Sound production learning across species: Beyond the vocal learning dichotomy

Pedro Tiago da Silva Gonçalves Martins

Sound production learning
across species:
Beyond the vocal learning dichotomy

Pedro Tiago da Silva Gonçalves Martins

Thesis submitted to the

University of Barcelona

in partial fulfillment of the requirements of the degree of

Doctor of Philosophy

in Cognitive Science and Language

under the supervision of

Dr. Cedric Boeckx
ICREA/University of Barcelona

Dr. Bridget D. Samuels
University of Southern California

Universitat de
BARCELONA

September 2020

ii

Abstract

All humans have language. This capacity is a complex biological trait whose evolution is currently an active research topic, especially in recent years. This thesis is an attempt at contributing to this enterprise in two ways, from different angles. The first one is a critical assessment of a prominent theory of language evolution, whereby the "core properties of language" can be reduced to a single computational operation, formally irreducible, which evolved suddenly as a result of a single genetic mutation. The second one is an exploration of vocal learning, a trait present in many species which in the case of humans is part of the language capacity as the functional provider of speech.

Chapter 2 identifies a fallacious line of argument associated with the aforementioned theory of language evolution, i.e. that from the formal simplicity of an operation — in the case of this particular hypothesis, the MERGE operation — one can derive the evolutionary steps it took for it to emerge. This argument is named here the "no nalf-MERGE" fallacy. After a summary of independent reasons to doubt this hypothesis for how language evolved, it is shown why the argument is biologically untenable in the first place. This chapter lends support to the idea that language evolved gradually.

The greater part of this thesis (chapters 3 and 4) focuses on vocal learning. Vocal learning, the capacity to modify auditory output on the basis of experience, is displayed by several species, across different families. In humans, it is crucial for speech. This thesis looks at vocal learning in two ways, by first offering an extension to the *Vocal Learning Continuum*, an influential framework, and secondly by using genomic information in the human lineage to suggest that vocal learning could have been present in at least some of our ancestors, narrowing the gap between them and modern humans regarding language components.

The *Vocal Learning Continuum* helped move past the dichotomic view according to which species are either vocal learners or non vocal learners, proposing instead a typology with a more nuanced, gradual distribution of this phenotype. However, several issues remain, namely the reliance on a particular brain connection for establishing a vocal learning circuit (forebrain control of phonatory muscles), as well as the primacy given to imitation, which is but one example of vocal learning and not

the whole behavioral spectrum of this phenotype. In chapter 3. after identifying and assessing these limitations, by pointing out conceptual and empirical problems, an extension to the framework is offered, called the *Vocal Learning Contiguum*, which eschews reliance on particular brain circuits and behaviors, favoring a broader perspective and welcoming more factors as sources of variation across species.

Chapter 4 explores a link between the *SRGAP2* gene and the emergence of vocal learning in the *Homo* lineage. *SRGAP2C*, a duplication of this gene found in Neanderthals and Denisovans and also in Modern Humans (but no other extant mammals), inhibits *SRGAP2A*, the ancestral version of the gene, which modulates axon guidance associated with the SLIT-ROBO molecular pathway. A connection is drawn between the downregulatory effect on axon guidance and the formation of a cortico-laryngeal connection associated with the human vocal learning circuit.

The thesis is complemented by a series of appendices that delve in more detail into some conceptual issues surrounding the field of language evolution, namely those coming from linguistics.

Resum

Tots els humans tenen llenguatge. Aquesta capacitat és un tret biològic complex, l'evolució del qual és actualment un tema de recerca molt actiu, especialment en els darrers anys. Aquesta tesi és un intent de contribuir en aquesta línia de recerca de dues maneres.

La primera és una apreciació crítica d'una teoria prominent d'evolució del llenguatge, segons la qual les "propietats nuclears del llenguatge" es poden reduir a una sola operació computacional, formalment irreductible, la qual va evolucionar de sobte com a resultat d'una sola mutació genètica. La segona és una exploració del aprenentatge vocal, un tret que forma part de la capacitat del llenguatge com a proveïdor funcional de la parla.

El capítol 2 identifica una línia argumental fal·losa associada amb la teoria susdita d'evolució del llenguatge, i.e. que des de la simplicitat formal d'una operació — en el cas d'aquesta hipòtesi, la operació Fusió — es poden derivar els passos evolutius necesaris per al seu sorgiment. Aquest argument s'anomena en aquest capítol "no nalf-Merge fallacy". Després d'exposar raons independents per dubtar d'aquesta hipòtesi de la evolució del llenguatge, es mostra per què l'argument és biològicament insostenible. Aquest capítol dona suport a la idea de que el llenguatge va evolucionar gradualment.

La part preponderant d'aquesta tesi (capítols 3 i 4) es centra en l'aprenentatge vocal: la capacitat de modificar l'output auditiu en base a l'experiència. L'aprenentatge vocal està present en diverses espècies, de diferents famílies taxonòmiques. En humans, és crucial per a la capacitat de la parla. Aquesta tesi analitza aquest tema de dues maneres. En primer lloc s'ofereix una extensió d'un marc influent, el *Contínuum de l'Aprenentatge Vocal*. En segon lloc s'utilitza informació genòmica del llinatge humà per suggerir que l'aprenentatge vocal podria haver estat present en com a mínim alguns dels nostres ancestres, reduint el buit entre aquests i els humans moderns pel que fa a components del llenguatge.

El *Contínuum de l'Aprenentatge Vocal* va ajudar a superar la visió dicotòmica segons la qual les espècies són "aprenedors vocals" o no ho són, proposant en el seu lloc una tipologia amb una distribució més matisada i gradual d'aquest fenotip. Tanma-

teix, queden certs aspectes per resoldre. Un d'ells és la dependència en una connexió cerebral particular per establir un circuit de aprenentatge vocal (control prosencèfalic dels músculs fonatoris). Un altre és la primacia donada a l'imitació, la qual, tot i ser un exemple de aprenentatge vocal, no representa la totatilat de l'espectre comportamenmtal d'aquest fenotip. En el capítol 3, després d'identificar i valorar aquestes limitacions, assenyalant problemes conceptuals i empírics, s'ofereix una extensió a aquest marc. Aquesta extensió rep aquí el nom de *Contíguum de l'Aprenentatge Vocal*. Aquest abordatge evita la centralitat d'un circuit cerebral o comportament particulars, afavorint una perspectiva més ampla i donant la benvinguda a més factors com a fonts de variació entre espècies.

El capítol 4 explora un nexe entre el gen *SRGAP2* i el sorgiment de l'aprenentatge vocal en el llinatge *Homo*. *SRGAP2C*, una duplicació d'aquest gen, present tant en neandertals i denissovans com en humans moderns (però no en altres mamífers existents), inhibeix la versió ancestral del gen (*SRGAP2A*), la qual modula el guiatge axonal associat amb la ruta molecular SLIT-ROBO. Com a resultat d'aquesta exploració, es proposa una associació entre l'efecte de regulació negativa en el guiatge axonal i la formació d'una connexió cortico-laríngia vinculada al circuit humà d'aprenentatge vocal.

La tesi es complementa amb una sèrie d'apèndixs que s'enfoquen amb més detall en algunes qüestions conceptuals que envolten el camp de l'evolució del llenguatge, principalment aquelles que provenen del àmbit de la lingüística.

Contents

Contents

List of Figures

List of Tables

Full List of Abbreviations

AAC	central nucleus of the anterior arcopallium
ABSL	Al-Sayyid Bedouin Sign language
Am	nucleus ambiguus
BA	Brodmann's area
DM	dorsal medial nucleus of the midbrain
DST	Developmental System Theory
E-Language	External Language
EM	External MERGE
Evo-devo	Evolutionary Developmental Biology
FLB	Faculty of Language in the Broad Sense
FLN	Faculty of Language in the Narrow Sense
HEK	human embryonic kidney
I-Language	Internal Language
IM	Internal MERGE
KJ	Kuypers/Jürgens hypothesis
LMC	laryngeal motor cortex
MN	motor neurons
mRNA	messenger RNA
mya	million years ago
NCBI	National center for Biotechnology
NVL	Non vocal learner
PAG	periaqueductal gray
RA	robust nucleus of the arcopallium
shRNA	short hairpin RNA
UG	Universal Grammar
VA	vocal nucleus of the arcopallium
VL	Vocal learning/Vocal learner
VLC	Vocal Learning Continuum
VTA	ventral tegmental area

Acknowledgments

I must start by thanking Cedric Boeckx, my supervisor. Cedric has played this role from the first day I moved to Barcelona in 2012, to start a Masters where he teaches. My graduate years are synonymous with Cedric. Before I set foot in a classroom, he had accepted me into his research group and given me office space, no questions asked. That was just the beginning; after that he offered much more. Throughout all these years, Cedric has been the most demanding supervisor but at the same time the most open-minded mentor. He reads everything I send him almost before I finish writing it, and I know that by the next day or even before I will have enough reasons to go back to the drawing board, rinse and repeat, and eventually ideas start to take shape. He works harder than anyone I know, is more knowledgeable than anyone I know, and at the same time always finds the time to meet with his students, or anyone who cares to ask. I never had to think twice before knocking on his door. If I add to this the fact that I have learned more from him than from anyone else, the only natural conclusion is that I couldn't have had a better supervisor.

Secondly, I must thank Bridget Samuels, my co-supervisor. A long time ago I reviewed her book, Phonological Architecture, which I was completely fascinated by and quickly became my favorite phonology book. It showed me new kinds of questions that could be asked, and indeed it sparked my interest in vocal learning, the topic of this thesis. Little did I know I would one day meet Bridget, and little did I know I would have this world-class phonologist be part of my PhD. Over the years we could never meet very often due to a minor detail (she's in California and I'm in Barcelona), but we did meet in some unlikely places, and on those occasions I could attest not only to her expertise but also to how fun she is to be around.

I have been fortunate enough to have met or otherwise interacted with some of the people who have influenced my own thinking and work. Some have perhaps given me more than they realized. I would like to thank: Adriano Lameira, Alex Gomez-Marin, Andrea Ravignani, Asif Ghazanfar, Bart de Boer, Dan Dediu, David Poeppel, Erich Jarvis, Juan Manuel Toro, Kazuo Okanoya, Koji Fujita, Luca Bonatti, Martin Haspelmath, Mirjam Knörnschild, Rie Asano, Simon Kirby, Sonja Vernes, and Tecumseh Fitch.

I've had the luck and pleasure of sharing an office with many colleagues, which made work not always feel like work (with varying results). I have learned and shared a lot with some of them, who deserve special mention: Evelina Leivada, Anna Martínez Álvarez, Saleh Alamri, Edward Shi, Elizabeth Qing Zhang, Alejandro Andirkó, Thomas O'Rourke, Stefanie Sturm, Sara Silvente i Font, and Juan Moriano. I would like to thank as well those who visited for shorter periods and made the office a cooler place to be in: Simone Gastaldon, Marcel Ruland, Mireia Rumbo i Roig, and Lucia Troiani.

I would like to thank Toni Martí and Mariona Taulé for their help with administration. They were often the light at the end of the bureaucratic tunnel. They also happen to be really nice people.

I would like to Nada Lavrač, from the Jožef Stefan Institute, for letting me take a peek into knowledge technologies and most importantly for providing support when I least expected it, but very much needed it.

Over the years, I've met many people in conferences, schools, etc. who I wish were gonna be at the next conference I was attending, or be in the same place as me for one reason or another. The list is very long, and some people I have mentioned already, but I would like to thank especially: Bill Thompson, Ella Lattenkamp, Felix Haiduk, Marianne de Heer Kloots, Marko Simonović, Mat Cieśla, Maurício Martins, Shanti Ulfsbjorninn, and Tomás Goucha.

I would like to thank my friends in Barcelona for making me feel like there is a life outside of academia (even though half of them are also in it and that's what we end up talking about anyway): Alejandro, Claudia, Jasna, Jone, Juan, Marc Lluís, Sara, Stefanie, and Tom.

Thank you Stefanie, Andrea and Mecki for everything. You are all very special to me. I don't know how I'll survive without heated discussions about everything and nothing, and without the hora del sol.

I want to thank my friends from Aveiro, a small group of people who rarely see each other but when they do it's as if no one had ever left.

I must thank João Veloso for more reasons than I can list. He introduced me to language as a scientific topic, and to phonology as a very cool part of it. He did so unknowingly at first, by allowing me to sit in his class, which I wasn't supposed to. Already then I realized he was the best teacher I had ever seen, and this is still true. Soon after that he helped me very much knowingly, and crucially. He made me want to pursue a career in academia by making me dive head-first in it, sometimes against the statu quo. It's not always easy for undergrads to get started with research or even realize that's an option, but João made sure to show me the way, and sometimes had to create new ways. He continues to do so. Without João my life would have turned out very differently. But more important than this, his unfailing friendship and support, far outside the scope of university, deserve much more than however many words I

can fit here. Thank you for everything.

My family has always been there for me, often to the best of their ability and sometimes beyond that. Any form of gratitude I express towards them will be inadequate. *Obrigado por tudo.*

Finally, I must express my gratitude to Liza, for having more patience with me than I am capable of myself, and for making me think more about and ultimately realize what the really important things in life are. *Hvala lepa.*

This thesis is dedicated to the memory of my father, Eugénio Martins (1957—2019).

Chapter 1

Introduction

The ultimate object of study of this thesis is language, a uniquely human trait. Particularly, I am interested in studying its nature as a biological capacity, and how it evolved. There are two main ways of going about this question. One way is to try to explain how language as a whole emerged in the human species, and derive hypotheses about a "key" or "basic" component of language and how humans suddenly became endowed with it, setting us apart from all other species. This approach is widespread in generative linguistics, the tradition most concerned with the biological nature of language. The most popular articulation of such a hypothesis (Berwick and Chomsky, 2016) posits that an essential computational operation emerged suddenly, by means of a single genetic mutation. One other way of going about this question, which in this thesis is argued to be the most productive, is to take different components of language, and look at them comparatively, across domains and species, while recognizing that the factors that shape their evolution are of different kinds, all working in tandem. This approach is reminiscent work in comparative cognition (de Waal and Ferrari, 2010), which decomposes abilities and looks around to see what is shared with other species and cognitive domains. As a cognitive ability language is indeed amenable to study along these lines Fitch (2017b). Choosing this latter approach does not yield a definite answer to the question of how language evolved; instead, it gets us closer to understanding it partially, contributing bit by bit to this complex puzzle, and uncovering other bits of information about humans and other species in the process. The first approach leads researchers to pick a very broad topic (language) and derive a very narrow theory of it, impervious to significant influence from the outside. The second approach forces researchers to pick a very narrow topic of their interest which is involved in language, focus on it, and then contextualize it in a very broad picture, very much informed by results and insights from a host of fields, mirroring the broadness of language itself as capacity.

The particular component of language this thesis dwells on is vocal learning, the

ability to modify auditory output on the basis of experience. This ability is crucial to speech, a major part of language, for it is what allows us to dynamically acquire and produce the speech units we use in vocal communication, going beyond our innate repertoire. However, vocal learning is not unique to humans. Several other species have this ability as well (refs), even though they don't have language. In other species, vocal learning serves communicative functions as well, among others. Crucially, however, not all species are alike in this behavior. Traditionally, species have been considered to either have or lack this ability, much like humans are considered to have language while other species simply lack it, but it is now recognized that this typology cannot take the form of a dichotomy. More specifically, this thesis explores how vocal learning can inform our understanding of language evolution.

1.1 Reasons for studying vocal learning in language evolution

There are several reasons to choose this topic. The first is tractability. Most of language is abstract, and indeed most of what linguists are interested in are things that cannot be seen. Syntax, semantics, pragmatics, phonology etc. are for the most part looked at through formal analysis, by positing relevant units, structures, and interactions between them, none of which have physical existence, though there are physical manifestations that allow for inquiry to take place. Vocal learning, which can be seen as functionally providing speech in the case of humans, offers a point of entry into language that is very much physical. There are of course challenges that are of an abstract nature when dealing with vocal learning behavior, which become more apparent as one gets closer to its cognitive aspects (for example whether or not learning is really taking place in a given setting), but its more obvious connection with the anatomy necessary to vocalize and the fact that its identification has a lot to do with acoustic measurement makes it more tractable than other components.

The tractability of vocal learning in comparison with other elements that make up language leads us to the second reason for studying it: cross-species comparison. Since vocal learning is not specific to humans and manifests itself in a relatively overt manner, aspects of it can be probed in other species at levels which are not practical or allowed with human subjects. For example, Gaub et al. (2010) show that the *Foxp2* gene does not seem to play a role for in the production of innate vocalizations in mice. In a classic study, Konishi (1965) showed that a white-crowned will not produce normal song if it is deafened early in development. Such tampering with the genetics or developmental trajectory of humans is of course not a possibility, but it tells us more about the phenotype.

This leads to the third reason: there is independent work on vocal learning of the

kind that simply cannot exist in more traditional conceptions of language, whose research programs look inwards. Many researchers have studied vocal learning in several species for different reasons (see chapter 3 for several examples of such work), both with and without language as an ultimate research goal (for example, Ravignani et al. 2016 for pinnipeds, Tyack and Sayigh 1997 for cetaceans, Lameira 2017 for primates, etc.). This offers a constant influx of useful information that helps guide research.

A more conceptual reason for someone interested in language evolution to look at vocal learning is that research on these two topics, despite their differences, has followed similar trends, and in this sense the successes and pitfalls of each can inform the other. Since the inception of generative linguistics (Chomsky, 1957), there has been a focus on syntax as the central aspect of language, and more recently (Chomsky, 1995) of a particular computational operation (Merge) as the basic property that defines human language. Theories of language evolution following this conception of language are therefore theories of the evolution of a very specific part of language, which is moreover stipulated. Its evolution is often ascribed to a genetic mutation that could have result in a slight rewiring of the brain. Similarly (although perhaps not as extremely) in vocal learning, there are two elements that are considered central by many, one behavioral and one neurological. The behavioral aspect is imitation, which has been assumed to be the relevant phenotypical manifestation of vocal learning. The neurological aspect is a particular circuit subserving the capacity for imitation, namely direct forebrain control of the phonatory muscles (see Fitch 2010; Petkov and Jarvis 2012 for prominent examples of work in this vein.)

