraints, such as the energy expended

for the transition, which also constrain the nature and number of potential signifiers. The constraints that arise from properties of the articulators and ease of articulation influence what phonemes occur in adjacent positions as early as babbling (MacNeilage and Davis, 2000). The overall complexity of a signifier is also likely to be limited by memory and retrieval capacities.

Discrete speech sounds and their combinations emerged because they are consequences of material laws that apply to a certain kind of organism hosting DR systems that can represent elements of their perceptual and conceptual systems and links between them. The chaotic system deriving from these brain systems must have the properties that we observe because the building materials channel the way the system becomes structured into specific self-organizations.

## Contrastive Dispersion of Meanings and Combinatorial Semantics

Segmentation is also a design feature of the human cognitive make-up. We digitize the world and events into discrete chunks, action packages varying from 0.3 to 12 s, mostly 1 to 4 s long (Schleidt and Kien, 1997). As for the ontology of the cognitive units, our perceptual attention systems treat the world as containing two basic kinds of entities (Hurford, 2007a, p. 527):

1. objects ("something is there");
2. properties of these objects ("what is there" "what is happening to it").

Another aspect of cognitive segmentation is found in the two types of attention discussed by Humphreys (1998). Global attention captures the gist of the whole scene. In language, this corresponds to something like the main predicate and its arguments. Local attention is subsequent focal attention on local features of individual objects. In language, this corresponds to secondary predicates such as nouns, adjectives, etc.

By allowing detachment, DR systems introduce a chaotic expansion on the meaning side of language: there is an extremely large if not infinite number of potential (offline) concepts. First, the vast number of objects and situations we perceive can all be represented offline as concepts, as well as their properties. This is compounded by the various perspectives we can have on them (Quine, 1960). Moreover, the potential for concept formation is multiplied by the affordance of intra-brain interactions where some neuronal systems are triggered by other brain events. In addition, a particular language can partition the conceptual substance in countless possible ways to delimit its lexical meanings (Saussure's radical arbitrariness). In a system of arbitrary signs, any of these elements treated by the cognitive system could be a meaning represented by DR systems and could be linked to a vocal form.

But this unbridled expansion in meanings is constrained by design features of our cognition. For instance, our global

attention process is constrained as to the number of participants that it can take in at a glance: we can subitize at most four salient objects at a time (cf. the “magical number 4” in Hurford, 2007b). Though actual events can involve any number of actants, the chaos of what we observe is organized by subitizing and chunks of four or fewer actants. The recurrence of the perception of these chunks in the environment creates accumulations, and language has settled on predicates with at most four arguments. The chaotic expansion that could potentially arise from linguistic arbitrariness in the meanings of words is also limited in a more general way by material properties and self-organization. Here, too, order arises out of chaos and clusters are formed in the mass of the conceptual substance as a result of frequency and accumulation. In this case, accumulation depends on the material conditions that make the situations denoted by the concept relevant for the organisms. The more a situation has some importance and/or is encountered frequently by the organism, the more frequently concepts associated with it will be activated. The accumulations self-organize around the concepts most used by the organism. It is this usefulness that makes the meanings tend to correspond to fairly broad and/or usual categories of things, actions, qualities, etc. (an observation already found in Locke, 1690/1964, p. 15). Similarly, Nowak et al. (2002, p. 2131) note that “[t]he evolutionary optimum is achieved by using only a small number of signals to describe a few valuable concepts.”

Usefulness is also the motivation for the important role played by basic level concepts (Rosch and Mervis, 1975; Rosch et al., 1976). Murphy and Lassaline (1997) argue that the basic level is an optimal compromise between informativeness and distinctiveness: this level is informative, because we can infer many unobserved properties once we know which basic category something belongs to, and distinctive because it is a relatively easy categorization to make. Thus, if you ask someone What are you sitting on?, you are more likely to get the answer chair rather than a subordinate such as kitchen chair or a superordinate such as furniture. Names for basic-level concepts are among the first common nouns learned by children (Brown, 1958).

In fact, we can construct so many particular objects and events and their properties out of reality, potentially an infinite number, that it would not be useful (in a general as well as in an evolutionary sense) since most of them recur only very rarely, if at all. This is likely why meanings tend to converge on these hot spots of accumulation.

Even with an important number of U-signs and the possibility of combining them by means of C-signs, the resulting meanings are nevertheless generally quite broad and may correspond to several different situations in the world, including the meanings of sentences. Trying to remedy this underdetermination would force language into ever more complex constructions, to a point where it would be extremely unwieldy. Humans have another prior mental attribute that avoids this problem and favors the cumulative use of broad concepts: a system of pragmatic inferences that derives from a full Theory of Mind (ToM). Given the pragmatic inferences that derive from ToM and the context of utterances, expressions need not have fully determined meanings in order to convey information that is sufficiently precise to

be of current use. When two human beings interact, they each have a full ToM, similar cognitive and perceptual systems, and similar contextual information. Therefore, they both know that they have an enormous amount of information in common, and their language faculty does not operate in a vacuum. Using and understanding language involves intensive reliance by speakers on their shared conceptual and contextual knowledge. Pragmatic theories from authors as diverse as Ducrot (1984), Grice (1975), Levinson (2000), and Sperber and Wilson (1986) all share this observation that comprehension is inferential and it draws on both sentence meaning and context (in a very broad sense). Since the inferential system is independently grounded, languages do not drift into an unbridled multiplication of meanings redundant with contextual information, but converge on broad, sufficiently informative meanings (Bouchard, 1995; Hoefler, 2009). A similar argument can be made from the perspective of language's other main function, i.e., thought organization.

To sum up, discrete meanings are clusters formed in the mass of the conceptual substance as a result of maximizing contrastive dispersion across the space for signifieds under the effects of frequency and accumulation due to relevance/usefulness. These clusters are relatively few in number and signs tend to have fairly broad meanings. This does not adversely affect the communicative or thinking functions of language because linguistic signs reside in organisms that independently have an inferential system that supplies the required complementary information.

## Syntax

### The Source of Syntax

Syntactic combination of words and phrases raises the same question as phonological concatenation. Where do the combinatorial tools come from?

If we try to determine what brain systems enable the formal properties of syntactic combinations and the plausibility of these systems given known laws of evolution, it is likely that we will not get very far, because formal systems are only very remotely related to factors involved in evolutionary changes. The system that forms signs (lexicon) and the system that combines signs (syntax) have properties that are so different in current models that they seem quite disconnected. For instance, Chomsky (1995, p. 8) says that matters concerning “the sound–meaning pairing for the substantive part of the lexicon [...] appear to be of limited relevance to the computational properties of language.” But that is not so in the approach I adopt. If we look at the physiological and cognitive properties of the elements being combined, a hypothesis emerges with means and a method of confirmation that are clear enough to be verifiable. Since I argue that the syntax of a language is a set of particular combinatorial signs, each with its signified and signifier, I change the ontology of syntax from a formal computational system to a set of neurophysiological elements.

Syntactic compositional processes, i.e., C-signs, are simply functional uses of universal pre-existing properties of vocal sounds and universal pre-existing properties of our cognitive system. Combinatorial syntax is due to the self-organization of



these prior vocal and cognitive elements. On the conceptual side, the most frequently represented element is the relation of predication, since it is common to all the attributions of properties. This is compounded by the fact that human brains with DR systems have extended this cognitive process: DR systems can not only attribute properties from sensory inputs to perceived objects but, by operating offline, they can also attribute abstract conceptual properties, not linked to immediate sensory inputs. Predication is the broad meaning par excellence. It is a relation that is broad enough to apply to almost all possible meanings and it is omnipresent in our cognitive system. So it is the meaning that creates by far the strongest concentration point in the chaos of semantic DR systems. The fact that our linguistic system has integrated the predicative function at its core simply reflects the place of this readily exaptable concept in our cognitive system, its high rate of frequency and accumulation.

On the perceptual-physical side, words being made of concatenated phonemes, i.e., of elements with properties of vocal sounds, the most frequent elements are temporal sequencing, and superimposition such as intonation, stress, and length. These traits are always present, so they are by far the most frequent elements in the vocal perceptual system. Thus, the hottest accumulation point in the mass of the conceptual substance is the relational concept of predication, and the hottest accumulation points in the mass of the perceptual substance are the two relational percepts of juxtaposition and superimposition. These accumulation points are so overwhelmingly dominant in their respective domains that they increase the frequency of links involving them to the point where these links inevitably accumulate and crystallize. It follows that when human organisms develop signs due to properties of their prior DR systems, they inescapably develop combinatorial signs involving predication as a meaning and juxtaposition and/or one of the forms of superimposition as a signifier. In short, syntax is a consequence of self-organization arising out of the chaos created by DR systems, as is the linguistic sign.

Syntactic combination arises from prior properties of the conceptual and perceptual substances involved, given general laws of nature concerning highly complex systems, à la Prigogine and Zipf. These cognitive and material design properties have a very strong canalizing effect. In particular, they are all primitive combinatorial processes: predication combines an object and its property; order and juxtaposition hold of two segments; intonation, length, and stress apply to segments. As a result, the sign itself introduces combinatorial systems into the linguistic system, and from these primitive combinatorial systems derive concatenation in phonology and combination in syntax. The logically prior properties of the physical and conceptual components of signs are the source of key design features of language, including the particular type of combinatorial system that it has. Syntax happens to have functional effects that are useful for communication and thought, but they are not the factors that triggered its emergence; they are just fortunate consequences.