We are faced in both cases with very particular and limiting definitions of a phenotype, and with a preference for a very particular neurological mechanism. These stipulations guide research on both language and vocal learning.

In the case of language, this state of affairs has led to the dismissal of research on aspects of language which are not related to Merge to be relegated to a secondary position, both in priority and importance. The main evolutionary hypothesis following this view focus on how Merge evolved, and conclude that it had to emerge suddenly, since Merge is so simple formally. Dialog between proponents of this view and everyone else who works on the evolution of language and its components is very hard, if possible at all. This is attributed to how misguided all work on evolution of language that does not speak to Merge is(Hauser et al., 2014).

In the case of vocal learning the situation is similar. A species must be capable of imitation and this capacity must be subserved by a particular brain circuit for it to be considered a vocal learner. Species that don't seem to show either are considered uninteresting and research on them deemed misguided. Indeed, a lot of work on species who don't display imitation at the level of, say, songbirds, does not make the "textbook" descriptions of the vocal learning phenotype. Moreover, even though

a typology of vocal learning has been devised under the *vocal learning continuum* framework (e.g. non vocal learner, moderate vocal learner, complex vocal learner), the bar is still set at imitation, and the particular brain circuit must be present in some form in any species.

This attempt at radically reducing phenotypes tends to hinder progress (the phenotype becomes easier to define, but the boundary conditions are also more biased). I touch on this issue in chapter 2 for language in general and chapter 3 for vocal learning.

1.2   Two tenets that guide the work in this thesis

The work in this thesis is held together by two main tenets, which have many connections. In this section I describe them briefly.

1.2.1   The notion of vocal learning adopted

Vocal learning as a phenotype can be subdivided into three main types, which I reproduce here from chapter 3, since in the published version they erroneously appear without labels:

Vocal comprehension learning ability to associate a sound with a behavioral response (example: dog (Canis familiaris) response to human commands)

Vocal usage learning ability to learn the context in which a vocalization can be used (example: vervet monkey (*Chlorocebus pygerythrus*) vocalizations in response to predators)

Vocal production learning ability to modify vocalizations on the basis of experience (example: birdsong), which can converge or diverge from a model

I focus mainly on the third subtype, *vocal production learning*, which I define in such a way so as to avoid the premium put on imitation several authors favor. I argue for this notion in chapter 3, the main reason being that several behaviors in several species do not necessarily involve imitation (see Tchernichovski and Marcus 2014 or Ghazanfar et al. 2019, for example).

For reasons also explored in depth in chapter 3, I reject the notion that a particular brain circuit is necessary for a vocal learning system to be in place in any species. namely involving forebrain control of phonatory muscles (e.g., the larynx in mammals or syrinx in birds).

Rejecting these two criteria amounts to rejecting the Kuypers/Jürgens hypothesis (Fitch, 2010), which requires them both to be fulfilled, and disqualifies any species lacking them from being considered. In chapter 3 I point to empirical evidence that

in my view leaves no option but to indeed abandon the KJ hypothesis, with examples of species that:

- display vocal behavior that clearly shows learning but not imitation

- display vocal behavior that they are "not supposed to" since they lack the relevant brain circuit

This forces a broader definition of vocal learning, one that looks at this phenotype as a behavior, not to be confused with the mechanisms that may subserve it, nor with a particular case of that behavior which is more impressive and more readily identified (imitation). It yields a definition whereby vocal behavior is learned from experience, which is not necessarily auditory (it could also be of a social nature), and whose manifestation does not have to be the perfect copy of a sound unit, but rather the modification of aspects of vocalizations (which could as one possibility, among others, yield imitation).

1.2.2 Doing away with dichotomies

Another tenet that guides the work on this thesis is the rejection of dichotomies in evolutionary theories, if they are to be plausible.

Classic dichotomies in biology seem to not have held up to modern times. Most would agree now that phenotypes are not the result of either "nature" or "nurture", not due to to either "genetics" or the "environment", either "biology" or "culture". They are not clearly "innate" or "learned", or either "novel" or an "adaptation". Dichotomies are good starting points, but the picture is always more complex and overlaps are more common than not. Linguistics is rife with dichotomic views and proposals, most of which seem to be a attempt to draw a line between humans vs. other species, or a special capacity of humans vs. those of other species. We have I-language vs. E-language (explanation, refs), there are two completely separate systems fed by I-language, the conceptual-intentional system for cognitive processes and the sensorimotor system for externalization processes (refs), competence and performance, core vs. periphery, basic vs ancillary properties of language, and the list goes on. All of these have the goal of defining what's important and what's secondary.

A famous example is a distinction between Faculty of Language in the Narrow Sense and Faculty of Language in the Broad Sense, or FLN/FLB distinction for short (Hauser et al., 2002). The distinction is supposed to contrast between properties of language that are unique to it and to humans, on the one hand, and properties of language which are shared with other domains and/or species. Unsurprisingly, Hauser et al. (2002) put forward that the FLN corresponds to the Merge operation, while all other properties of language can be ascribed to FLB. For these authors, understanding FLN is the real challenge of language evolution. This of course results in a state of

affairs whereby i) there is an assumption that there is set of properties that are unique to humans and indeed language, and ii) that all other properties will not tell us much about how language evolved and therefore are not worth studying if language is what we are after. This has indeed been the case in practice, with FLB becoming a rug, under which anything which is not the magic bullet that gave humans language can be swept, and deemed irrelevant (see, for example, a series of papers published in recent years, with several authors in common, e.g. Berwick et al. 2013; Hauser et al. 2014; Bolhuis et al. 2014; Everaert et al. 2015; Bolhuis et al. 2015; Everaert et al. 2017, to name but a few). This distinction makes comparative work virtually impossible by definition. See Martins et al. (2016) for a detailed discussion on this matter.

The notion of vocal learning adopted in this thesis, as mentioned in the previous section (1.2.1), is very much related to this issue. In vocal learning studies, strong dichotomies are less obvious in the current state of research, with frameworks such as the *Vocal Learning Continuum* (VLC) (Petkov and Jarvis, 2012) calling attention to the need to be more nuanced where thinking about the distribution of this phenotype. But dichotomies still remain. Having a particular kind of brain circuit as the predictor for vocal learning ability, as per the aforementioned KJ hypothesis, is a dichotomic criterion, which is especially problematic if it also fails empirically. Behaviorally, having imitation as the one aspect that makes the phenotype interesting has the same effect. This leaves many species and behaviors in a limbo, while it is clear what they display display is indeed vocal behavior that is learned.

Furthermore, and again as per the VLC, there is still a species for which a special label is reserved: *high vocal learners*, even though nothing in the vocal learning circuit of humans sets us completely apart from other species. Other things do, but they are not in the realm of the vocal learning phenotype as pursued in the VLC. The extension to the VLC proposed in 3 attempts to do away with dichotomies of this sort.

1.3   Thesis Outline

This thesis tries to tackle some questions that arise in the study of human language as a biological capacity and vocal learning more specifically, with the tenets described previously in mind.

In chapter 2, published as Martins and Boeckx (2019), I look at what is arguably the most prominent hypothesis for language evolution coming from linguistics, namely that the central component of language is an operation MERGE and that it emerged suddenly by means of a single mutation. This hypothesis makes the following assumptions. I offer evidence against these assumptions, and most crucially why logically the hypothesis fails. I identify what I call the "no half-MERGE fallacy", the argument that from the formal simplicity of MERGE one can derive the number of biological steps that led to its emergence, and argue that it cannot hold and that therefore a different

view of language evolution is warranted.

In chapter 3, published as Martins and Boeckx (2020), I offer an extension to the *Vocal Learning Continuum*, and influential framework for the study of Vocal Production Learning. I point its conceptual and empirical limitations and propose a multidimensional framework that does not incur these same problems.

In chapter 4, published as Martins et al. (2018), I draw on data from archaic human genomes to support the idea that complex learning could already have been present in some of our extinct ancestors. I do so by looking at the *SRGAP2C*, an *SRGAP2* duplication also present in Neanderthal and Denisovans, inhibiting the ancestral *SRGAP2A* gene, which modulates axon guidance. I claim this could contribute to a cortico-laryngeal connection already in these ancient humans, eliminating one difference between them and modern humans with respect to language components.

These three chapters form the bulk of thesis and follow a conceptual order: how to think about language evolution (not a single trait, but rather a multitude of traits), how to apply the same rationale to a sub-component of language, vocal learning, and how looking at a very specific source of information can give clues about a subcomponent of language and ultimately can language evolution, bringing us back to our original goal.

After the main chapters, there are a handful of appendices dealing with issues less central to the topic of the thesis, but conceptually important for research on language evolution. These are all published papers, which together are representative of my work on the conceptual problems in the field of language evolution.

In Appendix A, published as Martins and Boeckx (2016b), I go over the different ways in which biolinguistics has been used as a term, and argue that most of these senses do not add anything conceptually or practically to the field of language evolution. In other words, I argue that biolinguistic research has been misguided for the most part, and that it no longer serves the biological and evolutionary aspirations it purports to.

In Appendix B, published as Martins and Boeckx (2016a), I go over the idea that language evolution is a mystery in principle, and that no important insights can be gained. I argue that this is only true if one adheres to a particular hypothesis of language evolution (namely, one that eschews multiple factors and steps leading to its emergence). I argue instead that language evolution is a problem, like any other scientific problem, and that we can indeed derive conclusions about it if we follow the comparative method.

Appendix C, published as Silvente i Font et al. (2020) looks at data on birds (approximately 180 species) with the goal of checking whether different traits (life-history and allometric variables) correlate with vocal learning ability. Brain-body mass ratio seems to distinguish species typically considered to be vocal leaners from those that aren't. In the context of this thesis, a possible interpretation is that vocal learning is

not tied to a specific circuit evolutionary trajectory, since different bird orders have themselves different developmental trajectories leading to apparently similar phenotypes.

There is also one further Appendix D, unpublished, which goes over some issues raised by Berwick and Chomsky (2019) regarding the paper corresponding to Chapter 2. These issues were raised in a paper published simultaneously with ours. Since this chapter is conceptually important, I find it merits defense against prominent criticism. I offer clarification on why the logic of a single mutation for MERGE does not hold, and dispel the misrepresentations of my original points.

Chapter 2

Language and complexity considerations: The no half-Merge fallacy

ESSAY

Language evolution and complexity considerations: The no half-Merge fallacy

Pedro Tiago Martins[1,2]*, Cedric Boeckx[1,2,3]

1 Section of General Linguistics, University of Barcelona, Barcelona, Spain, **2** University of Barcelona Institute for Complex Systems (UBICS), Barcelona, Spain, **3** Catalan Institute for Advanced Studies and Research (ICREA), Barcelona, Spain

* pt.martins@ub.edu
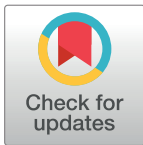
Abstract

Recently, prominent theoretical linguists have argued for an explicit scenario for the evolution of the human language capacity on the basis of its computational properties. Concretely, the simplicity of a minimalist formulation of the operation Merge, which allows humans to recursively compute hierarchical relations in language, has been used to promote a sudden-emergence, single-mutation scenario. In support of this view, Merge is said to be either fully present or fully absent: one cannot have half-Merge. On this basis, it is inferred that the emergence of our fully fledged language capacity had to be sudden. Thus, proponents of this view draw a parallelism between the formal complexity of the operation at the computational level and the number of evolutionary steps it must imply. Here, we examine this argument in detail and show that the jump from the atomicity of Merge to a single-mutation scenario is not valid and therefore cannot be used as justification for a theory of language evolution along those lines.

Introduction

The capacity for language is a defining trait of the human species. Understanding the nature of this capacity and how it came to be is a major topic of research (see [1] for a recent special issue on the topic). A leading proposal on the nature of the capacity, coming from the work of Chomsky [2], is that humans are equipped with some form of innate circuitry that allows for recursive computation over hierarchical structures. The theory describing this capacity has changed over the decades, with the most recent major articulation [3] proposing a basic operation named Merge. In its minimal expression, this operation takes two linguistic units (say, α and β) and forms a set $\{\alpha,\beta\}$, which can, in turn, function as a unit to be further combined: $\{\ldots\{\gamma,\{\alpha,\beta\}\}\ldots\}$. For example, Merge can take the units *the* and *book* and form the set $\{the, book\}$ and further merge that set with *bought* and form the set $\{bought, \{the, book\}\}$ and so on. Merge is claimed to be sufficient to yield grammatical structure and to be unique to humans.

As for the question of evolution, in a recent book, Berwick and Chomsky [4] propose that Merge, being such a simple operation, had to be the result of a single genetic mutation that endowed one individual with the necessary biological equipment for language. This idea is also defended in other recent work (e.g., [5–7]).

10

There are different parts to the position in [4], to which we will return briefly. But the key argument that interests us here is the claim that, because Merge is either fully present or fully absent, the human language faculty had to emerge suddenly, as the result of a single mutation.

The argument here is that because there can be no intermediate steps between "not having Merge" and "having Merge" as a formal operation underlying recursion—in other words, there cannot be such a thing as half-Merge—there can be no multiple, gradual evolutionary steps accounting for its emergence. Thus, Merge and, with it, a full-blown modern language faculty must have been the result of a sudden, single mutation. We will call this evolutionary scenario the "no half-Merge" argument.

In what follows, we will analyze this argument and show that it rests on tenets that do not hold (thus becoming the "no half-Merge fallacy"). We will conclude that this argument cannot be used as justification for a single-mutant theory of Merge, nor of human language, and that a different view is warranted.

The no half-Merge argument

The single-mutant theory of language evolution in [4] rests on a number of points that are presented as tightly connected. In a nutshell: there was a Great Leap Forward, an unprecedented explosion of symbolic capacity and production sometime between the appearance of anatomically modern humans and a single exodus from Africa, roughly 100,000 years ago [4]. This can only be explained by a sudden (and single) genetic change that endowed one or a very small number of individuals with very advantageous capacities, the clearest expression of which are reflected in language. The actual result of that change was the operation Merge. This operation is said to be optimal and undecomposable. Furthermore, the authors in [4] state that things could not have happened otherwise, because there was not enough time for a more complex multistep evolutionary scenario to happen in a short time span. It is very important for this proposal for each of these tenets to hold, for one rests upon the other.

Even though the present paper focuses on the atomicity of Merge and its evolutionary implications, we see evidence for doubting the other strands of the evolutionary narrative in [4]. The Great Leap Forward, single-group exodus out-of-Africa narrative, taken for granted in [4], has lost its original appeal, with mounting evidence in favor of a multigroup, multistep evolutionary trajectory of *Homo sapiens* [8–10]. Recent work has put forward models that are more consistent with the diversity evident in the fossil record, advancing the idea that several populations from different regions within Africa gave rise to anatomically modern humans [10–12]. The out-of-Africa exodus, which, it is now thought, did not consist of a single event, has been pushed as far back as approximately 120,000 years ago, because fossils do not fit the original timeline [13]. The chronologically staggered and dispersed nature of the archaeological record used to infer cognitive modernity also points to this view [8, 14]. Moreover, many of the artifacts once associated with *H. sapiens*' cognitive modernity have been attributed to then-coexisting human species [15].

The idea that Merge was the result of a single mutation and that there was not enough time for multiple mutations to give rise to it has recently been modeled, and, contrary to expectations, a multistep scenario turns out to be much more plausible [16]. The model in [16] is based on the assumptions of [4] and other information consistent with them, such as the presupposition of a single-mutation event, maximum population size at that time, the extremely large fitness advantage the change would confer, and number of offspring that would be expected. By using standard population genetic approaches (diffusion models [17] and extreme value theory [18]), the authors show that a single macromutation scenario is much less likely than one whereby several mutations have smaller fitness advantages. Therefore,

11

there seems to be no independent evolutionary-dynamics motivation for the single-mutation scenario that in [4] is called the "simplest assumption." Thus, it seems that both evolutionary dynamics and the inadequacy of the Great Leap Forward idea are independent reasons for doubting key aspects of the single-mutant theory of the evolution of language.

Let us then turn to another aspect of the proposal of [4], namely, that because Merge is atomic, it could only have evolved as the result of a single mutation, for this "phenotype" does not allow for intermediate steps. In [4], it is put as follows:

"A plausible speculation is that some small rewiring of the brain provided the core element of the Basic Property: an optimal computational procedure, which yields an infinite array of hierarchically structured expressions, each interpreted systematically at the conceptual inter-face with other cognitive systems. . . .It is, in fact, not easy to conceive of a different possibility, since there can be no series of small steps that leads to infinite yield."

The argument has been stated most succinctly (and endorsed) by [19], who makes the same inference from formal complexity (or simplicity) to evolutionary steps: "There's no such thing as half-recursion. It's an all or nothing software trick" (p. 290); "it's not totally implausible that such a faculty might have come about in a single mutation, which we should probably call a macro-mutation" (p. 382).

We now focus on the argument itself and articulate the reasons why it can't be used to justify a single-mutant theory of language evolution. We think it is worth examining this argument in detail because, in our experience, this is presented as "the last bastion of retreat" for linguists when a scenario like [4] is challenged.

The no half-Merge fallacy

The language phenotype is defined in [4] as equivalent to Merge. Under this view, theories of language evolution are theories of the evolution of Merge, and everything else is deemed peripheral.

Theories of language competence (that is, what goes on in the "head" of a speaker) rest mainly on formalization. Under the assumption that the system we are interested in is a biological one, formalizing a linguistic mechanism is equivalent to describing it at the computational level in the sense of David Marr's influential "three levels of analysis" [20]. The computational level describes what is being done. The other two levels are the algorithmic (how something is being done, by which processes) and the implementational (the physical implementation in the brain, and all the way down to the genome). It is recognized in the literature that the formal simplicity of an operation deemed crucial to language cannot be conflated with simplicity at the biological level [21, 22]. And yet, this is precisely what accounts like [4] do: they extend the atomicity of Merge (computational description) down to the implementational level (single neural circuit rewiring; single mutation).