## Type-recursion

In addition to concatenation-recursion, as found in phonology, the syntax of human language exhibits a particular kind of recursion, where an element of type X can be embedded within other X elements indefinitely. I refer to this as type-recursion. We want to know not only why language has recursion, but also why it has type-recursion.

Type-recursion involves more than recognizing nested attributes of objects (an ability that some animals have) (Penn et al., 2008, p. 117). To have type-recursion, you need an additional property: the complex signs must have a label; they must belong to a category. If a phrase did not have a labeled category, it could not contain another phrase of the same category.

Since properties of signifiers are essentially those of phonological elements, the types cannot come from these. The source of the typological distinction must be in the signified/meaning. Whether these categories are determined ontologically or functionally is an important question that has been debated for centuries. I will not address it here since it is tangential to the issue. I will simply assume the broad hypothesis that lexical items have categories when they interact in syntax. It is also broadly assumed that the phrasal categories are identical to the lexical categories (Noun, Verb, Adjective, Preposition, Tense, etc.). This is due to the fact that syntactic phrases are endocentric: the category of a phrase always comes from one distinctive component, which we refer to as the head—ultimately a lexical head, a U-sign.

The syntactic properties of headedness and endocentricity derive from prior properties. First, asymmetries in syntactic relations, such as the asymmetry between heads and dependents, come from the fact that predication, the meaning of C-signs, is asymmetrical (Venneman, 1974; Keenan, 1978; Bouchard, 2002): the property expressed by the dependent is attributed to the head. Second, endocentricity derives from the way we cognitively process property-attribution (predication): in our cognitive perception of the world, an object to which we attribute a property remains an element of the same type; in a way, it remains the same object. In language, this means that a noun to which we add an adjective remains a nominal thing; a verb to which we add an argument remains a verbal thing, etc. There is a kind of hyperonymic relation between the head and the phrase (cf. Bauer, 1990; Croft, 1996). Assuming the parsimonious hypothesis that the only syntactic primitives are lexical and combinatorial signs, we derive the Endocentricity Theorem:

### (8) Endocentricity Theorem

The category of a constituent X is the category of the element that receives a property by the predication of the C-sign that formed X.

In other words, if X is formed by a C-sign that assigns the property of A to B, then X's category is the category of B (B being the "object" that receives the property of A).

We now see why language has type-recursion: type-recursion occurs whenever a restraining sign or one of its elements happens to be of the same type as the restrained sign whose category projects and determines the category of the complex sign. Type-recursion is a side effect of the combinatorial properties of the substances of signs, interacting with a general cognitive principle of property attribution.

Combinatorial syntax is not a hard-wired property that has evolved at some time. This ability ultimately derives from the particular representational capacity of DR systems that allows the formation of signs. Discrete infinity is a side effect of limitations on chaotic systems like arbitrary language, in interaction with material properties of the sensory-motor and conceptual substances. Both concatenation-recursion and type-recursion derive from the resulting self-organization that takes place. The reason that other animals do not have anything like combinatorial syntax in their communication systems is that they do not have DR systems, which also explains why they do not have unit-signs/words. The crucial leap for a language-ready brain was the development of DR systems that enable the linking of elements of two substances with no logical or natural connection between their elements, so that the linking is purely symbolic. The sole distinctive trait of human language is the capacity to form Saussurean signs. Recursion and discrete infinity are just side effects of this trait.

The ontology of syntax is not a formal computational system, but a set of neurophysiological elements. These elements have high evolvability, in contrast with formal systems. Interestingly, it is by attributing a non-central place to recursion that we can explain how language became type-recursive. My account is in the spirit of Evo-Devo proposals: type-recursion is not due to a specific genetic change but to logically prior properties of the building materials of language.

The Sign Theory of Language has high evolvability with respect to signs and combinatoriality. But of course, a linguistic theory must also pass the test of accounting for the collection of properties that linguists have uncovered about language. I am fully aware that if I am to make the radical claim that syntax is just a small set of C-signs determined by the nature of the sensory-motor and conceptual substances, I must show how that proposal can account for the numerous claims made about the syntax of human languages over the years. Space limitations prevent me from doing that here. But the linguistically inclined reader will find a long discussion in Part IV of Bouchard (2013) that tackles a representative sample of some of the constructions that have been most influential for theoretical argumentation over several decades:

- subject-auxiliary inversion and structure dependence;
- binding conditions on referential relations;
- existential *there*-constructions;
- subject raising constructions;
- long distance dependencies and bounding conditions.

In addition, Bouchard (2002) analyses the distribution and interpretation of adjectives in French and English in exquisite detail, as well as bare noun phrases and bare determiners (clitics).

In all these cases, the unification proposed in STL leads to new insights that allow us to progress in our understanding of language. Many properties that we know about make sense in this model, whereas they just existed, were described but left unaccounted for, in classic models.

## Conclusion

We must understand precisely what language is and have well defined linguistic phenotypes to search for the neural substrates that enable these phenotypes.

From a linguistic perspective, there are strong reasons to assume that the central trait of human language is the capacity to form signs by linking perceptual forms and meanings, rather than the currently prominent view that puts computational tools with discrete infinity at the core of language. Recursive syntax turns out to be a side effect of sign formation: due to general principles, an organism that develops signs inevitably develops combinatorial signs. The Sign Theory of Language offers a comprehensive and unifying approach to the functioning of the main subsystems of language: by calling on the perceptual and conceptual substances of signs and the self-organization that it triggers, the theory explains specific properties of signs, as well as the basic structuring of language in its phonology, semantics, and syntax.

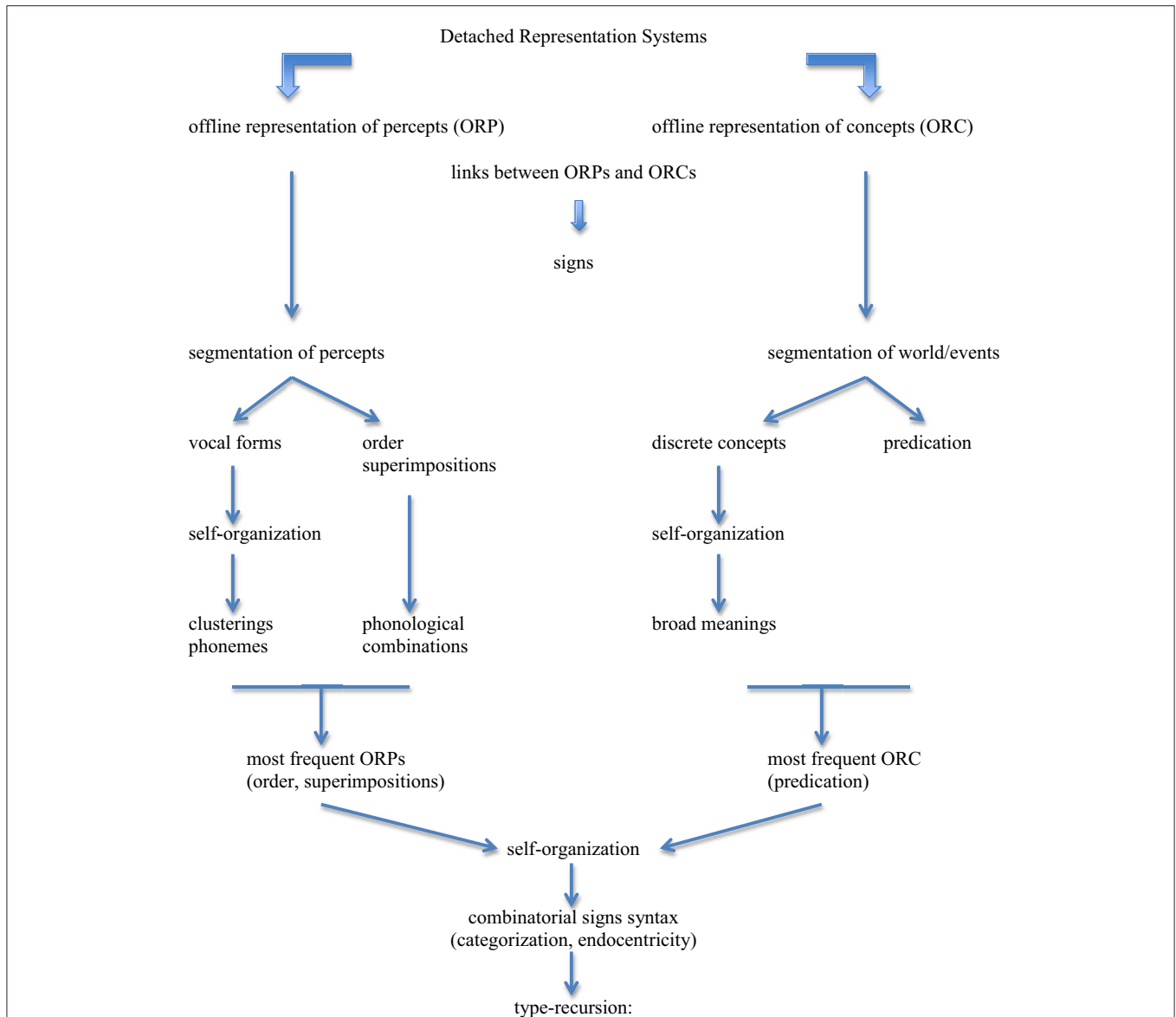
This change of perspective regarding linguistic phenotypes suggests to direct research on neural substrates that enable meta-representational functionalities, detached from sensory input and motoric output. Such Detached Representational systems need not be language specific. Given biological continuity, it is likely that the neural mechanisms will have broad effects at the functional level. While looking for the neural substrates that enable the formation of linguistic signs, it may therefore be useful to consider possible effects of these substrates on non-linguistic traits that may also depend on DR systems, such as the traits discussed under the Human-specific Adaptive Suite.