An additional problem for an account like [4] concerns the simplicity of Merge (essentially, set formation, as described in the introduction). Such simplicity is only apparent: for Merge to adequately capture the core structural traits of linguistic competence, it must be formulated in such a way as to capture the distinction known to linguists as "external merge" (forming nested dependencies) and "internal Merge" (forming cross-serial/crossing dependencies) (Fig 1). Both kinds of dependencies occur in natural language, but the latter type, in which dependencies between items cross one another, requires more memory resources to keep track of all open dependencies across intervening elements [23].

If we go back to the hierarchy of formal languages [25] (Table 1), which we still take to be a useful categorization of the kinds of grammars that are computable, crossing dependencies were argued to require a level of complexity (mildly context sensitive) over and above that

12
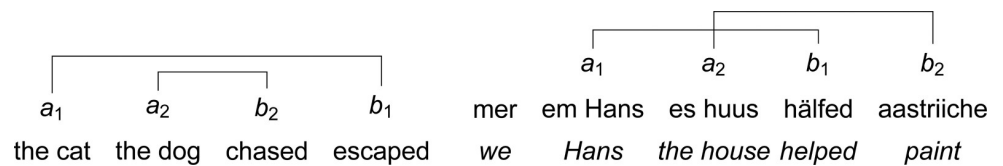
Fig 1. Nested dependencies (left) versus crossed dependencies (right). In the English example to the left, "*the cat the dog chased escaped*," the dependencies do not cross. In the Swiss-German example (from [24]), to the right, "*mer Hans es huus hälfed aastriiche*" (we helped Hans paint the house), the dependencies cross.

https://doi.org/10.1371/journal.pbio.3000389.g001

required for nested dependencies (context free). That is to say, crossing dependencies require more computational memory resources. Accordingly, they cannot simply be assumed to be part of the default Merge definition.

Thus, it is perfectly reasonable to entertain a multistep scenario for Merge, with at least two steps: one step (effectively, external Merge in the terminology mentioned earlier) taking us beyond the range of resources attested in other species' communication systems (limited to dependencies that can be captured by finite-state automata [26]). This would allow for the introduction of nested dependencies as described previously. A second step, corresponding to internal Merge, would make it possible for crossing dependencies to be part of the species' communication system (technically corresponding to the characteristics of a linear-bounded automaton; Table 1).

Note, then, that even if we grant the claim that there is no such thing as half-recursion, it doesn't follow that Merge is equally atomic. It is perfectly possible that external-Merge and internal-Merge steps took place at different times, requiring at the very least two (macro)mutations. It is also possible under this view that only one of the macromutations would be unique to modern humans.

This is where Berwick and Chomsky's [4] argument concerning the evolution of Merge and the modern language faculty rests on the accuracy of the Great Leap Forward view and the claim that there was not enough evolutionary time to accumulate the relevant mutations. Even if we grant that there cannot be such a thing as half-external-Merge, the macromutation giving rise to it could have taken place thousands of years before other mutations could affect the brain in ways that gave rise to the computational regime supporting the internal-Merge step.

To be clear, we are not suggesting that it actually took exactly two steps for Merge to arise. We simply use Berwick and Chomsky's methodological approach to try to derive evolutionary steps by looking only at formal properties and conclude that these don't entail a single mutation. Besides the reasons we mention in the previous section, there seems to be no logical necessity for a scenario such as the one in [4].

The evolution of something as complex as human language deserves integration of results and insights from different corners of the research landscape, namely the fields of neurobiology, genetics, cognitive science, comparative biology, archaeology, psychology, and linguistics. This is hard because it requires compatible levels of granularity between all fields involved, but

Table 1. The hierarchy of formal languages and corresponding automata.

Class	Grammar	Automaton
Type-3	Regular	Finite-state
Type-2	Context-free	Pushdown
Type-1	Context-sensitive	Linear bounded
Type-0	Unrestricted	Turing machine

https://doi.org/10.1371/journal.pbio.3000389.t001

13
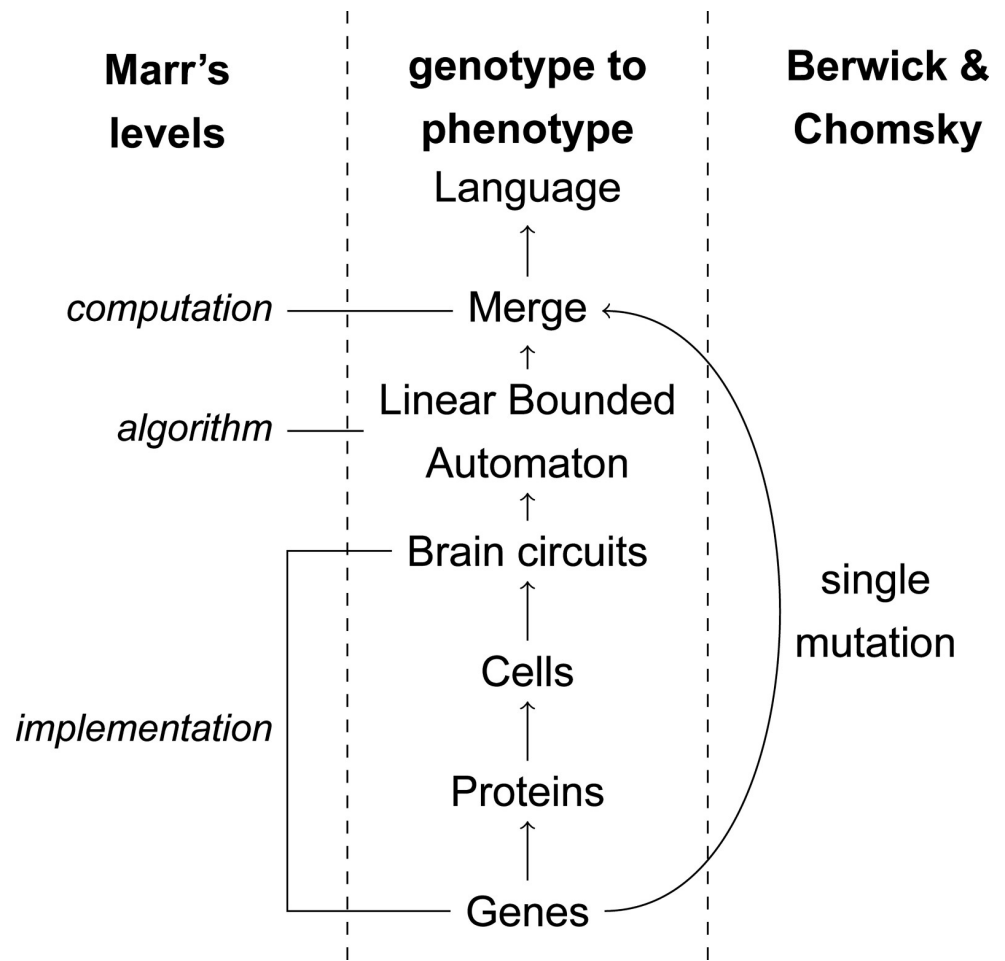
Fig 2. Berwick and Chomsky's theory of language evolution in the context of Marr's levels.

it is the only way of achieving meaningful understanding [27, 28]. This is where the a priori value of the single-mutant theory of language evolution of [4] lies. It offers a computational characterization of language that can serve as a boundary condition for other fields interested in addressing the evolutionary question. In the terms of [29], this potentially turns the question of language evolution into a "problem," as opposed to a "mystery."

Computational considerations must come to grips with both the insights and the data from other disciplines. In the case at hand, we believe theories of language evolution will benefit from taking on board the archaeological evidence questioning the Great Leap Forward. Similarly, the apparent simplicity of single-mutant scenarios should be reevaluated in light of simulations showing that multiple-mutation scenarios are more plausible, even over reduced temporal windows [16].

These empirical considerations are bypassed by work such as [4], and a single-mutation scenario is presented as arising from virtual necessity, because there cannot be such a thing as half-recursion. This eschews the fact that Merge is intended to capture a specific sense of recursion that encompasses the full range of dependencies in natural languages. Such dependencies have been known since the 1950s not to be uniform, something that the cover-term Merge obscures. This nonuniformity certainly admits a layered, mosaic-like evolutionary history.

14

Marr's levels of analysis are of great importance to our understanding of cognitive traits. The implementational level, especially, must be given more attention than it has received in works such as [4] (Fig 2). This is the level at which the farthest-reaching claims are made when devising theories of the evolution of language. Eventually, the computational description must be linked to algorithmic and implementational descriptions that connect to the neural wetware and its molecular basis. Although this is an extremely hard problem, one can be fairly confident that there won't be a one-to-one mapping between the genotype and the phenotype [30]. Moreover, we now have the possibility to look at the complete catalog of changes between modern humans and archaic humans that reached fixation [31, 32], and it is unlikely that a single mental operation could be the direct result of any of these changes.

A single computational step need not correspond to a single-mutation or a single-rewiring event. In fact, to our knowledge, there is not a single case of a novel behavior arising from a single genetic mutation. Instead, each gene deemed important is but one cog in a network of genes [32, 33]. Even in domains that are easier to probe than cognition, such as concrete physical traits, it is extremely hard to find true evolutionary novelty and even more so to attribute it to single gene changes [34, 35].

We find it problematic to rely on "logical necessity" based on the formal complexity of a trait to motivate evolutionary scenarios. It is this fallacy that we draw attention to in this paper. If one were to follow the same logic, one would put forward single-mutation evolutionary scenarios for many phenotypic traits (say, counting or bipedalism), because it is hard to conceive of what the intermediate steps of the behavior might be (can there be such a thing as half-counting or half-bipedalism?). Evolutionary studies give us daily reasons to embrace complex scenarios, and we see no reason to abandon them in the context of language. Indeed, we think that decomposing the species-specific trait of modern language into a mosaic of less exceptional ingredients, each with its own evolutionary trajectory, is the only way to open inquiry into its emergence to empirical investigation.

References

1. Fitch WT, editor. Special Issue on the Biology and Evolution of Language. Psychonomic Bulletin & Review. 2017; 24(1):1–244.

2. Chomsky N. Aspects of the Theory of Syntax. Cambridge, MA: MIT Press; 1965.

3. Chomsky N. The Minimalist Program. Cambridge, MA: MIT Press; 1995.

4. Berwick RC, Chomsky N. Why Only Us; Language and Evolution. Cambridge, MA: MIT Press; 2016.

5. Berwick RC, Friederici AD, Chomsky N, Bolhuis JJ. Evolution, Brain, and the Nature of Language. Trends in Cognitive Sciences. 2013; 17(2):89–98. https://doi.org/10.1016/j.tics.2012.12.002 PMID: 23313359

6. Bolhuis JJ, Tattersall I, Chomsky N, Berwick RC. How Could Language Have Evolved? PLoS Biol. 2014; 12(8):e1001934. https://doi.org/10.1371/journal.pbio.1001934 PMID: 25157536

7. Friederici AD, Chomsky N, Berwick RC, Moro A, Bolhuis JJ. Language, mind and brain. Nature Human Behaviour. 2017; 1(10):713. https://doi.org/10.1038/s41562-017-0184-4 PMID: 31024099

8. Mcbrearty S, Brooks AS. The Revolution That Wasn't: A New Interpretation of the Origin of Modern Human Behavior. Journal of Human Evolution. 2000; 39(5):453–563. https://doi.org/10.1006/jhev.2000.0435 PMID: 11102266

9. Stringer C. The Origin and Evolution of Homo Sapiens. Phil Trans R Soc B. 2016; 371 (1698):20150237. https://doi.org/10.1098/rstb.2015.0237 PMID: 27298468

10. Scerri EML, Thomas MG, Manica A, Gunz P, Stock JT, Stringer C, et al. Did Our Species Evolve in Subdivided Populations across Africa, and Why Does It Matter? Trends in Ecology & Evolution. 2018; 33 (8):582–594. https://doi.org/10.1016/j.tree.2018.05.005 PMID: 30007846

11. Shea JJ, Adler D, Barham L, Conard NJ, Eren MI, Lombard M, et al. Homo sapiens is as Homo sapiens was: Behavioral variability versus "behavioral modernity" in Paleolithic archaeology. Current Anthropology. 2010; 52(1):000–000.

12. Henn BM, Steele TE, Weaver TD. Clarifying distinct models of modern human origins in Africa. Current Opinion in Genetics & Development. 2018; 53:148–156. https://doi.org/10.1016/j.gde.2018.10.003 PMID: 30423527

13. Bae CJ, Douka K, Petraglia MD. On the origin of modern humans: Asian perspectives. Science. 2017; 358(6368):eaai9067. https://doi.org/10.1126/science.aai9067 PMID: 29217544

14. Henshilwood CS, Marean CW. The Origin of Modern Human Behavior. Current Anthropology. 2003; 44 (5):627–651. https://doi.org/10.1086/377665 PMID: 14971366

15. D'Errico F. The invisible frontier. A multiple species model for the origin of behavioral modernity. Evolutionary Anthropology: Issues, News, and Reviews. 2003; 12(4):188–202. https://doi.org/10.1002/evan. 10113

16. de Boer B, Thompson B, Ravignani A, Boeckx C. Evolutionary Dynamics Do Not Motivate a Single-Mutant Theory of Human Language. bioRxiv 517029 [Preprint]. 2019 [cited 2019 Mar 8] Available from: https://www.biorxiv.org/content/10.1101/517029v1

17. Kimura M. Average time until fixation of a mutant allele in a finite population under continued mutation pressure: Studies by analytical, numerical, and pseudo-sampling methods. Proceedings of the National Academy of Sciences. 1980; 77(1):522–526.

18. Orr HA. The distribution of fitness effects among beneficial mutations. Genetics. 2003; 163(4):1519–1526. PMID: 12702694

19. Dawkins R. Brief Candle in the Dark: My Life in Science. New York, NY: Ecco; 2015.

20. Marr D. Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. San Francisco: W.H. Freeman; 1982.

21. Johnson M. Marr's Levels and the Minimalist Program. Psychonomic Bulletin & Review. 2017; 24 (1):171–174. https://doi.org/10.3758/s13423-016-1062-1 PMID: 27368624

22. Perfors A. On Simplicity and Emergence. Psychonomic Bulletin & Review. 2017; 24(1):175–176. https://doi.org/10.3758/s13423-016-1157-8 PMID: 28120317

23. Uriagereka J. Syntactic Anchors: On Semantic Structuring. Cambridge: Cambridge University Press; 2008.

24. Shieber SM. Evidence Against the Context-Freeness of Natural Language. In: The Formal Complexity of Natural Language. Springer Netherlands; 1985. p. 320–334. https://doi.org/10.1007/978-94-009-3401-6_12

25. Chomsky N. Three Models for the Description of Language. IRE Transactions on Information Theory. 1956; 2(3):113–124. https://doi.org/10.1109/TIT.1956.1056813

26. Berwick RC, Okanoya K, Beckers GJL, Bolhuis JJ. Songs to Syntax: The Linguistics of Birdsong. Trends in Cognitive Sciences. 2011; 15(3):113–121. https://doi.org/10.1016/j.tics.2011.01.002 PMID: 21296608

27. Poeppel D, Embick D. Defining the Relation between Linguistics and Neuroscience. In: Cutler A, editor. Twenty-First Century Psycholinguistics. Mahwah, NJ/London: Lawrence Erlbaum; 2005. p. 103–118.

28. Krakauer JW, Ghazanfar AA, Gomez-Marin A, MacIver MA, Poeppel D. Neuroscience Needs Behavior: Correcting a Reductionist Bias. Neuron. 2017; 93(3):480–490. https://doi.org/10.1016/j.neuron.2016. 12.041 PMID: 28182904

29. Hauser MD, Yang C, Berwick RC, Tattersall I, Ryan MJ, Watumull J, et al. The Mystery of Language Evolution. Frontiers in Psychology. 2014; 5. https://doi.org/10.3389/fpsyg.2014.00401 PMID: 24847300

30. Fisher SE. Evolution of Language: Lessons from the Genome. Psychonomic Bulletin & Review. 2017; 24(1):34–40. https://doi.org/10.3758/s13423-016-1112-8 PMID: 27432000

31. Pääbo S. The Human Condition—A Molecular Approach. Cell. 2014; 157(1):216–226. https://doi.org/ 10.1016/j.cell.2013.12.036 PMID: 24679537

32. Kuhlwilm M, Boeckx C. A catalog of single nucleotide changes distinguishing modern humans from archaic hominins. Scientific Reports. 2019; 9(1). https://doi.org/10.1038/s41598-019-44877-x PMID: 31186485

33. Konopka G, Friedrich T, Davis-Turak J, Winden K, Oldham MC, Gao F, et al. Human-Specific Transcriptional Networks in the Brain. Neuron. 2012; 75(4):601–617. https://doi.org/10.1016/j.neuron.2012. 05.034 PMID: 22920253

34. Muller GB, Wagner GP. Novelty in Evolution: Restructuring the Concept. Annual Review of Ecology and Systematics. 1991; 22(1):229–256. https://doi.org/10.1146/annurev.es.22.110191.001305

35. Moczek AP. On the Origins of Novelty in Development and Evolution. BioEssays. 2008; 30(5):432–447. https://doi.org/10.1002/bies.20754 PMID: 18404691

16

Chapter 3

Vocal learning: Beyond the contiguum

ESSAY

Vocal learning: Beyond the continuum

Pedro Tiago Martins [1,2] *, Cedric Boeckx [2,3]

1 Section of General Linguistics, Universitat de Barcelona, Barcelona, Spain, **2** University of Barcelona Institute of Complex Systems (UBICS), Barcelona, Spain, **3** Catalan Institute for Research and Advanced Studies (ICREA), Barcelona, Spain

* pt.martins@ub.edu
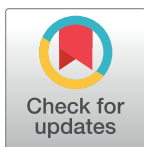
Abstract

Vocal learning is the ability to modify vocal output on the basis of experience. Traditionally, species have been classified as either displaying or lacking this ability. A recent proposal, the vocal learning continuum, recognizes the need to have a more nuanced view of this phenotype and abandon the yes–no dichotomy. However, it also limits vocal learning to production of novel calls through imitation, moreover subserved by a forebrain-to-phonatory-muscles circuit. We discuss its limitations regarding the characterization of vocal learning across species and argue for a more permissive view.