Another important question in probing the language-ready brain in adults is how much of the mechanisms are in place at birth and how much of the language system takes form during infant development.

My hope is that the change in perspective that we are led to from purely linguistic considerations will enlighten the search for the neural substrates of language, so we will eventually have a better understanding of what makes, and made, the human brain language-ready.

The general outline of the model is as follows:

### (9) Figure 3



**FIGURE 3 | General outline of the model.** Detached Representation Systems produce offline representations of percepts (ORP) and offline representations of concepts (ORC). Links between ORPs and ORCs create linguistic signs. These signs have restricted traits because the prior properties of the substances of percepts and concepts severely constrain them. The segmentation of percepts produces a potentially infinite set of vocal forms; these forms undergo self-organization that creates clusterings that result from frequency and accumulation due to ease of production and distinctness of perception: this delimits the set of potential phonemes for languages. The segmentation of percepts also introduces order and superimposition (intonation, length, stress), which derive phonological combinations. The segmentation of world/events produces a potentially infinite set of discrete concepts; these concepts undergo self-organization that creates clusterings that result from frequency and accumulation due to relevance/usefulness for the organism: this delimits the set of broad meanings for languages. Concepts fall in two broad classes, objects and properties, the latter introducing the notion of predication in a broad sense. The most frequent ORP elements are order and superimposition, and the most frequent ORC element is predication. These accumulation points are so overwhelmingly dominant in their respective domains that they increase the frequency of links involving them to the point where these links inevitably develop into combinatorial signs (syntax). Given categorization and endocentricity (due to object permanence), the syntax of languages has the formal property of type-recursion.

## References

- Baron-Cohen, S. (1995). *Mindblindness: An Essay on Autism and Theory of Mind*. Cambridge, MA: MIT Press.
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behav. Brain Sci.* 22, 577–660. doi: 10.1017/s0140525x99002149
- Bauer, L. (1990). Be-heading the word. *J. Linguist.* 26, 1–31. doi: 10.1017/S00222226700014407
- Belletti, A. (2004). *Structures and Beyond: The Cartography of Syntactic Structures (Oxford Studies in Comparative Syntax)*. Oxford: Oxford University Press.
- Bloom, P. (2000). *How Children Learn the Meanings of Words*. Cambridge, MA: MIT Press.