Introduction

Humans (and no other species) have language. An important component for language is speech, which refers to the auditory/vocal medium we use to communicate linguistic units among ourselves, and apart from auditory perception and processing, it requires a vocal tract with a wide range of possibilities, such as ours (but not necessarily exactly like ours [1]), and the capacity for vocal learning. Vocal learning broadly construed is the ability to modify vocal output on the basis of experience. Unlike language, however, vocal learning is not unique to humans: it is present in several, distantly related species. There are different ways in which species modify their vocalizations (see Box 1). This could involve either a modification of an aspect of vocalization (as long as it can be shown that such a modification is learned from experience, usually auditory but not necessarily so [2]) or the production of novel vocalizations altogether.

Vocal learning is indeed a very productive area of study across disciplines [6]. Species that display vocal learning abilities are a relevant source of information on the nature and evolution of language in humans, chiefly regarding phonological aspects [7]. However, not everyone agrees on what constitutes vocal learning as a phenotype, and this greatly affects how work on vocal learning is carried out.

The "canonical" list of vocal learners

There is a general trend in the literature (e.g., [8]) that limits vocal learning only to species that can produce novel calls through imitation, subserved by a direct connection between the

Box 1. Simple vocal learning typology

When looking at vocal abilities, there are ways of teasing apart which behaviors require some form of learning and which don't. There is a major split between vocalizations that are innate, whose employment does not depend on experience, and those that do require experience and go beyond the innate repertoire. The former kind of vocal behavior is common to most animals. It includes crying and laughing [3], for example, and does not require experience (though see [4] for how experience influences prosodic aspects of crying). The latter kind is less widespread, and it consists broadly of three subtypes, following [5]:

- ability to associate a sound with a behavioral response (example: dog [*Canis familiaris*] response to human commands)

- ability to learn the context in which a vocalization can be used (example: vervet monkey [*Chlorocebus pygerythrus*] vocalizations in response to predators)

- ability to modify vocalizations on the basis of experience (example: birdsong), which can converge or diverge from a model

The behavior that interests us here is vocal production learning, which is what most researchers refer to when they refer to vocal learning. But what constitutes vocal production learning as a phenotype is far from agreed upon by researchers, both in contrast to the other subtypes and on its own. This naturally affects which species are considered capable of it.

forebrain and phonatory muscles (e.g., the larynx in mammals or syrinx in birds). This has led to a canonical list of vocal learning species. It comprises three bird orders and some mammals. The birds—by far the longest list when counting individual species (in the thousands [9])—comprise songbirds (Passeriformes), parrots (Psittaciformes), and hummingbirds (Trochiliformes) [10, 11]. The mammals include humans, some cetaceans [12, 13], pinnipeds [14, 15], elephants [16], and bats [17, 18]. Birds are considered closest to humans in vocal learning abilities, even though they are phylogenetically the most remote. Humans are the only primate uncontroversially considered to be vocal learners, whereas nonhuman primates are considered of little relevance in this regard. Refinements to this list usually consist of looking inward for finer distinctions in the families already established (e.g., bats [18] or parrots [19]), and rarely outward.

It is possible that this focus on imitation and novel vocalizations is due to it being the clearest case of something being "transmitted" and then "learned." That is, it could be that for those who put a premium on imitation (e.g., [8]), learned (as opposed to "innate") entails that there has to be imitation of something that wasn't there before in any form (in the repertoire), as opposed to improvisation or other ways in which sounds in a repertoire can change (for example, through social feedback or modifications of aspects of calls that do not entail an entirely novel output).

It is also possible that the appeal of "neuro-reductionism" (to virtually equate a behavioral phenotype with a neural implementation [20]), might have had an influence in the establishment of this take on vocal learning. The existence of a direct connection from forebrain to

19

phonatory muscles, allowing for fine control of those structures, is indeed an appealing idea on which to build.

There is, however, empirical evidence of vocal learning abilities in other species outside of the canonical list. Such evidence is usually behavioral and not an attempt to show direct forebrain control of phonatory muscles. Indeed, there is work questioning that such a connection is a necessary condition for vocal learning in the first place [21, 22].

The vocal learning continuum and beyond

In light of this, we think it worth discussing the "vocal learning continuum hypothesis" (VLC) [11], which categorizes species along a continuum of increasing vocal learning complexity. This is a valuable idea that goes against the traditional dichotomous view of vocal learning, according to which a species is either definitely a vocal learner or not at all. However, it too relies on production of novel calls through imitation, subserved by forebrain control of phonatory muscles, to determine the distribution of vocal learning abilities across species. Such an approach is therefore not representative of the diversity of vocal learning behavior across the animal kingdom [23, 24]. This diversity pertains not only to species for which there is recent evidence of vocal learning but also to the "well-established" vocal learning species, namely birds [25].

Indeed, we find that the way species learn to produce their communication signals should form the set of criteria that makes a species a vocal learner. The exact nature of the vocalizations and the neurobiology are of course extremely important, and they will allow for much more precise evolutionary work, but one must not lose sight of the fact that vocal learning is a behavioral phenotype, with learning as the most striking aspect.

Given this state of affairs, two ways offer themselves for future studies on vocal learning: (1) perpetuating the bifurcation between canonical and "negligible" vocal learning species or (2) turning attention to the behaviors observed and assessing them in the context of a broad sense of vocal learning, as opposed to dismissing them on neurobiological grounds alone.

Recently proposed frameworks and reflections also show some concern with this question and call for more wide-ranging perspectives on vocal learning (e.g., [26–28]).

In the remainder of the present paper, we start by going over the VLC and point out some limitations. We then point to evidence from "noncanonical" species that reinforces these limitations and conclude by outlining an extension to the VLC highlighting approaches to vocal learning that can help overcome them.

Limitations of the vocal learning continuum

The VLC proposes that species can be placed along a continuum, yielding a gradual as opposed to dichotomous classification [11, 29–31]. The categories in the VLC are as follows: vocal non-learners, limited vocal learners, moderate vocal learners, complex vocal learners, and high vocal learners.

The motivation for the VLC is that some species cannot be clearly categorized as nonvocal learners or vocal learners (in the all-or-nothing sense), with the mouse (*Mus musculus*) being such a case: they seem to have some form of song (ultrasonic vocalizations), but it is not clear whether they are learned or innate. Some aspects of it, however, seem to be contingent on social feedback, which highlights the role of experience. In addition, species with very impressive but not identical vocal learning abilities, such as songbirds versus parrots, can also more safely be placed a notch apart so that they are not equated and the nuances that distinguish them are not lost.

Although the VLC is a very welcome and important proposal for the study of vocal learning, we find that it has some limitations, which must be overcome in order to achieve a full(er) picture of the vocal learning morphospace and an understanding of its evolutionary history.

Some of the limitations of the VLC are of a conceptual nature, and some are empirical. The conceptual limitations are independent of what the VLC is actually about and, instead, have to do with the validity of establishing a two-dimensional model of a complex trait, which had to evolve. The empirical limitations have to do with applying the model to vocal learning specifically and how its predictions don't pan out, for different reasons.

Bidimensionality

Because vocal learning categories are determined by the existence and strength of a particular brain circuit in the VLC, this makes it a bidimensional system (see [28] for some recent discussion on the same issue).

The particular brain circuit is, to put it in simple terms, a direct connection from the forebrain to phonatory muscles, and it is thought to be present in some form in canonical vocal learning species. The Kuypers/Jürgens (KJ) hypothesis [8, 32] posits that such a connection is necessary for the kind of motor control that is required for vocal learning, and the VLC tacitly relies on it. This idea had already been made popular before (e.g., [33], among others), but it was perhaps made more widespread in the work of Jarvis (e.g., [11, 34]), and Fitch (e.g., [8, 35]), who named the hypothesis after two scientists who made important contributions to primate neurobiology [36–38]. In other words, even though the VLC is a more nuanced conception of how to ascribe vocal learning across species because it allows intermediate steps, it is still limited in the sense that it has the forebrain-to-phonatory-muscles connection as the sole predictor and allows for variation only in that dimension. It is relevant here to recognize the role of a direct forebrain-to-phonatory-muscles connection as a necessary ingredient in the VLC; proponents of the VLC are of course well aware that it alone cannot explain away vocal learning as a whole. Other abilities and traits are involved, such as auditory learning [11], but the VLC is not concerned with them.

An analogy to the bidimensional nature of the VLC would be a slider in a physical machine or a computer program that controls a parameter, and by sliding it back and forth, the output is changed. In this case, the parameter would be the strength of a forebrain-to-phonatory-muscles connection, and the output would be "less" or "more" vocal learning. If the slider is at position 0, we get no vocal learning. If the slider is at the maximum value, we get "high-end" vocal learning.

There are two ways in which this bidimensionality is problematic. The first is that it leaves out capacities and constraints at other levels of analysis [18, 24, 28], which might or might not go hand in hand with this brain circuit. This is well captured by the following questions, taken from [28]: (1) What makes a species a vocal learner? (2) When is vocal learning employed? (3) How can vocal learning be expressed by the organism? (4) Who (else) is capable of vocal learning? And (5) why did vocal learning evolve?

Recent empirical evidence shows that species that do not or are thought to not have relevant forebrain control of phonatory muscles can be vocal learners. This brings to the fore other ways in which species may achieve what is in effect vocal learning behavior. Testing of species whose vocal learning capacities are unknown or supposed not to exist still yields surprises.

The other problem is that if a goal (or even the main goal) of comparative work is to derive information about the evolution of traits and species, we cannot expect a single aspect (e.g., a single genetic change or a single brain connection) to offer a realistic picture of how the trait evolved [39]. Even if the empirical evidence established that only vocal learning species in any

21

one sense consistently have a certain brain connection and vice versa, evolutionary aspirations would still require a more complex explanation. In the realm of complex traits, there is always a cascade of effects with far-reaching implications [40]. It is also the case that even homologous behaviors don't necessarily share a neural mechanism: there can be genetic changes affecting circuitry with no change in behavior [41].

More on brain wiring

The discovery of particular wiring (see Box 2) made it possible to attempt a principled, brain-based separation of strictly innate calls in a way that's shared among all mammals tested from calls that are controlled volitionally. However, even in the very strict sense of learning of novel vocalizations through imitation, it is not known beyond doubt that this is a necessary condition. For example, there are reports of learned, voiced calls in the orangutan (*Pongo* spp., a

Box 2. Two major pathways

There are two major pathways believed to be specifically involved in vocal behavior: a general, "primal" one that is associated with all vocalizing animals and, in addition, a more specific one that is associated with vocal learners.

The primal pathway goes from the anterior cingulate cortex to the PAG, to the reticular formation of the pons and medulla, and from there to the phonatory neurons [3]. It seems that the PAG pathway is not involved in vocal motor coordination but, instead, is responsible for initiation and intensity of what is in effect a vocal reaction. It is not involved in its patterning.

Besides this pathway, used for "reactive" or "affective" vocalizations, it is hypothesized that vocal learners also have a direct connection from the laryngeal motor cortex to the nucleus ambiguus (Am) and, from there, to the phonatory muscles. In birds, similar pathways are thought to exist. There is a connection from the dorsal medial nucleus of the midbrain (DM) to the 12th nerve nucleus, which controls the syrinx. This is the vocalization pathway analogous to the PAG pathway in, say, humans. In vocal learning birds, there is also a connection from the robust nucleus of the arcopallium to the 12th nerve nucleus [46]. Nonvocal learning birds are thought to not have such connections (e.g., pigeons [*Columbia livia*] [47], but evidence is scarce). This direct telencephalic connection in birds is analogous to the cortical connection in humans.

Not much is known about the presence of these connections in some of the families included in the canonical list of vocal learners (e.g., cetaceans and pinnipeds).

This association between medial pathways and innate vocalizations, on the one hand, and cortical pathways and vocal learning, on the other, has become established in the literature, but the claim made by the KJ hypothesis is not without challenges [21, 22]. Most relevant here are perhaps the criticisms by Lameira [22] because they are presented in light of comparative evidence. One argument has to do with attribution: the work by Kuypers and Jürgens does not show or entail what the hypothesis states. For example, Kuypers [36] is assumed to have shown that great apes did not have the required forebrain-to-larynx connection, when in reality, he did in fact identify it in a chimpanzee (*Pan troglodytes*) subject, and Jürgens [38] used monkeys and not great apes in his work. This casts some doubt on our understanding of direct vocal control in chimpanzees and, potentially, other primates. The second argument has to do with evidence against what

the hypothesis predicts: nonhuman primates should not in any way display vocal learning. Yet evidence for primate vocal learning is accumulating (see, e.g., [22, 24, 48]). We go into more detail in section S1_Text. Evidence in the opposite direction also exists: mice seem to have the required machinery, yet they are not vocal learners in the KJ sense [29]. From a neurobiological point of view, this should mean that either this particular connection is not necessary in principle for vocal learning or that nonhuman primates actually have it and that interpretations of the few data on this matter are incorrect. The third argument has to do with the very mechanical requirements the KJ hypothesis put forward for vocal learning, which rely heavily on vocal fold control. Also in the formalization of the VLC, this is assumed explicitly: "Vocal learning is the ability to modify the spectral and syntactic composition of vocalizations generated by the vocal organ (larynx in mammals or syrinx in bird)" [29]. This requirement leaves out supralaryngeal vocal production—equivalent to voiceless consonants in humans. These vocalizations, which in humans are the most widespread [49], involve the control of several structures above the "vocal organ," such as the lips and jaw, and are used as well to expand the vocal repertoire. This might seem like a minor point, but it is worth emphasizing that human language, through speech, makes use of both voiced and voiceless sounds in all known languages. It is also the case that whispered speech, for the most part supralaryngeal, is intelligible, and there is evidence for the use of different acoustic cues in the absence of fundamental frequency [50]. Direct control of phonatory muscles—which produce voiced sounds—alone will leave a great deal unexplained. It has been suggested that, because control of these supralaryngeal structures is clearly present in nonhuman primates, laryngeal control is the extra neurobiological ingredient (a "derived trait" or autapomorphy) that made humans vocal learners [32, 51], but as far as we can assess, this only says something about the sound source and ultimately the acoustics, not about learning and, therefore, not about how ancient or widespread the ability would be in nonhuman species.

species that supposedly lacks the relevant connection [42]). It is also not clear whether the connection is sufficient (within reason) either; mice (*M. musculus*) apparently have the circuit but do not produce novel sounds through imitation [29, 30], and perhaps more interestingly, recent work shows that female zebra finches (*Taeniopygia guttata*), which do not produce learned song, have "male-like" song pathways [43], so the narrative is not totally compelling. Furthermore, there is work showing the involvement of other structures and pathways in the learning of vocal behavior in a relevant manner, such as the cerebellum [44], the periaqueductal gray (PAG) [2], or the ventral tegmental area (VTA) [45]. It is also not entirely clear why vocal learning, a phenotype whose most interesting aspect is arguably the learning part, must be limited to a certain kind of vocalization, namely the kind that requires fine control of the phonatory muscles (what is usually referred to as "phonation"). Moreover, it is important to bear in mind that any one connection does not exist in isolation; each brain region involved will be part of several other connections, each with its own complex evolutionary history.

It thus seems that, although phonatory muscle control is obviously a very useful ability, relying on the KJ hypothesis alone might not give us a good indication of the basis of vocal learning and how widespread the phenotype is.

23

Complexity considerations

The VLC also aims to represent various degrees of vocal learning complexity. But as we will see on at least three counts, it does not do so satisfactorily.

Bengalese finches (*Lonchura striata domestica*) are the domesticated strain of the white-rumped munia (*L. striata*). Domestication has been claimed to increase vocal learning complexity: if a "wild" species is already a good vocal learner, it becomes a more complex vocal learner after undergoing domestication [52]. In the case of the Bengalese finch, for example, this happens despite the fact that this bird species was not bred for its song. It is possible that imitation—crucial per the KJ hypothesis and, concomitantly, the VLC—could actually be detrimental to syntactic complexity. Compared with their wild counterparts, Bengalese finches display higher unpredictability and syntactic complexity in their song because not only do they imitate their tutors (partially) but they also improvise, resulting in what is, in effect, lower imitation fidelity [53].

The full "classic" circuit of vocal learning involves a posterior pathway for vocalization and an anterior pathway for learning [34]. Besides those pathways, parrots were discovered to have a "shell" song system in addition to the "core" song system in all vocal learning birds [19]. A larger shell system relative to the core system is associated with parrot species that have more "complex" vocal learning abilities, and vice versa. Moreover, this shell system has mostly intercortical connections, as opposed to the direct connection to the motor neurons characteristic of the core system. It seems, then, that parrot species with a larger shell system have an edge in the VLC, but this edge is not related to the direct connection the VLC rests on. This poses a challenge to the VLC as it stands because it requires the addition of an extra factor (say, adjacent "song/speech" nuclei with intercortical connections or even just "strong intercortical connections"), opening way to a much less restrictive VLC because more factors would be added as needed for capturing differences between species, departing from the bidimensionality we have already alluded to. Indeed, Chakraborty and Jarvis [54] acknowledge it might not be straightforward to reconcile the core/shell system with the VLC.

Finally, in the VLC, humans alone are considered high-end vocal learners, whereas parrots are classified at a level just below, referred to as complex vocal learners. This is purportedly because of the higher syntactic complexity in human vocalizations, but this does not rest on the criteria for categorizing species along the VLC (presence and strength of direct connection to the phonatory organ and imitation). Language complexity need not even be instantiated in vocal behavior; it is well established that the linguistic capacity is the same in sign language (see [55]). It could be that, indeed, humans are the most advanced vocal learner, but this is not possible to discern from the criteria used in the VLC. It might have more to do with the process of cultural transmission and not with anything "vocal" [56]. In a manner similar to Bengalese finches, it has been hypothesized that the increased prosociality that characterizes domestication allows for the jump in complexity to take place (see [56] for discussion).

A more permissive view

Recent work has a more wide-ranging view of what constitutes a vocal learning species and of what plays a role in it. This, we contend, is necessary in order to extend the idea of the VLC and overcome its limitations.