- Boeckx, C. (2012). "Homo combinans," in *Paper Presented at the Evolution of Language Conference 9 (EVLANG9)* (Kyoto).
- Bookheimer, S. (2002). Functional MRI of language: new approaches to understanding the cortical organization of semantic processing. *Ann. Rev. Neurosci.* 25, 151–188. doi: 10.1146/annurev.neuro.25.112701.142946
- Bouchard, D. (1995). *The Semantics of Syntax*. Chicago, IL: University of Chicago Press.
- Bouchard, D. (1996). Sign languages and language universals: the status of order and position in grammar. *Sign. Langu. Stud.* 91, 101–160. doi: 10.1353/sls.1996.0018
- Bouchard, D. (2001). Review of Richard S. Kayne 2000, *Parameters and Universals*. *The Linguist Network*, #12.1602. Available online at: <http://linguistlist.org/>
- Bouchard, D. (2002). *Adjectives, Number and Interfaces: Why Languages Vary*. Oxford: Elsevier Science.
- Bouchard, D. (2013). *The Nature and Origin of Language*. Oxford: Oxford University Press.
- Bouquet, S., and Engler, R. (2002). *Écrits de Linguistique Générale par Ferdinand de Saussure*. Paris: Gallimard.
- Brentari, D. (2002). "Modality differences in sign language phonology and morphophonemics," in *Modality and Structure in Signed and Spoken Language*, eds R. Meier, K. Cormier, and D. Quinto-Pozos (Cambridge: Cambridge University Press), 35–64.
- Brown, R. (1958). Linguistic determinism and the parts of speech. *J. Abnorm. Soc. Psychol.* 55, 1–5. doi: 10.1037/h0041199
- Buflin, E., and Carbonell, E. (2004). Are symbolic behavior and neuroplasticity an example of gene-culture coevolution? *Rev. Neurol.* 39, 48–55.
- Carré, R., and Mrayati, M. (1990). "Articulatory-acoustic-phonetic relations and modelling, regions, and modes," in *Speech Production and Speech Modelling*, eds A. Marchal and J. Hardcastle (Dordrecht: Kluwer Academic Publishers), 211–240.
- Chomsky, N. (1959). Review of *Verbal Behavior* by B. F. Skinner. *Language* 35, 26–58. doi: 10.2307/411334
- Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, N. (2005). Three factors in language design. *Linguist. Inquiry* 36, 1–22. doi: 10.1162/0024389052993655
- Chomsky, N. (2007). Of minds and language. *Biolinguistics* 1, 9–27.
- Chomsky, N. (2010). "Poverty of stimulus: unfinished business," in *Transcription of Oral Presentation* (Mainz: Johannes-Gutenberg University Mainz).
- Chomsky, N. (2011). Language and other cognitive systems. What is special about language? *Langu. Learn. Develop.* 7, 263–278. doi: 10.1080/15475441.2011.584041
- Cinque, G. (1994). "On the evidence for partial N movement in the Romance DP," in *Paths Towards Universal Grammar*, eds G. Cinque, J. Koster, J.-Y. Pollock, L. Rizzi, and R. Zanuttini (Georgetown: Georgetown University Press), 85–110.
- Cinque, G. (1999). *Adverbs and Functional Heads*. Oxford: Oxford University Press.
- Cinque, G. (2002). *The Structure of DP and IP: The Cartography of Syntactic Structures (Oxford Studies in Comparative Syntax)*. Oxford: Oxford University Press.
- Croft, W. (1996). "What's a head?" in *Phrase Structure and the Lexicon*, eds J. Rooryck and L. Zaring (Dordrecht: Kluwer Academic Publishers), 35–75.
- Deacon, T. W. (2006). "Evolution of language systems in the human brain," in *Evolution of Nervous Systems. The Evolution of Primate Nervous Systems*, ed J. Kaas (New York, NY: Academic Press), 1–26.
- Dehaene, S. (2005). "Evolution of human cortical circuits for reading and arithmetic: the 'neuronal recycling' hypothesis," in *From Monkey Brain to Human Brain*, eds S. Dehaene, J. R. Duhamel, M. D. Hauser, and G. Rizzolatti (Cambridge, MA: MIT Press), 133–157.
- Dehaene, S., and Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron* 56, 384–398. doi: 10.1016/j.neuron.2007.10.004
- DeWitt, I., and Rauschecker, J. P. (2011). Phoneme and word recognition in the auditory ventral stream," in *Proceedings of the National Academy of Sciences of the USA*. Available online at: <http://www.pnas.org/cgi/doi/10.1073/pnas.1113427109>
- Donald, M. (1998). "Mimesis and the executive suite: missing links in language evolution," in *Approaches to the Evolution of Language*, eds J. Hurford, M. Studdert-Kennedy, and C. Knight (Cambridge: Cambridge University Press), 44–67.