Imitation and de novo vocalizations are not the whole story

An important step, in our view, is to adopt a view of vocal learning behavior not necessarily focused on imitation.

The production of de novo vocalizations (new in a species repertoire) and, furthermore, doing so through imitation is often taken as the golden standard when assessing vocal learning abilities. This is problematic because imitation is one possible means of displaying vocal learning behavior. Indeed, diverging from imitation is also a common phenomenon in vocal and cultural development [23].

Perhaps a more productive conception of vocal learning is looking at learned vocal behavior as having to be acquired in some manner over developmental time, especially in contingent ways (that is, dependent on experience and not a "certainty" given the initial state of the organism).

There are interesting cases that illustrate vocal development by diverging from the tutor song; that is, by the countering of or lack of imitation. Infant marmosets, for example, develop vocal learning abilities through social reinforcement from parents, not imitation. This leads to more control of the vocal apparatus, which allows them to produce lower entropy calls [24, 57].

Canaries (*Serinus canaria domestica*) trained on atypical song imitate it at first but, when reaching maturity, shape it into the species-specific song they were never exposed to [58]. Another well-known example is the de novo emergence of zebra finch song not by imitation but, instead, by the approximation of wild-type song over a couple of generations by birds reared in isolation, with no exposure to singing tutors [59].

Evidence of this kind is good indication that vocal learning is not driven (solely) by imitation and that vocal learning ability is characterized also by behaviors that suppress imitation.

Evidence from species outside of the canonical vocal learners list

Opening up to more-permissive definitions of vocal learning goes hand in hand with opening up to the study of more species. A decent amount of evidence for vocal learning outside of the canonical list has been put forward, especially in recent years, with primates as the most representative of this trend, and some work on rodents. They moreover deserve special attention because there is resistance to taking this kind of evidence into account. Other species are more quickly accepted, perhaps because they employ imitation, and neurobiological information on these species is given a great deal of importance, given its scarcity. A good example of this is the African elephant (*Loxodonta africana*), which quickly entered the accepted list of vocal learning species [16].

The logistic difficulties in keeping and studying larger species as opposed to birds and other (usually smaller) species might also bias positions against, say, primate vocal abilities, leading to a situation in which absence of evidence might be mistaken for evidence of absence. Rodents present challenges on their own (e.g., several species produce ultrasonic vocalizations, which pose further challenges, and there is a lot of interindividual variation [60]). Studies on birds have unsurprisingly dominated vocal learning research (see data in [28]). Great ape language acquisition projects (e.g., [61]) might also have contributed to this state of affairs, given their varying goals and approaches, which usually had to do with finding some form of human language, as well as difficulty in interpreting their results. In Box 3, we summarize some evidence that we think deserves, at the very least, attention if vocal learning is to be understood as a phenotype that's characterized by the learned modification of calls, with the exact nature of the calls being an important but not (dis)qualifying feature. A more complete (yet not exhaustive) list can be checked in S1_Text.

We believe that evidence of the kind we review here has only been neglected because of its nonconformity with the KJ hypothesis. We find that one way of getting a full picture of vocal learning is placing the focus on observing the behavior, without preconceptions of what should

25

Box 3. Evidence of vocal learning abilities in species outside of the canonical vocal learners list

Primates and rodents are not usually considered to be vocal learners, yet they display behavioral traits that fall within vocal learning in some sense. This is at odds with what circuitry is thought to be required for vocal learning behavior under the KJ hypothesis and the vocal learning continuum and warrants a rethinking of what is really known about the neurobiology of vocal learning. Moreover, some of these species are as well understood as others that do count as vocal learners in literature, warranting in this case a rethinking of the motivations for including some species but not others in the canonical list. Turning first to primates, both monkeys and apes display relevant behavior. Marmosets, a New World monkey who engages in turn taking [62], uses different proportions of affiliative call types depending on social distance [63], as well as loudness relative to physical distance [64]. They can also convey identity through aspects of their calls [65]. Their calls change from infancy into adulthood, much like in humans, and limiting parent feedback disrupts this development [66]. Several Old World monkeys display relevant vocal learning abilities. Diana monkeys show call converge in social interactions [67]. Campbell's monkeys, also a turn-taking species [68], sequence the sounds in their repertoire in a nonrandom way in different situations [69]. Rhesus monkeys have a juvenile period of volitional vocalizing, disappearing once adulthood is reached [70]. Apes show striking vocal learning abilities. Orangutans can learn voiced calls [42, 71] and whistles [72, 73]. They also employ "instrumental gesture calls," whereby they volitionally use their hands or leaves in front of their mouth to lower the maximum frequency of calls [74, 75]. Gorillas [76] and chimpanzees [77] have also been shown to display vocal learning behavior. Turning now to rodents, there is promising evidence for vocal learning as well. Mice, who produce complex ultrasonic vocalizations, display variation in syllable type, which can distinguish between individuals [78]. They have also been shown to require feedback to maintain certain features of their song [30] and of changes in song development [79]. There is also Alston's mice, who engage in vocal bouts, which because of their length and patterning, have been deemed worthy of being called song [80, 81]. See S1 Text for an expanded list of species and abilities.

allow it, and then proceeding with the mechanisms. As put by Krakauer and colleagues [20]: "The neural basis of behavior cannot be properly characterized without first allowing for independent detailed study of the behavior itself."

The relationship between a neural structure and a behavior is not one of explanation of the behavior [20]. This is the case even if that relationship is consistent, which in the case of vocal learning and according to evidence we reviewed, it might not be. Although some neuronal implementation will of course be in place, there is no indication that the vocal learning phenotype can be equated with a particular one given that there is not even a consensus on what the behavior encompasses. In the face of paradoxical evidence (e.g., primates displaying vocal behavior they are not "supposed to" have), rejecting the behavioral evidence instead of revising the neural hypothesis will not lead to understanding of the trait. It is in this sense that we think it is important to have behavior as the entry point to the study of vocal learning. If the VLC is
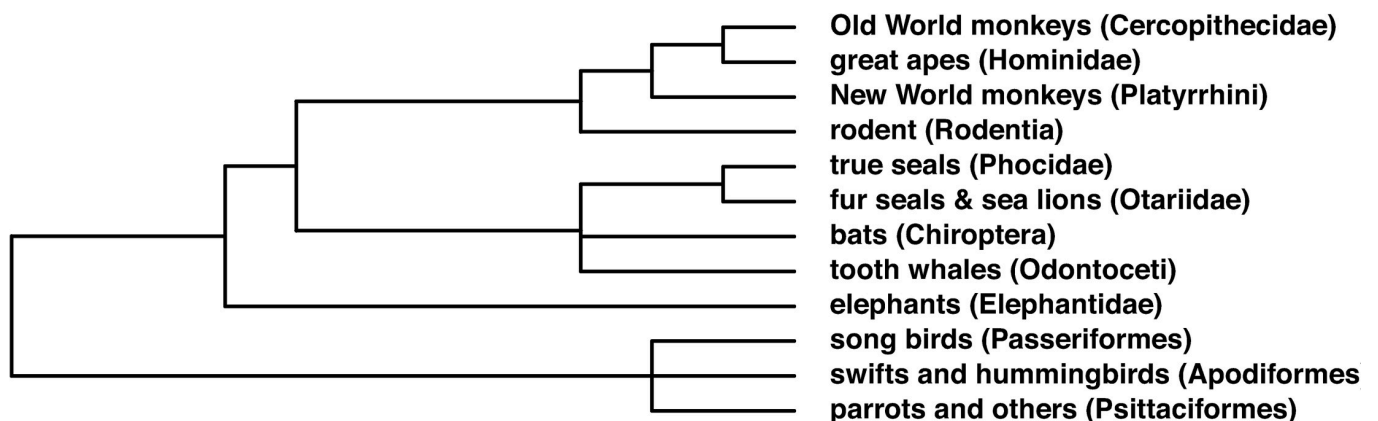
Fig 1. Tree of vocal learners with a focus on behavior. Differences between this tree and the canonical tree are reflected in the presence of primates and rodents. The inclusion of whole families or orders is made under the assumption that all members thereof are at least worth studying and is not a claim about their actual vocal learning abilities, which are an empirical question. Common and scientific names from NCBI (https://www.ncbi.nlm.nih.gov/taxonomy). Tree built with taxize R package [82]. NCBI, National Center for Biotechnology Information.

https://doi.org/10.1371/journal.pbio.3000672.g001

extended beyond the specific neural substrate that is taken to allow the direct control of the vocal organ, we could gain a better understanding of the phenotype.

Tree of vocal learners with a focus on behavior

Taking all the evidence available into account, and placing the focus in the behavior observed, we believe a more accurate "vocal learners list" (albeit with some tentative cases) would be the one we sketch in Fig 1.

We can see in this new list that it is possible to reduce the gap between us and the other vocal learners in a principled way. Although a direct forebrain–larynx connection is maybe not shared, there is much that is shared: similar patterns of early postnatal vocal development [66], volition [24], both voiceless and even voiced calls [42], socially reinforced vocal production, etc.

The canonical list of vocal learners, although much more manageable, is in effect a list of species for which there is, on the one hand, evidence of imitation and, on the other hand, evidence of direct connection from forebrain to phonatory muscles or an assumption of its existence (Fig 2, left). Assumption of its existence relies on two other assumptions: that this circuit is crucial for vocal learning and that nonhuman primates cannot have this circuit. But the fact of the matter is that there is no demonstration of this circuit for some species routinely considered vocal learners (Fig 2, center) that do show vocal imitation (Fig 2, right). In terms of evidence, there is nothing separating, say, orangutans from seals: there is evidence of imitation for both species, yet only one is an accepted vocal learner. Although one could argue about the strength of the evidence for some species over others, as well as ease of elicitation or perceived quality of the behavior, this disparity in the way different species are categorized seems to be stipulated by the KJ hypothesis and therefore warrants further reflection.

In previous work, we suggested the term "sound production" learning as opposed to vocal (production) learning as what might be a better term for defining the capacity we are interested in here, given the association in the literature of vocal with the phonatory muscles [83]. This might allow for a more encompassing definition, regardless of the mechanics involved. This would dilute a distinction that, as a characterization of the behavior, is not very relevant: if a species can change its repertoire, be it through imitation or not, the exact structures of the

Accepted vocal learners

Evidence of forebrain control of phonatory muscles

Evidence of imitation
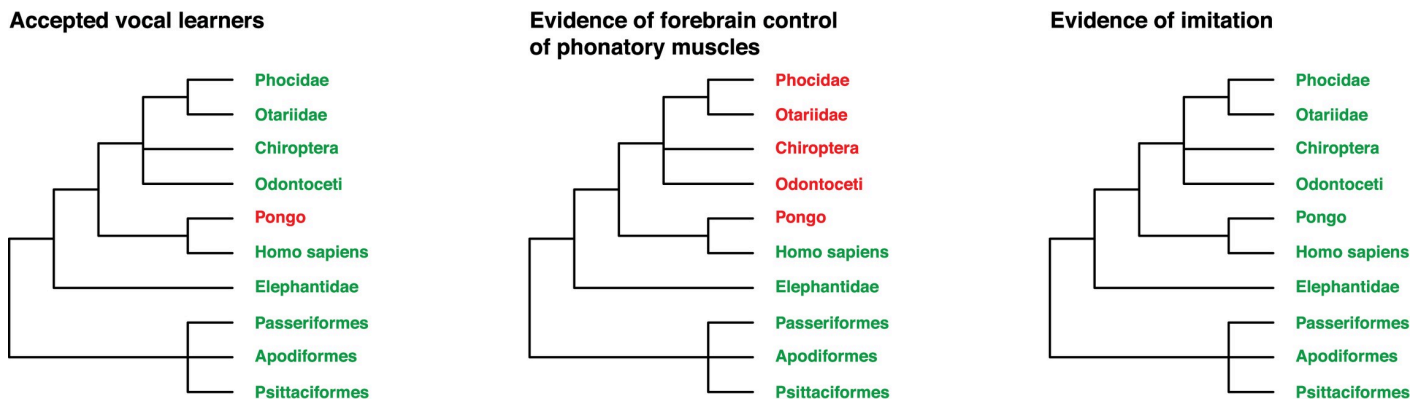
Fig 2. Comparison of simplified phylogenies of vocal learning: The canonical list with the addition of the orangutan (*Pongo*) for contrast. Key: green indicates species or orders considered to conform to the criterion on each tree, and red indicates species or orders that are considered not to. Left: list of accepted vocal learners. Center: list of species for which there is evidence of a direct connection between forebrain and phonatory muscles. Right: list of species for which there is (some) evidence of vocal imitation.

vocal tract that are used to doing so are not grounds for a big divide in the classification of the behavior, at least not with regard to learning. This is similar to what happens with the exact brain structures used in different vocal learning species: the fact that birds have no cortex and no larynx but, instead, telencephalon and syrinx does not warrant a strong divide as far as behavior is concerned, and indeed, birds are considered the prime model species for studying vocal learning in humans. Using the term sound production learning could lead to the inclusion of sound sources not limited to the vocal tract or other orofacial structures. In our view, it depends on the degree to which vocal learning is grounded in behavior and learning as opposed to the sound sources and the pathways expected to control them.

Vocal learning contiguum: An outline

Focusing on just one measure of any one cognitive trait (that is, a "two-dimensional" continuum [28]), seems to be an attempt to find neat cognitive phylogenies (in the sense of Fitch and colleagues [35]) for what is a complex behavior (see [84] for discussion). This becomes an easier task if vocal learning is reduced to a single circuit because it allows one to conceive of single events (nodes in a phylogeny) that confer the behavior to a species and its descendants (e.g., a whole order of birds). If other factors are considered, however, different cognitive phylogenies could be devised. If we ask all the questions posed by Lattenkamp and Vernes [28] for each species, we will see that there will be gaps, but we will see as well that each species provides answers to at least some of the questions. This is therefore in our view an adequate list given all the evidence.

In Fig 3, we outline an extension of the VLC, which we call the vocal learning contiguum, to capture the notion of a space of neighboring and overlapping factors, as opposed to a linear scale (as in the VLC). In this conception, vocal learning is understood as a morphospace, and a species can be represented as displaying vocal abilities of a certain type without a necessary association with either a specific neural implementation or a specific set of functional pressures. If groups of species congregate in particular areas of the morphospace, one could take this as a good indication of which factors help shape them as vocal learners and to which degree. This is in a way a simplification of the several factors that contribute to a complex behavior. A more realistic picture would comprise several dimensions. Here, for representational purposes, we collapse them into three: evidence of vocal learning behavior in a broad sense, evidence of a specialized neuronal implementation, and evidence of functional pressures
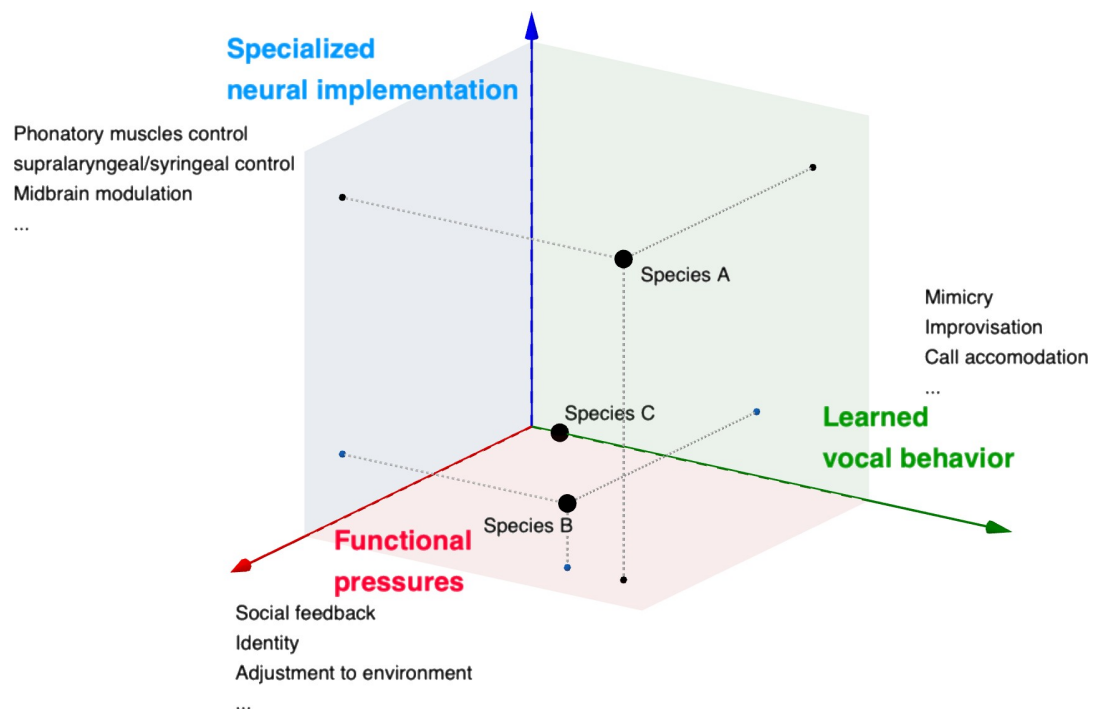
Fig 3. Outline of a vocal learning contiguum. Three main dimensions are considered: learned vocal behavior (green), functional pressures (red), and specialized neural implementation (blue). Examples are given to different factors that can be included in each of these dimensions. Black dots represent hypothetical species placed in the morphospace according to evidence. Species A displays vocal learning abilities and is thought to have a particular neuronal implementation in place, as well as subject to considerable functional pressures that help shape the behavior. Species B is a vocal learner, but evidence for a particular specialized implementation is scarce. For species C, there is no behavioral evidence of vocal learning, and no specialized neural implementation for vocal learning is thought to be present (and concomitantly no functional pressures).

(environment, social feedback [auditory or not], etc.) that help shape the behavior. The "position" of each of these species along each axis can be understood as identifying how much evidence there is for this factor playing a role. Each of the axes, which can be understood as "macrodimensions," can be further decomposed, and each of the subcomponents would also be subject to different factors.