- Ducrot, O. (1984). *Le dire et le dit*. Paris: Éditions de Minuit.
- Erwin, D. (2003). The goldilocks hypothesis. *Science* 302, 1682–1683. doi: 10.1126/science.1092859
- Fedor, A., Ittész, P., and Szathmáry, E. (2009). "The biological background of syntax evolution," in *Biological Foundations and Origin of Syntax*, eds D. Bickerton and E. Szathmáry (Cambridge, MA: MIT Press), 15–39.
- Fitch, T., Arbib, M., and Merlin, D. (2010). "A molecular genetic framework for testing hypotheses about language evolution," in *The Evolution of Language*, eds A. Smith, M. Schouwstra, B. de Boer, and K. Smith (Singapore: World Scientific Publishing), 137–144. doi: 10.1142/9789814295222\_0018
- Flavell, J. H. (1992). "Perspectives on perspective-taking," in *Piaget's Theory: Prospects and Possibilities*, eds H. Beilin and P. Pufall (Hillsdale, NJ: Erlbaum), 107–39.
- Gärdenfors, P., and Osvath, M. (2005). *The Evolution of Anticipatory Cognition as a Precursor to Symbolic Communication*. Lund: Lund University Cognitive Science.
- Giraud, A.-L., and Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nat. Neurosci.* 15, 511–517. doi: 10.1038/nn.3063
- Gould, S. J., and Lewontin, R. (1979). The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proc. R. Soc. Lond.* 1161, 581–598. doi: 10.1098/rspb.1979.0086
- Grice, P. (1975). "Logic and conversation," in *Syntax and Semantics 3: Speech Acts*, eds P. Cole and J. Morgan (New York, NY: Academic Press), 41–58.
- Hagoort, P. (2009). "Reflections on the neurobiology of syntax," in *Biological Foundations and Origin of Syntax*, eds D. Bickerton and E. Szathmáry (Cambridge, MA: MIT Press), 279–298.
- Hauser, M. D., Chomsky, N., and Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science* 298, 1569–1579. doi: 10.1126/science.298.5598.1569
- Hauser, M. D., Yang, C., Berwick, R. C., Tattersall, I., Ryan, M. J., Watumull, J., et al. (2014). The mystery of language evolution. *Front. Psychol.* 5:401. doi: 10.3389/fpsyg.2014.00401
- Hein, G., and Knight, R. T. (2008). Superior temporal sulcus—it's my area: Or is it? *J. Cogn. Neurosci.* 20, 2125–2136. doi: 10.1162/jocn.2008.20148
- Hoefler, S. (2009). *Modelling the Role of Pragmatic Plasticity in the Evolution of Linguistic Communication*. Ph.D. thesis, University of Edinburgh.
- Hornstein, N., and Boeckx, C. (2009). "Approaching universals from below: I-universals in light of a Minimalist Program for linguistic theory," in *Language Universals*, eds M. Christiansen, C. Collins, and S. Edelman (Oxford: Oxford University Press), 79–98.
- Humphreys, G. W. (1998). Neural representation of objects in space: a dual coding account. *Philos. Trans. R. Soc. Lond. B* 353, 1341–1351. doi: 10.1098/rstb.1998.0288
- Hurford, J. (2007a). The origin of Noun Phrases: reference, truth and communication. *Lingua* 117, 527–542. doi: 10.1016/j.lingua.2005.04.004
- Hurford, J. (2007b). *The Origins of Meaning: Language in the Light of Evolution*. Oxford: Oxford University Press.
- Hurley, S. (2008). The shared circuits model (SCM): how control, mirroring, and simulation can enable imitation, deliberation, and mindreading. *Behav. Brain Sci.* 31, 1–22. doi: 10.1017/S0140525X07003123
- Jackendoff, R. (2002). *Foundations of Language*. Oxford: Oxford University Press.
- Jacob, F. (1982). *The Possible and the Actual*. New York, NY: Pantheon.
- Karmiloff-Smith, A. (2006). The tortuous route from genes to behavior: a neuroconstructivist approach. *Cogn. Affect. Behav. Neurosci.* 6, 6–17. doi: 10.3758/CABN.6.1.9
- Kayne, R. (1994). *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.
- Kayne, R. S. (2010). *Comparisons and Contrasts*. New York, NY: Oxford University Press.
- Keenan, E. (1978). "Language variation and the logical structure of universal grammar," in *Language Universals*, ed H. Seiler (Tübingen: Gunter Narr Verlag), 89–123.
- Lakatos, I. (1970). "Falsification and the methodology of scientific research programmes," in *Criticism and the Growth of Knowledge*, eds I. Lakatos and A. Musgrave (Cambridge: Cambridge University Press), 91–196.



- Levinson, S. (2000). *Presumptive Meaning: The Theory of Generalized Conversational Implicature*. Cambridge, MA: MIT Press.
- Lewontin, R. (1998). "The evolution of cognition: questions we will never answer," in *Methods, Models, and Conceptual Issues*, eds D. Scarborough and S. Sternberg (Cambridge, MA: MIT Press), 107–132.
- Lieberman, P. (1968). Primate vocalizations and human linguistic ability. *J. Acoust. Soc. Am.* 44, 1157–1164. doi: 10.1121/1.1910950
- Liljencrants, J., and Lindblom, B. (1972). Numerical simulation of vowel quality systems: the role of perceptual contrast. *Language* 48, 839–862. doi: 10.2307/411991
- Lindblom, B. (1992). "Phonological units as adaptive emergents of lexical development," in *Phonological Development: Models, Research, Implications*, eds C. Ferguson, L. Menn, and C. Stoel-Gammon (Timonium, MD: York Press), 565–604.
- Locke, J. (1690/1964). *An Essay Concerning Human Understanding*. Cleveland: Meridian Books.
- MacNeillage, P. F., and Davis, B. L. (2000). On the origin of internal structure of word forms. *Science* 288, 527–531. doi: 10.1126/science.288.5465.527
- Mayberry, R. (2002). "Cognitive development in deaf children: the interface of language and perception in neuropsychology," In *Handbook of Neuropsychology*, eds S. J. Segalowitz and I. Rapin (Oxford: Elsevier Science), 71–107.
- Meister, I. G., and Iacoboni, M. (2007). No language-specific activation during linguistic processing of observed actions. *PLoS ONE* 2:e891. doi: 10.1371/journal.pone.0000891
- Meltzoff, A. N., and Moore, M. K. (1997). Explaining facial imitation. A theoretical model. *Early Develop. Parent.* 6, 179–192.
- Ménard, L. (2013). "Sensorimotor constraints and the organization of sound patterns," in *New Perspectives on the Origins of Language*, eds C. Lefebvre, B. Comrie and H. Cohen (Amsterdam: John Benjamins), 257–277.
- Murphy, G. L., and Lassaline, M. E. (1997). "Hierarchical structure in concepts and the basic level of categorization," In *Knowledge, Concepts and Categories*, eds K. Lamberts and D. R. Shanks (Cambridge, MA: MIT Press), 93–131.
- Narita, H., and Fujita, K. (2010). A naturalist reconstruction of Minimalist and evolutionary biolinguistics. *Biolinguistics* 4, 356–376.
- Neville, H. J. (1993). "Neurobiology of cognitive and language processing: effects of early experience," in *Brain Development and Cognition*, ed M. H. Johnson (Cambridge, MA: Blackwell), 424–447.
- Neville, H. J., and Bavelier, D. (1998). Neural organization and plasticity of language. *Curr. Opin. Neurobiol.* 8, 254–258. doi: 10.1016/S0959-4388(98)80148-7
- Nowak, M. A., Komarova, N. L., and Niyogi, P. (2002). Computational and evolutionary aspects of language. *Nature* 417, 611–617. doi: 10.1038/nature00771
- Oudeyer, P.-Y. (2005). Self-organization in the evolution of speech. *J. Theor. Biol.* 233, 435–449. doi: 10.1016/j.jtbi.2004.10.025
- Oudeyer, P.-Y. (2006). *Self-organization in the Evolution of Speech*. Oxford: Oxford University Press.
- Oudeyer, P.-Y. (2007). "From vocal replication to shared combinatorial speech codes: a small step for evolution, a big step for language," in *Emergence of Communication and Language*, eds C. Lyon, L. Nehaviv, and A. Cangelosi (Berlin: Springer-Verlag), 207–222.
- Penn, D. C., Holyoak, K. J., and Povinelli, D. J. (2008). Darwin's mistake: explaining the discontinuity between human and nonhuman minds. *Behav. Brain Sci.* 31, 109–130. doi: 10.1017/S0140525X08003543
- Pinker, S. (1994). *The Language Instinct*. New York, NY: William Morrow.
- Poeppel, D. (2001). Pure word deafness and the bilateral processing of the speech code. *Cogn. Sci.* 25, 679–693. doi: 10.1207/s15516709cog2505\_3
- Poeppel, D. (2003). The analysis of speech in different temporal integration windows: cerebral lateralization as 'asymmetric sampling in time'. *Speech Commun.* 41, 245–255. doi: 10.1016/S0167-6393(02)00107-3
- Povinelli, D. J. (2000). *Folk Physics for Apes*. Oxford: Oxford University Press.
- Premack, D. (2004). Is language the key to human intelligence? *Science* 303, 318–320. doi: 10.1126/science.1093993
- Prigogine, I., and Stengers, I. (1984). *Order Out of Chaos*. New York: Bantam Books.
- Quine, W. V. O. (1960). *Word and Object*. Cambridge, MA: MIT Press.
- Rizzi, L. (2004). *The Structure of CP and IP: The Cartography of Syntactic Structures (Oxford Studies in Comparative Syntax)*. Oxford: Oxford University Press.
- Rizzolatti, G., and Craighero, L. (2004). The mirror-neuron system. *Ann. Rev. Neurosci.* 27, 169–192. doi: 10.1146/annurev.neuro.27.070203.144230
- Rosch, E., and Mervis, C. B. (1975). Family resemblances: studies in the internal structure of categories. *Cogn. Psychol.* 7, 573–605. doi: 10.1016/0010-0285(75)90024-9
- Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., and Boyes-Braem, P. (1976). Basic objects in natural categories. *Cogn. Psychol.* 8, 382–439. doi: 10.1016/0010-0285(76)90013-X
- Russo, E., and Treves, A. (2011). An uncouth approach to language recursivity. *Biolinguistics* 5, 133–150.
- Sanders, L. D., and Poeppel, D. (2007). Local and global auditory processing: behavioral and ERP evidence. *Neuropsychologia* 45, 1172–1186. doi: 10.1016/j.neuropsychologia.2006.10.010
- Saussure, F. (1916). *Cours de Linguistique Générale*. Critical edition prepared by T. de Mauro. Paris: Payot.
- Schleidt, M., and Kien, J. (1997). Segmentation in behavior and what it can tell us about brain function. *Hum. Nat.* 8, 77–111. doi: 10.1007/s12110-997-1005-7
- Sperber, D., and Wilson, D. (1986). *Relevance: Communication and Cognition*. Oxford: Blackwell.
- Studdert-Kennedy, M. (2005). "How did language go discrete?" in *Language Origins*, ed M. Tallerman (Oxford: Oxford University Press), 48–67.
- Szathmáry, E. (2008). "Towards an understanding of language origins," in *Codes of Life*, ed M. Barbieri (New York, NY: Springer-Verlag), 283–313.
- Thompson, D. W. (1917). *On Growth and Form*. Cambridge: Cambridge University Press.
- Venneman, T. (1974). Theoretical word order studies: results and problems. *Pap. Ling.* 7, 5–25.
- Wang, W. (1976). "Language change. in *Origins and Evolution of Language and Speech*, eds S. Harnad, H. Steklis, and J. Lancaster (New York, NY: New York Academy of Sciences), 61–72.
- Zipf, G. K. (1965/1949). *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology*. New York, NY: Hafner.

**Conflict of Interest Statement:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Bouchard. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