The choice of "contiguum" for our conception of vocal learning is intentionally similar to "continuum," used in the VLC. We do not intend ours to be understood as something completely separate from the VLC but rather as an extension. In this sense, instead of a line, or even a ladder, taking us from category A to B to C and so on, we imagine a morphospace in which different forces bring a species closer to the behavior or pull apart. Like the VLC, we acknowledge the nondichotomous nature of vocal learning abilities. But we go beyond it in considering more factors than VLC for determining such abilities. The VLC would in effect be a vector in our contiguum: one of several "forces" pushing a species toward one or another phenotype. We name it "specialized neural implementation" in Fig 3. The female zebra finch, for example, which possesses pathways that some considered conducive to vocal learning [34], lacks the behavior readily displayed by males [43], which goes to show that a specific neurobiological pathway cannot be used as a predictor.

Other authors have recently put forward ideas and frameworks that we believe go in a similar direction to ours. Lattenkamp and Vernes [28] and Vernes and Wilkinson[18], though focusing on bats, call attention to the importance of behavioral, developmental, social and

motivational, neurobiological, and ecological factors that play a role in vocal learning. We think this is the right approach to take and that it can be extended to other species: other species can qualify as vocal learners if we accept that neuronal wiring is just one aspect contributing to vocal learning.

These questions of the kind posed by these authors [18, 28], in our view, follow a pedigree of influential work that has ultimately shaped biological research and contributed to a better understanding of cognition. We are referring here specifically to Tinbergen's four questions [85], which ask about mechanism, evolution, ontogeny, and function, and Marr's three levels of analysis [86]: the computational, algorithmic, and implementational levels. These frameworks have forced researchers not be to tied to any one level of description, and keeping all of them in mind when seeking understanding contributes to what Krakauer and colleagues [20] have recently called a pluralistic notion of neuroscience.

Wirthlin and colleagues [26] have a very recent proposal whereby vocal learning can be understood as being made up of different subcomponents, or "modules," and they start by looking at three: vocal coordination (ability to flexibly modify the temporal production of vocal output), vocal production variability (ability to dynamically change acoustic variability throughout development), and vocal versatility (repertoire size versus degree to which it can be modified with experience). Though not exhaustive, these three modules encapsulate several aspects commonly associated with the vocal learning phenotype. Species can be placed along "axes" for each module, and precise comparative and evolutionary characterizations can be attained.

These proposals differ in their details but find commonality in advocating for a multidimensional view of the vocal learning phenotype, which will lead to a more complex but also more accurate representation of its distribution and characteristics. Marrying ideas of this kind with evidence of the kind we review, we believe a more permissive view of vocal learning will start taking shape, encouraging further comparative studies.

Conclusions

Like other aspects of cognition, vocal learning is a mosaic, made up of different parts. The shared aspects of it should make this even less controversial than, say, language because no one can claim—as they do for language—that what other species have is very different and hard to compare to what we have.

As with any trait, an encompassing view of vocal learning makes it harder to pin down its evolutionary history and the mechanisms behind it. But reducing it to a very specific phenotype and mechanism limits the scope of comparative work, and although it might give the impression that the phenotype becomes more tractable and well-defined, it invariably forces one to subscribe to a very narrow conception that relies on a single driver. Language in general is a good (if extreme) illustration of this. Attempts have been made to reduce language to a very narrow phenotype in order to better study it (e.g., [87]). However, such approaches prevent comparative work almost by definition. As far as we can tell, reductionist views of the language phenotype have not been fruitful and have led to implausible scenarios for the evolution of language [39, 88]. The case of vocal learning is not as extreme because virtually any definition of it yields more than one species with the trait, therefore allowing for some comparative work (although according to the VLC, humans are the only high-end vocal learners). Including more species under the umbrella of any one phenotype should not be a goal in and of itself, but there is much room in our view for casting a wider net and capturing the behaviors that are now being uncovered in other species.

Supporting information

S1 Text. Evidence of vocal learning in species not traditionally considered to display vocal learning abilities.
(PDF)

References

1. Fitch WT, de Boer B, Mathur N, Ghazanfar AA. Monkey Vocal Tracts Are Speech-Ready. Science Advances. 2016; 2(12):e1600723. https://doi.org/10.1126/sciadv.1600723 PMID: 27957536

2. Tanaka M, Sun F, Li Y, Mooney R. A Mesocortical Dopamine Circuit Enables the Cultural Transmission of Vocal Behaviour. Nature. 2018; 563(7729):117–120. https://doi.org/10.1038/s41586-018-0636-7 PMID: 30333629

3. Jürgens U. The Neural Control of Vocalization in Mammals: A Review. Journal of Voice. 2009; 23(1):1–10. https://doi.org/10.1016/j.jvoice.2007.07.005 PMID: 18207362

4. Mampe B, Friederici AD, Christophe A, Wermke K. Newborns' Cry Melody Is Shaped by Their Native Language. Current Biology. 2009; 19(23):1994–1997. https://doi.org/10.1016/j.cub.2009.09.064 PMID: 19896378

5. Janik VM, Slater PJB. The Different Roles of Social Learning in Vocal Communication. Animal Behaviour. 2000; 60(1):1–11. https://doi.org/10.1006/anbe.2000.1410 PMID: 10924198

6. Shen HH. News Feature: Singing in the Brain. Proceedings of the National Academy of Sciences. 2017; 114(36):9490–9493. https://doi.org/10.1073/pnas.1711799114 PMID: 28874614

7. Samuels BD. Can a Bird Brain Do Phonology? Frontiers in Psychology. 2015; 6. https://doi.org/10.3389/fpsyg.2015.01082 PMID: 26284006

8. Fitch WT. The Evolution of Language. Cambridge: Cambridge University Press; 2010.

9. Gill F, Donsker D (eds). IOC World Bird List (v 9.2). 2019; 10.14344/IOC.ML.9.2.

10. Nottebohm F. The Neural Basis of Birdsong. PLoS Biol. 2005; 3(5):e164. https://doi.org/10.1371/journal.pbio.0030164 PMID: 15884976

11. Petkov CI, Jarvis E. Birds, Primates, and Spoken Language Origins: Behavioral Phenotypes and Neurobiological Substrates. Frontiers in Evolutionary Neuroscience. 2012; 4. https://doi.org/10.3389/fnevo.2012.00012 PMID: 22912615

12. Tyack PL, Sayigh LS. Vocal Learning in Cetaceans. In: Snowdon CT, Hausberger M, editors. Social Influences on Vocal Development. Cambridge: Cambridge University Press; 1997. p. 208–233.

13. Janik VM. Cetacean Vocal Learning and Communication. Current Opinion in Neurobiology. 2014; 28:60–65. https://doi.org/10.1016/j.conb.2014.06.010 PMID: 25057816

14. Ralls K, Fiorelli P, Gish S. Vocalizations and Vocal Mimicry in Captive Harbor Seals, *Phoca Vitulina*. Canadian Journal of Zoology. 1985; 63(5):1050–1056. https://doi.org/10.1139/z85-157

15. Ravignani A, Fitch WT, Hanke FD, Heinrich T, Hurgitsch B, Kotz SA, et al. What Pinnipeds Have to Say about Human Speech, Music, and the Evolution of Rhythm. Frontiers in Neuroscience. 2016; 10. https://doi.org/10.3389/fnins.2016.00274 PMID: 27378843

16. Stoeger AS, Manger P. Vocal Learning in Elephants: Neural Bases and Adaptive Context. Current Opinion in Neurobiology. 2014; 28:101–107. https://doi.org/10.1016/j.conb.2014.07.001 PMID: 25062469

17. Prat Y, Taub M, Yovel Y. Vocal Learning in a Social Mammal: Demonstrated by Isolation and Playback Experiments in Bats. Science Advances. 2015; 1(2):e1500019. https://doi.org/10.1126/sciadv.1500019 PMID: 26601149

18. Vernes SC, Wilkinson GS. Behaviour, Biology, and Evolution of Vocal Learning in Bats. Animal Behavior and Cognition. 2019; 375(1789):20190061. https://doi.org/10.1098/rstb.2019.0061 PMID: 31735153

19. Chakraborty M, Walløe S, Nedergaard S, Fridel EE, Dabelsteen T, Pakkenberg B, et al. Core and Shell Song Systems Unique to the Parrot Brain. PLoS ONE. 2015; 10(6):e0118496. https://doi.org/10.1371/journal.pone.0118496 PMID: 26107173

20. Krakauer JW, Ghazanfar AA, Gomez-Marin A, MacIver MA, Poeppel D. Neuroscience Needs Behavior: Correcting a Reductionist Bias. Neuron. 2017; 93(3):480–490. https://doi.org/10.1016/j.neuron.2016.12.041 PMID: 28182904

21. Lieberman P. Vocal Tract Anatomy and the Neural Bases of Talking. Journal of Phonetics. 2012; 40(4):608–622. https://doi.org/10.1016/j.wocn.2012.04.001

22. Lameira AR. Bidding Evidence for Primate Vocal Learning and the Cultural Substrates for Speech Evolution. Neuroscience & Biobehavioral Reviews. 2017; 83(Supplement C):429–439. https://doi.org/10.1016/j.neubiorev.2017.09.021 PMID: 28947156

23. Tchernichovski O, Marcus G. Vocal Learning beyond Imitation: Mechanisms of Adaptive Vocal Development in Songbirds and Human Infants. Current Opinion in Neurobiology. 2014; 28:42–47. https://doi.org/10.1016/j.conb.2014.06.002 PMID: 25005823

24. Ghazanfar AA, Liao DA, Takahashi DY. Volition and Learning in Primate Vocal Behaviour. Animal Behaviour. 2019; https://doi.org/10.1016/j.anbehav.2019.01.021

25. Brenowitz EA, Beecher MD. Song Learning in Birds: Diversity and Plasticity, Opportunities and Challenges. Trends in Neurosciences. 2005; 28(3):127–132. https://doi.org/10.1016/j.tins.2005.01.004 PMID: 15749165

26. Wirthlin M, Chang EF, Knörnschild M, Krubitzer LA, Mello CV, Miller CT, et al. A Modular Approach to Vocal Learning: Disentangling the Diversity of a Complex Behavioral Trait. Neuron. 2019; 104(1):87–99. https://doi.org/10.1016/j.neuron.2019.09.036 PMID: 31600518

27. Takahashi DY. Vocal Learning: Shaping by Social Reinforcement. Current Biology. 2019; 29(4):R125–R127. https://doi.org/10.1016/j.cub.2019.01.001 PMID: 30779900

28. Lattenkamp EZ, Vernes SC. Vocal Learning: A Language-Relevant Trait in Need of a Broad Cross-Species Approach. Current Opinion in Behavioral Sciences. 2018; 21:209–215. https://doi.org/10.1016/j.cobeha.2018.04.007

29. Arriaga G, Zhou EP, Jarvis ED. Of Mice, Birds, and Men: The Mouse Ultrasonic Song System Has Some Features Similar to Humans and Song-Learning Birds. PLoS ONE. 2012; 7(10):e46610. https://doi.org/10.1371/journal.pone.0046610 PMID: 23071596

30. Arriaga G, Jarvis ED. Mouse Vocal Communication System: Are Ultrasounds Learned or Innate? Brain and language. 2013; 124(1). https://doi.org/10.1016/j.bandl.2012.10.002 PMID: 23295209

31. Jarvis ED. Evolution of Vocal Learning and Spoken Language. Science. 2019; 366(6461):50–54. https://doi.org/10.1126/science.aax0287 PMID: 31604300

32. Fitch WT, Zuberbühler K. Primate Precursors to Human Language: Beyond Discontinuity. In: Altenmüller E, Schmidt S, Zimmermann E, editors. Evolution of Emotional Communication. Oxford University Press; 2013. p. 26–48.

33. Deacon TW. The Symbolic Species. New York: Norton; 1997.

34. Jarvis ED. Learned Birdsong and the Neurobiology of Human Language. Annals of the New York Academy of Sciences. 2004; 1016(1):749–777. https://doi.org/10.1196/annals.1298.038 PMID: 15313804

35. Fitch WT, Huber L, Bugnyar T. Social Cognition and the Evolution of Language: Constructing Cognitive Phylogenies. Neuron. 2010; 65(6):795–814. https://doi.org/10.1016/j.neuron.2010.03.011 PMID: 20346756

36. Kuypers HGJM. Some Projections from the Peri-central Cortex to the Pons and Lower Brain Stem in Monkey and Chimpanzee. Journal of Comparative Neurology. 1958; 110(2):221–255. https://doi.org/10.1002/cne.901100205 PMID: 13654557

37. Jürgens U, von Cramon D. On the Role of the Anterior Cingulate Cortex in Phonation: A Case Report. Brain and Language. 1982; 15(2):234–248. https://doi.org/10.1016/0093-934x(82)90058-x PMID: 7074343

38. Jürgens U. Neural Pathways Underlying Vocal Control. Neuroscience & Biobehavioral Reviews. 2002; 26(2):235–258. https://doi.org/10.1016/S0149-7634(01)00068-9

39. Martins PT, Boeckx C. Language Evolution and Complexity Considerations: The No Half-Merge Fallacy. PLoS Biol. 2019; 17(11):e3000389. https://doi.org/10.1371/journal.pbio.3000389 PMID: 31774810

40. Konopka G, Bomar JM, Winden K, Coppola G, Jonsson ZO, Gao F, et al. Human-Specific Transcriptional Regulation of CNS Development Genes by FOXP2. Nature. 2009; 462(7270):213. https://doi.org/10.1038/nature08549 PMID: 19907493

41. Katz PS. Evolution of Central Pattern Generators and Rhythmic Behaviours. Philosophical Transactions of the Royal Society B: Biological Sciences. 2016; 371(1685):20150057. https://doi.org/10.1098/rstb.2015.0057 PMID: 26598733

42. Lameira AR, Hardus ME, Mielke A, Wich SA, Shumaker RW. Vocal Fold Control beyond the Species-Specific Repertoire in an Orang-Utan. Scientific Reports. 2016; 6:30315. https://doi.org/10.1038/srep30315 PMID: 27461756

43. Shaughnessy DW, Hyson RL, Bertram R, Wu W, Johnson F. Female Zebra Finches Do Not Sing yet Share Neural Pathways Necessary for Singing in Males. Journal of Comparative Neurology. 2019; 527 (4):843–855. https://doi.org/10.1002/cne.24569 PMID: 30370534

32

44. Pidoux L, Le Blanc P, Levenes C, Leblois A. A Subcortical Circuit Linking the Cerebellum to the Basal Ganglia Engaged in Vocal Learning. eLife. 2018; 7:e32167. https://doi.org/10.7554/eLife.32167 PMID: 30044222

45. Kearney MG, Warren TL, Hisey E, Qi J, Mooney R. Discrete Evaluative and Premotor Circuits Enable Vocal Learning in Songbirds. Neuron. 2019; 104(3):559–575.e6. https://doi.org/10.1016/j.neuron.2019.07.025 PMID: 31447169

46. Nottebohm F, Stokes TM, Leonard CM. Central Control of Song in the Canary, Serinus Canarius. Journal of Comparative Neurology. 1976; 165(4):457–486. https://doi.org/10.1002/cne.901650405 PMID: 1262540

47. Wild JM, Li D, Eagleton C. Projections of the Dorsomedial Nucleus of the Intercollicular Complex (DM) in Relation to Respiratory-Vocal Nuclei in the Brainstem of Pigeon (Columba Livia) and Zebra Finch (Taeniopygia Guttata). Journal of Comparative Neurology. 1997; 377(3):392–413. https://doi.org/10.1002/(sici)1096-9861(19970120)377:3<392::aid-cne7>3.0.co;2-y PMID: 8989654

48. Cheney DL, Seyfarth RM. Flexible Usage and Social Function in Primate Vocalizations. Proceedings of the National Academy of Sciences. 2018:201717572. https://doi.org/10.1073/pnas.1717572115 PMID: 29432157

49. Lameira AR, Maddieson I, Zuberbühler K. Primate Feedstock for the Evolution of Consonants. Trends in Cognitive Sciences. 2014; 18(2):60–62. https://doi.org/10.1016/j.tics.2013.10.013 PMID: 24238780

50. Heeren WFL. Vocalic Correlates of Pitch in Whispered versus Normal Speech. The Journal of the Acoustical Society of America. 2015; 138(6):3800–3810. https://doi.org/10.1121/1.4937762 PMID: 26723334

51. Fitch WT. The Biology and Evolution of Speech: A Comparative Analysis. Annual Review of Linguistics. 2018; 4(1):255–279. https://doi.org/10.1146/annurev-linguistics-011817-045748

52. Okanoya K. Sexual Communication and Domestication May Give Rise to the Signal Complexity Necessary for the Emergence of Language: An Indication from Songbird Studies. Psychonomic Bulletin & Review. 2017; 24(1):106–110. https://doi.org/10.3758/s13423-016-1165-8 PMID: 28120319

53. Takahasi M, Okanoya K. Song Learning in Wild and Domesticated Strains of White-Rumped Munia, *Lonchura Striata*, Compared by Cross-Fostering Procedures: Domestication Increases Song Variability by Decreasing Strain-Specific Bias. Ethology. 2010; 116(5):396–405. https://doi.org/10.1111/j.1439-0310.2010.01761.x

54. Chakraborty M, Jarvis ED. Brain Evolution by Brain Pathway Duplication. Phil Trans R Soc B. 2015; 370(1684):20150056. https://doi.org/10.1098/rstb.2015.0056 PMID: 26554045

55. Baker A, editor. The Linguistics of Sign Languages: An Introduction. Amsterdam; Philadelphia: John Benjamins Pub. Company; 2016.

56. Thomas J, Kirby S. Self Domestication and the Evolution of Language. Biology & Philosophy. 2018; 33 (1–2). https://doi.org/10.1007/s10539-018-9612-8 PMID: 29606782

57. Takahashi DY, Liao DA, Ghazanfar AA. Vocal Learning via Social Reinforcement by Infant Marmoset Monkeys. Current Biology. 2017; 27(12):1844–1852.e6. https://doi.org/10.1016/j.cub.2017.05.004 PMID: 28552359

58. Gardner TJ, Naef F, Nottebohm F. Freedom and Rules: The Acquisition and Reprogramming of a Bird's Learned Song. Science (New York, NY). 2005; 308(5724):1046–1049. https://doi.org/10.1126/science.1108214 PMID: 15890887

59. Fehér O, Wang H, Saar S, Mitra PP, Tchernichovski O. *De Novo* Establishment of Wild-Type Song Culture in the Zebra Finch. Nature. 2009; 459(7246):564–568. https://doi.org/10.1038/nature07994 PMID: 19412161

60. Ferhat AT, Torquet N, Le Sourd AM, de Chaumont F, Olivo-Marin JC, Faure P, et al. Recording Mouse Ultrasonic Vocalizations to Evaluate Social Communication. Journal of Visualized Experiments. 2016; 112:53871. https://doi.org/10.3791/53871 PMID: 27341321

61. Hayes KJ, Hayes C. The Intellectual Development of a Home-Raised Chimpanzee. Proceedings of the American Philosophical Society. 1951; 95(2):105–109.

62. Takahashi DY, Fenley AR, Ghazanfar AA. Early Development of Turn-Taking with Parents Shapes Vocal Acoustics in Infant Marmoset Monkeys. Philosophical Transactions of the Royal Society B: Biological Sciences. 2016; 371(1693):20150370. https://doi.org/10.1098/rstb.2015.0370 PMID: 27069047

63. Liao DA, Zhang YS, Cai LX, Ghazanfar AA. Internal States and Extrinsic Factors Both Determine Monkey Vocal Production. Proceedings of the National Academy of Sciences. 2018; 115(15):3978–3983. https://doi.org/10.1073/pnas.1722426115 PMID: 29581269

64. Choi JY, Takahashi DY, Ghazanfar AA. Cooperative Vocal Control in Marmoset Monkeys via Vocal Feedback. Journal of Neurophysiology. 2015; 114(1):274–283. https://doi.org/10.1152/jn.00228.2015 PMID: 25925323

65. Miller CT, Wren Thomas A. Individual Recognition during Bouts of Antiphonal Calling in Common Marmosets. Journal of Comparative Physiology A. 2012; 198(5):337–346. https://doi.org/10.1007/s00359-012-0712-7 PMID: 22277952

66. Gultekin YB, Hage SR. Limiting Parental Feedback Disrupts Vocal Development in Marmoset Monkeys. Nature Communications. 2017; 8:14046. https://doi.org/10.1038/ncomms14046 PMID: 28090084

67. Candiotti A, Zuberbuhler K, Lemasson A. Convergence and Divergence in Diana Monkey Vocalizations. Biology Letters. 2012; 8(3):382–385. https://doi.org/10.1098/rsbl.2011.1182 PMID: 22237503

68. Lemasson A, Glas L, Barbu S, Lacroix A, Guilloux M, Remeuf K, et al. Youngsters Do Not Pay Attention to Conversational Rules: Is This so for Nonhuman Primates? Scientific Reports. 2011; 1:22. https://doi.org/10.1038/srep00022 PMID: 22355541

69. Lemasson A, Ouattara K, Petit EJ, Zuberbühler K. Social Learning of Vocal Structure in a Nonhuman Primate? BMC Evolutionary Biology. 2011; 11:362. https://doi.org/10.1186/1471-2148-11-362 PMID: 22177339

70. Hage SR, Gavrilov N, Nieder A. Developmental Changes of Cognitive Vocal Control in Monkeys. The Journal of Experimental Biology. 2016; 219(11):1744–1749. https://doi.org/10.1242/jeb.137653 PMID: 27252457

71. Lameira AR, Hardus ME, Bartlett AM, Shumaker RW, Wich SA, Menken SBJ. Speech-Like Rhythm in a Voiced and Voiceless Orangutan Call. PLoS ONE. 2015; 10(1):e116136. https://doi.org/10.1371/journal.pone.0116136 PMID: 25569211

72. Wich SA, Swartz KB, Hardus ME, Lameira AR, Stromberg E, Shumaker RW. A Case of Spontaneous Acquisition of a Human Sound by an Orangutan. Primates. 2009; 50(1):56–64. https://doi.org/10.1007/s10329-008-0117-y PMID: 19052691

73. Lameira AR, Hardus ME, Kowalsky B, de Vries H, Spruijt BM, Sterck EHM, et al. Orangutan (*Pongo* Spp.) Whistling and Implications for the Emergence of an Open-Ended Call Repertoire: A Replication and Extension. The Journal of the Acoustical Society of America. 2013; 134(3):2326–2335. https://doi.org/10.1121/1.4817929 PMID: 23967963

74. Hardus ME, Lameira AR, Schaik CPV, Wich SA. Tool Use in Wild Orang-Utans Modifies Sound Production: A Functionally Deceptive Innovation? Proceedings of the Royal Society of London B: Biological Sciences. 2009; 276(1673):3689–3694. https://doi.org/10.1098/rspb.2009.1027 PMID: 19656794

75. Lameira AR, Hardus ME, Wich SA. Orangutan Instrumental Gesture-Calls: Reconciling Acoustic and Gestural Speech Evolution Models. Evolutionary Biology. 2012; 39(3):415–418. https://doi.org/10.1007/s11692-011-9151-6 PMID: 22923853

76. Perlman M, Clark N. Learned Vocal and Breathing Behavior in an Enculturated Gorilla. Animal Cognition. 2015; 18(5):1165–1179. https://doi.org/10.1007/s10071-015-0889-6 PMID: 26139343

77. Watson SK, Townsend SW, Schel AM, Wilke C, Wallace EK, Cheng L, et al. Vocal Learning in the Functionally Referential Food Grunts of Chimpanzees. Current Biology. 2015; 25(4):495–499. https://doi.org/10.1016/j.cub.2014.12.032 PMID: 25660548

78. Holy TE, Guo Z. Ultrasonic Songs of Male Mice. PLoS Biol. 2005; 3(12):e386. https://doi.org/10.1371/journal.pbio.0030386 PMID: 16248680

79. Grimsley JMS, Monaghan JJM, Wenstrup JJ. Development of Social Vocalizations in Mice. PLoS ONE. 2011; 6(3):e17460. https://doi.org/10.1371/journal.pone.0017460 PMID: 21408007

80. Miller JR, Engstrom MD. Vocal Stereotypy and Singing Behavior in Baiomyine Mice. Journal of Mammalogy. 2007; 88(6):1447–1465. https://doi.org/10.1644/06-MAMM-A-386R.1

81. Okobi DE, Banerjee A, Matheson AMM, Phelps SM, Long MA. Motor Cortical Control of Vocal Interaction in Neotropical Singing Mice. Science. 2019; 363(6430):983–988. https://doi.org/10.1126/science.aau9480 PMID: 30819963

82. Chamberlain SA, Szöcs E. Taxize: Taxonomic Search and Retrieval in R. F1000Research. 2013; 2:191. https://doi.org/10.12688/f1000research.2-191.v2 PMID: 24555091

83. Martins PT, Boeckx C. Sound Production Learning and the Contiguum Hypothesis. Proceedings of the 12th International Conference on the Evolution of Language (Evolang12). Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika; 2018. p. 284–286.

84. Theofanopoulou C, Boeckx C. Cognitive Phylogenies, the Darwinian Logic of Descent, and the Inadequacy of Cladistic Thinking. Frontiers in Cell and Developmental Biology. 2015; 3. https://doi.org/10.3389/fcell.2015.00064 PMID: 26528479

85. Tinbergen N. On Aims and Methods of Ethology. Zeitschrift für Tierpsychologie. 1963; 20(4):410–433.

86. Marr D. Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. San Francisco: W.H. Freeman; 1982.

87. Berwick RC, Chomsky N. Why Only Us; Language and Evolution. Cambridge, MA: MIT Press; 2016.

34

88. de Boer B, Thompson B, Ravignani A, Boeckx C. Evolutionary Dynamics Do Not Motivate a Single-Mutant Theory of Human Language. Scientific Reports. 2020; 10(1):451. https://doi.org/10.1038/s41598-019-57235-8 PMID: 31949223

35

S1 Text

Pedro Tiago Martins[1,2,*] and Cedric Boeckx[1,2,3]

[1]Section of General Linguistics, Universitat de Barcelona
[2]University of Barcelona Institute for Complex Systems (UBICS)
[3]Catalan Institute for Research and Advanced Studies (ICREA)

Description

This supplementary file lists evidence of vocal learning in species not traditionally considered to display vocal learning abilities, namely primates and rodents. It offers a more complete (yet not exhaustive) list of species we find relevant in the study of vocal learning. This list is summarized in Box 3 in the main paper [1].

Abbreviations: BA, Brodmann's area; VLC, Vocal learning continuum hypothesis.

1 New world monkeys

Marmosets use different proportions of affiliative call types depending on social distance, in a way that arousal levels alone cannot predict [2]. They also control loudness relative to distance, which they assess from the loudness of their vocalizing conspecifics [3]. There are also aspects of marmoset calls that convey gender and identity [4]. Marmoset calls change during the transition from infancy into adulthood. Even though part of this change is explained by growth, but also parental feedback, much like in humans [5]. They do this not through imitation but rather through the experience-dependent increase in the control of the vocal apparatus that allows them to more consistently produce low entropy (adult-like) calls [6]. Indeed, limiting parent feedback disrupts this development [7].

Other features of the vocalizations are unaffected by experience, but instead by growth [8]. It has recently been shown that adult marmosets also spectrally modify their calls to avoid clashing with interfering sounds [9], and another study shows call structure variation between different populations [10]. Moreover, recent work shows that marmoset calls, not unlike human speech, are made up of several small units, instead of longer discrete ones [11].

Marmosets engage in vocal turn-taking [12, 13], in a way similar to humans [14]. If evidence described in this section is taken into account, there seems to be a connection between turn-taking and vocal learning, reflected mainly in humans and non-great apes (though it is possible that other

1

primate orders will reveal interesting in this regard as well). Recent work highlights the importance of turn-taking behavior in the emergence of controlled vocalizations. The auditory input that "participating" individuals receive from one another leads to changes in their own vocalizations, and the reciprocation and continuation of the exchange ultimately shapes vocal behavior [15–17].

Another species of marmoset, the pygmy marmosets (*Cebuella pygmea*), has also been studied. Similarly to the common marmoset, they are also turn-takers [18]. They display within-subspecies vocal differences [19]. Individuals with different call structure converge when paired, with long lasting effects [20]. Pygmy marmosets also change call structure under new social environments [21].

Cotton-top tamarins (*Saguinus oedipus*) change several aspects of their vocalizations in when noise is present in their environment. [22].

2 Old world monkeys

Diana monkey (*Cercopithecus diana*) show call convergence in social interactions [23]. Campbell's monkey (*Cercopithecus campbelli*), also a turn-taking species [24] display an impressive repertoire for several different situations, with non-random transition probabilities when sequencing their calls [25].

Rhesus monkey (*Macaca mulatta*) can learn to vocalize in response to (experimenter) visual cues [26]. They also have a juvenile period of volitional vocal control, which seems to disappear when adulthood is reached [27].

Japanese monkeys (*Macaca fuscata*) match acoustic features of their calls to those of the calls they are "replying" to in interactions with conspecifics [28]. Olive baboons (*Papio anubis*) adapt their vocalizations with respect to their environment, with longer, lower frequency calls in closed habitats [29]. Mandrills (*Mandrillus sphinx*) approximate their calls to closer-related conspecifics [30].

3 Apes

Chimpanzees show call convergence: after being moved to a new place, food calls of a chimpanzee population changed towards those of the resident group [31]. This was shown again in wild chimpanzees [32]. This seems to occur even within a group [33]. They also produce grooming-related calls consistent with social learning and not with associative behavior [34]. There are also reports of modification of food call structure according to tree size, with smaller trees (of the same fruit, i.e. this is not a label for a tree type) corresponding to higher pitch size [35]. Chimpanzees are also more likely to produce attention-getting vocalizations if human experimenters and food are present [36]. Interestingly, chimpanzees that produce attention-getting sounds (e.g. raspberries) have greater gray matter density compared to individuals who don't, specifically in ventrolateral prefrontal cortex and the left dorsal premotor cortex (roughly corresponding to BA 44 in humans) [37]. This lends support to the idea that these calls — which are voiceless, and therefore are not subserved by vocal fold control — are due to some form of vocal learning, as opposed to innate.

Orang-utans, even those belong to the same subspecies and therefore presumably having the same genetic makeup, sometimes employ different calls in similar situations, so much so that different populations will have a different number of calls in their repertoire. The repertoire size is correlated with the amount of social interactions within each population [38, 39]. There are also cases of acquisition of new sounds. An orang-utan in captivity has been reported to acquire a

2

human-like whistle, which it had the capacity to modulate [40]. Whistles are completely absent from known (non-human) repertoires, and later more orang-utans were shown to also have learned whistling from conspecifics [41]. There is also evidence for the learning of a voiced call, likely from humans [42], which is more remarkable in a sense, since this should require forebrain projections which orang-utans supposedly lack. This case is not isolated, as another voiced call was later reported in a different individual [43]. A very recent study shows that orangutans are capable of producing sound using a membranophone, which requires vocal fold vibration. Moreover, they do so by producing species-atypical vocalizations, and continue to alter the vocalizations to enhance the output through the membranophone [44].

Orangutans also employ what came to be known as "instrumental gesture calls": they make use of their hands or objects such as leaves to lower the maximum frequency of calls, resulting in size exaggeration. This is a way of exerting vocal control by circumventing the lack of direct control of the larynx they might have, and seems to be a volitional behavior [45,46]. Gorillas (*Gorilla gorilla*) have been shown to produce "raspberries" in some populations but crucially not in others, with no known ecological factors that cause this. In other words, some form of vocal tradition, akin to a dialect, had to be learned and transmitted [47]. There is also a study reporting the volitional production of several "breathing related-behaviors", involving dynamic coordination of breathing, larynx, tongue and lips [48]. In gibbons (*Hylobates agilis agilis*), mother-daughter vocal interactions help the development of species-specific vocalizations [49].

4   Rodents

Rodents is perhaps too broad a category, but is too understudied an order (in terms of number of species) to make bold claims about which families definitely are or are not vocal learners.

The mouse (*Mus musculus*) shows features that may qualify it as a vocal learner, produces complex (ultrasonic) vocalizations, comprising different syllable types and patterning thereof, and also a striking similarity to birdsong when adjusted to the human freq range [50]. There is also variation in syllable type which distinguishes individuals [50]. Whether their songs are innate or learned is a matter of debate (this was, indeed, what sparked the idea of VLC, since it would be hard to be categorical). Cross fostering experiments between strains that produce different patterns of vocalizations yielded individuals whose vocalizations had aspects of both [51], suggesting that calls are innate. However, sub-strains seem to match each other's pitch under certain conditions [52]. Deafening experiments reached the conclusion that mouse song is innate [53]. However, there is work showing mouse require feedback to maintain certain features of their song [54], and of changes in song across development [55].

Alston's singing mice, (*Scotinomys teguina*), engage in vocal bouts, containing sonic and ultrasonic vocalizations, which due to their length and patterning have been deemed worthy of being called song [56]. Recent results show moreover the involvement of a certain degree of cortical control, important for social, turn taking vocal interactions [57].

The common degu (*Octodon degus*), which has some 15 distinct calls in its repertoire [58], can be trained to vocalize in order to obtain food. Interestingly, they have been shown to, after such training, building nested structures with objects, something they do not do without vocal training [59]. This suggests that external experience, perhaps at a certain developmental stage, even if usually absent in normal circumstances, can promote abilities that would suffice to classify a species as vocal learner, or at the very least not rule out the ability. Also, degus exposed to shocks vocalize differently depending on whether they are being observer by a close individual or

3

a stranger [60], which highlights aspects of social learning in the patterns of vocalization in this species.

The African mole rat (*Fukomys micklemi*), a social subterranean species, has an expanded vocal repertoire when compared to other, solitary species (for example, the silvery mole-rat (*Heliophobius argenteocinereus*), among others) [61]. While this is not evidence of vocal learning beyond doubt, the fact the major difference between these species is how social they are plausibly suggests there social factors are responsible for the large and diverse repertoire of the African mole rat.

References

[1] Martins PT, Boeckx C. Beyond the Vocal Learning Continuum. PLOS Biology. 2020. doi:10.1371/journal.pbio.3000672.

[2] Liao DA, Zhang YS, Cai LX, Ghazanfar AA. Internal States and Extrinsic Factors Both Determine Monkey Vocal Production. Proceedings of the National Academy of Sciences. 2018;115(15):3978–3983. doi:10.1073/pnas.1722426115.

[3] Choi JY, Takahashi DY, Ghazanfar AA. Cooperative Vocal Control in Marmoset Monkeys via Vocal Feedback. Journal of Neurophysiology. 2015;114(1):274–283. doi:10.1152/jn.00228.2015.

[4] Miller CT, Wren Thomas A. Individual Recognition during Bouts of Antiphonal Calling in Common Marmosets. Journal of Comparative Physiology A. 2012;198(5):337–346. doi:10.1007/s00359-012-0712-7.

[5] Takahashi DY, Fenley AR, Teramoto Y, Narayanan DZ, Borjon JI, Holmes P, et al. The Developmental Dynamics of Marmoset Monkey Vocal Production. Science. 2015;349(6249):734–738. doi:10.1126/science.aab1058.

[6] Takahashi DY, Liao DA, Ghazanfar AA. Vocal Learning via Social Reinforcement by Infant Marmoset Monkeys. Current Biology. 2017;27(12):1844–1852.e6. doi:10.1016/j.cub.2017.05.004.

[7] Gultekin YB, Hage SR. Limiting Parental Feedback Disrupts Vocal Development in Marmoset Monkeys. Nature Communications. 2017;8:14046. doi:10.1038/ncomms14046.

[8] Zhang YS, Ghazanfar AA. Vocal Development through Morphological Computation. PLOS Biology. 2018;16(2):e2003933. doi:10.1371/journal.pbio.2003933.

[9] Zhao L, Rad BB, Wang X. Long-Lasting Vocal Plasticity in Adult Marmoset Monkeys. Proceedings Biological Sciences. 2019;286(1905):20190817. doi:10.1098/rspb.2019.0817.

[10] Zürcher Y, Burkart JM. Evidence for Dialects in Three Captive Populations of Common Marmosets (Callithrix Jacchus). International Journal of Primatology. 2017;38(4):780–793. doi:10.1007/s10764-017-9979-4.

[11] Pomberger T, Risueno-Segovia C, Löschner J, Hage SR. Precise Motor Control Enables Rapid Flexibility in Vocal Behavior of Marmoset Monkeys. Current Biology. 2018;0(0). doi:10.1016/j.cub.2018.01.070.

4

[12] Takahashi DY, Narayanan DZ, Ghazanfar AA. Coupled Oscillator Dynamics of Vocal Turn-Taking in Monkeys. Current Biology. 2013;23(21):2162–2168. doi:10.1016/j.cub.2013.09.005.

[13] Takahashi DY, Fenley AR, Ghazanfar AA. Early Development of Turn-Taking with Parents Shapes Vocal Acoustics in Infant Marmoset Monkeys. Philosophical Transactions of the Royal Society B: Biological Sciences. 2016;371(1693):20150370. doi:10.1098/rstb.2015.0370.

[14] Levinson SC. Turn-Taking in Human Communication – Origins and Implications for Language Processing. Trends in Cognitive Sciences. 2016;20(1):6–14. doi:10.1016/j.tics.2015.10.010.

[15] Hasson U, Ghazanfar AA, Galantucci B, Garrod S, Keysers C. Brain-to-Brain Coupling: A Mechanism for Creating and Sharing a Social World. Trends in Cognitive Sciences. 2012;16(2):114–121. doi:10.1016/j.tics.2011.12.007.

[16] Hoffmann S, Trost L, Voigt C, Leitner S, Lemazina A, Sagunsky H, et al. Duets Recorded in the Wild Reveal That Interindividually Coordinated Motor Control Enables Cooperative Behavior. Nature Communications. 2019;10(1):2577. doi:10.1038/s41467-019-10593-3.

[17] Hasson U, Frith CD. Mirroring and beyond: Coupled Dynamics as a Generalized Framework for Modelling Social Interactions. Philosophical Transactions of the Royal Society B: Biological Sciences. 2016;doi:10.1098/rstb.2015.0366.

[18] Snowdon CT, Cleveland J. "Conversations" among Pygmy Marmosets. American Journal of Primatology. 1984;7(1):15–20. doi:10.1002/ajp.1350070104.

[19] de la Torre S, Snowdon CT. Dialects in Pygmy Marmosets? Population Variation in Call Structure. American Journal of Primatology. 2009;71(4):333–342. doi:10.1002/ajp.20657.

[20] Snowdon CT, Elowson AM. Pygmy Marmosets Modify Call Structure When Paired. Ethology. 1999;105(10):893–908. doi:10.1046/j.1439-0310.1999.00483.x.

[21] Elowson AM, Snowdon CT. Pygmy Marmosets, Cebuella Pygmaea, Modify Vocal Structure in Response to Changed Social Environment. Animal Behaviour. 1994;47(6):1267–1277. doi:10.1006/anbe.1994.1175.

[22] Hotchkin CF, Parks SE, Weiss DJ. Noise-Induced Frequency Modifications of Tamarin Vocalizations: Implications for Noise Compensation in Nonhuman Primates. PLOS ONE. 2015;10(6):e0130211. doi:10.1371/journal.pone.0130211.

[23] Candiotti A, Zuberbuhler K, Lemasson A. Convergence and Divergence in Diana Monkey Vocalizations. Biology Letters. 2012;8(3):382–385. doi:10.1098/rsbl.2011.1182.

[24] Lemasson A, Glas L, Barbu S, Lacroix A, Guilloux M, Remeuf K, et al. Youngsters Do Not Pay Attention to Conversational Rules: Is This so for Nonhuman Primates? Scientific Reports. 2011;1:22. doi:10.1038/srep00022.

[25] Lemasson A, Ouattara K, Petit EJ, Zuberbühler K. Social Learning of Vocal Structure in a Nonhuman Primate? BMC Evolutionary Biology. 2011;11:362. doi:10.1186/1471-2148-11-362.

[26] Hage SR, Gavrilov N, Nieder A. Cognitive Control of Distinct Vocalizations in Rhesus Monkeys. Journal of Cognitive Neuroscience. 2013;25(10):1692–1701.

5

[27] Hage SR, Gavrilov N, Nieder A. Developmental Changes of Cognitive Vocal Control in Monkeys. The Journal of Experimental Biology. 2016;219(11):1744–1749. doi:10.1242/jeb.137653.

[28] Sugiura H. Matching of Acoustic Features during the Vocal Exchange of Coo Calls by Japanese Macaques. Animal Behaviour. 1998;55(3):673–687. doi:10.1006/anbe.1997.0602.

[29] Ey E, Rahn C, Hammerschmidt K, Fischer J. Wild Female Olive Baboons Adapt Their Grunt Vocalizations to Environmental Conditions. Ethology. 2009;115(5):493–503. doi:10.1111/j.1439-0310.2009.01638.x.

[30] Levréro F, Carrete-Vega G, Herbert A, Lawabi I, Courtiol A, Willaume E, et al. Social Shaping of Voices Does Not Impair Phenotype Matching of Kinship in Mandrills. Nature Communications. 2015;6(1):7609. doi:10.1038/ncomms8609.

[31] Watson SK, Townsend SW, Schel AM, Wilke C, Wallace EK, Cheng L, et al. Vocal Learning in the Functionally Referential Food Grunts of Chimpanzees. Current Biology. 2015;25(4):495–499. doi:10.1016/j.cub.2014.12.032.

[32] Crockford C, Herbinger I, Vigilant L, Boesch C. Wild Chimpanzees Produce Group-Specific Calls: A Case for Vocal Learning? Ethology. 2004;110(3):221–243. doi:10.1111/j.1439-0310.2004.00968.x.

[33] Mitani J, Gros-Louis J. Chorusing and Call Convergence in Chimpanzees: Tests of Three Hypotheses. Behaviour. 1998;135(8):1041–1064. doi:10.1163/156853998792913483.

[34] Watts DP. Production of Grooming-Associated Sounds by Chimpanzees (Pan Troglodytes) at Ngogo: Variation, Social Learning, and Possible Functions. Primates. 2016;57(1):61–72. doi:10.1007/s10329-015-0497-8.

[35] Kalan AK, Mundry R, Boesch C. Wild Chimpanzees Modify Food Call Structure with Respect to Tree Size for a Particular Fruit Species. Animal Behaviour. 2015;101:1–9. doi:10.1016/j.anbehav.2014.12.011.

[36] Hopkins WD, Taglialatela JP, Leavens DA. Chimpanzees Differentially Produce Novel Vocalizations to Capture the Attention of a Human. Animal Behaviour. 2007;73(2):281–286. doi:10.1016/j.anbehav.2006.08.004.

[37] Bianchi S, Reyes LD, Hopkins WD, Taglialatela JP, Sherwood CC. Neocortical Grey Matter Distribution Underlying Voluntary, Flexible Vocalizations in Chimpanzees. Scientific Reports. 2016;6:34733. doi:10.1038/srep34733.

[38] van Schaik CP, Ancrenaz M, Borgen G, Galdikas B, Knott CD, Singleton I, et al. Orangutan Cultures and the Evolution of Material Culture. Science (New York, NY). 2003;299(5603):102–105. doi:10.1126/science.1078004.

[39] Wich SA, Krützen M, Lameira AR, Nater A, Arora N, Bastian ML, et al. Call Cultures in Orang-Utans? PLoS ONE. 2012;7(5):e36180. doi:10.1371/journal.pone.0036180.

[40] Wich SA, Swartz KB, Hardus ME, Lameira AR, Stromberg E, Shumaker RW. A Case of Spontaneous Acquisition of a Human Sound by an Orangutan. Primates. 2009;50(1):56–64. doi:10.1007/s10329-008-0117-y.

6

[41] Lameira AR, Hardus ME, Kowalsky B, de Vries H, Spruijt BM, Sterck EHM, et al. Orangutan (*Pongo* Spp.) Whistling and Implications for the Emergence of an Open-Ended Call Repertoire: A Replication and Extension. The Journal of the Acoustical Society of America. 2013;134(3):2326–2335. doi:10.1121/1.4817929.

[42] Lameira AR, Hardus ME, Bartlett AM, Shumaker RW, Wich SA, Menken SBJ. Speech-Like Rhythm in a Voiced and Voiceless Orangutan Call. PLoS ONE. 2015;10(1):e116136. doi:10.1371/journal.pone.0116136.

[43] Lameira AR, Hardus ME, Mielke A, Wich SA, Shumaker RW. Vocal Fold Control beyond the Species-Specific Repertoire in an Orang-Utan. Scientific Reports. 2016;6:30315. doi:10.1038/srep30315.

[44] Lameira AR, Shumaker RW. Orangutans Show Active Voicing through a Membranophone. Scientific Reports. 2019;9(1):1–6. doi:10.1038/s41598-019-48760-7.

[45] Hardus ME, Lameira AR, Schaik CPV, Wich SA. Tool Use in Wild Orang-Utans Modifies Sound Production: A Functionally Deceptive Innovation? Proceedings of the Royal Society of London B: Biological Sciences. 2009;276(1673):3689–3694. doi:10.1098/rspb.2009.1027.

[46] Lameira AR, Hardus ME, Wich SA. Orangutan Instrumental Gesture-Calls: Reconciling Acoustic and Gestural Speech Evolution Models. Evolutionary Biology. 2012;39(3):415–418. doi:10.1007/s11692-011-9151-6.

[47] Robbins MM, Ando C, Fawcett KA, Grueter CC, Hedwig D, Iwata Y, et al. Behavioral Variation in Gorillas: Evidence of Potential Cultural Traits. PLOS ONE. 2016;11(9):e0160483. doi:10.1371/journal.pone.0160483.

[48] Perlman M, Clark N. Learned Vocal and Breathing Behavior in an Enculturated Gorilla. Animal Cognition. 2015;18(5):1165–1179. doi:10.1007/s10071-015-0889-6.

[49] Koda H, Lemasson A, Oyakawa C, Rizaldi, Pamungkas J, Masataka N. Possible Role of Mother-Daughter Vocal Interactions on the Development of Species-Specific Song in Gibbons. PLoS ONE. 2013;8(8):e71432. doi:10.1371/journal.pone.0071432.

[50] Holy TE, Guo Z. Ultrasonic Songs of Male Mice. PLoS Biology. 2005;3(12):e386. doi:10.1371/journal.pbio.0030386.

[51] Kikusui T, Nakanishi K, Nakagawa R, Nagasawa M, Mogi K, Okanoya K. Cross Fostering Experiments Suggest That Mice Songs Are Innate. PLoS ONE. 2011;6(3):e17721. doi:10.1371/journal.pone.0017721.

[52] Arriaga G, Zhou EP, Jarvis ED. Of Mice, Birds, and Men: The Mouse Ultrasonic Song System Has Some Features Similar to Humans and Song-Learning Birds. PLoS ONE. 2012;7(10):e46610. doi:10.1371/journal.pone.0046610.

[53] Mahrt EJ, Perkel DJ, Tong L, Rubel EW, Portfors CV. Engineered Deafness Reveals That Mouse Courtship Vocalizations Do Not Require Auditory Experience. Journal of Neuroscience. 2013;33(13):5573–5583. doi:10.1523/JNEUROSCI.5054-12.2013.

7

[54] Arriaga G, Jarvis ED. Mouse Vocal Communication System: Are Ultrasounds Learned or Innate? Brain and language. 2013;124(1). doi:10.1016/j.bandl.2012.10.002.

[55] Grimsley JMS, Monaghan JJM, Wenstrup JJ. Development of Social Vocalizations in Mice. PLoS ONE. 2011;6(3):e17460. doi:10.1371/journal.pone.0017460.

[56] Miller JR, Engstrom MD. Vocal Stereotypy and Singing Behavior in Baiomyine Mice. Journal of Mammalogy. 2007;88(6):1447–1465. doi:10.1644/06-MAMM-A-386R.1.

[57] Okobi DE, Banerjee A, Matheson AMM, Phelps SM, Long MA. Motor Cortical Control of Vocal Interaction in Neotropical Singing Mice. Science. 2019;363(6430):983–988. doi:10.1126/science.aau9480.

[58] Long CV. Vocalisations of the Defu *OCTODON DEGUS* ,a Social Caviomorph Rodent. Bioacoustics. 2007;16(3):223–244. doi:10.1080/09524622.2007.9753579.

[59] Tokimoto N, Okanoya K. Spontaneous Construction of "Chinese Boxes" by Degus (*Octodon Degu*): A Rudiment of Recursive Intelligence?: Cup Nesting by Degus. Japanese Psychological Research. 2004;46(3):255–261. doi:10.1111/j.1468-5584.2004.00257.x.

[60] Lidhar NK, Insel N, Dong JY, Takehara-Nishiuchi K. Observational Fear Learning in Degus Is Correlated with Temporal Vocalization Patterns. Behavioural Brain Research. 2017;332:362–371. doi:10.1016/j.bbr.2017.06.011.

[61] Vanden Hole C, Van Daele PAAG, Desmet N, Devos P, Adriaens D. Does Sociality Imply a Complex Vocal Communication System? A Case Study for *Fukomys Micklemi* (Bathyergidae, Rodentia). Bioacoustics. 2014;23(2):143–160. doi:10.1080/09524622.2013.841085.

8

Chapter 4

SRGAP2 and the gradual evolution of the modern human language faculty

OXFORD

SRGAP2 and the gradual evolution of the modern human language faculty

Pedro Tiago Martins,[1,2] Maties Marí and Cedric Boeckx[1,2,3,]*

[1]Section of General Linguistics, Universitat de Barcelona, Gran Via de les Corts Catalanes 585, Barcelona 08007, Spain, [2]Universitat de Barcelona Institute for Complex Systems, Carrer de Martí Franquès 1, Barcelona 08028, Spain and [3]ICREA, Passeig Lluís Companys 23, Barcelona 08010, Spain

*Corresponding author: cedric.boeckx@ub.edu

Abstract

In this article, we examine a new source of evidence that draws on data from archaic human genomes to support the hypothesis that vocal learning in *Homo* preceded the emergence of anatomically modern humans. We build our claim on the evolutionary history of the SLIT-ROBO GTPase 2 gene (*SRGAP2*). The SLIT-ROBO molecular pathway has been shown to have an important role in the context of vocal learning. Though the relevance of the *SRGAP2* gene duplication in the emergence of some aspect of language has not gone completely unnoticed, recent results now allow us to articulate a mechanistic hypothesis of its role in the context of axon guidance. Specifically, *SRGAP2C*, a duplication of *SRGAP2* crucially also found in Neanderthals and Denisovans, but not in extant mammals, inhibits the ancestral *SRGAP2A*, which in turn modulates the axon guidance function of the SLIT-ROBO molecular pathway. This, we claim, could have contributed to the establishment of the critical cortico-laryngeal connection of the vocal learning circuit. Our conclusions support the idea that complex vocal learning could already have been part of the arsenal of some of our extinct ancestors.

Key words: SRGAP2; vocal learning; language evolution; FOXP2; birdsong

1. Introduction

There has been much controversy among scholars regarding when the faculty of language arose in the evolutionary history of our species. Proposals put forward in the last decades cover a range of dates as large as 100,000–500,000 years ago (Bickerton 2002; Mithen 2005; Chomsky 2010; Dediu and Levinson 2013). A recent special issue on the biology and evolution of language also reflects the disparity of competing positions (Fitch 2017).

When addressing this question, a part of the problem lies in the fact that many researchers continue to see the language faculty as a homogeneous organic object. But we believe that it is far more promising, from a biological point of view, to see our linguistic competence as a complex mosaic formed by a species-specific ('novel') combination of several inherited and phylogenetically heterogeneous traits, tinkered with along traditional Darwinian lines (West-Eberhard 2003; Boeckx 2013). We expect many of these pieces of the language mosaic to be fairly straightforwardly recognized in other species (homologies), whereas other pieces may have less transparent roots (Fitch 2017). Inasmuch as the appearance and development of these various traits is directly related to genetic factors, a crucial source of evidence for tracing the phylogenetic history of language, and ultimately timing its emergence, comes from the study of the genetic material remaining in fossils of ancient organisms. Progress in paleogenetics has dramatically

changed the testability of some evolutionary scenarios (Pääbo 2014). A famous example of this was given by Krause et al. (2007), who found that *FOXP2*, a gene associated with language impairments and hampered orofacial movements (Lai et al. 2001), has the same two unique mutations in both Neanderthals and humans, critically missing in our closest extant great ape relatives. To the extent that these two mutations contributed to the establishment of some aspects of our brain's language-readiness (Enard et al. 2009; Schreiweis et al. 2014), Krause et al.'s discovery strongly suggests that aspects of our language faculty had evolved prior to the divergence of the two lineages, some 600,000 years ago (Mendez et al. 2016). In this article, we focus on the evolutionary history of *SRGAP2*, which codes for the SLIT-ROBO Rho GTPase activating protein 2 (SRGAP2). We offer, on the basis of what we have learned from other species about vocal learning, another argument in support of the idea that vocal learning was established in *Homo* before the emergence of anatomically modern humans. While the link between *SRGAP2* duplication and language evolution has been mooted before (Chakraborty et al. 2015; Hillert 2015), we show how it has now become possible to provide a mechanistic articulation of this link, making the hypothesis fully testable.

1.1 Vocal learning in birds: a mirror for human language evolution

Vocal learning is the ability to learn to reproduce communicative signals from conspecifics. Such an ability is displayed in a limited number of lineages phylogenetically scattered across some groups of mammals (bats, elephants, cetaceans, pinnipeds, and humans) and birds (songbirds, parrots, and hummingbirds) (Petkov and Jarvis 2012; Shen 2017). Among the pieces interlocked within the language mosaic, we have decided to focus on vocal learning here because it is the best understood to date in light of the recent literature (Jarvis and Mello 2000; Jarvis 2004; Chakraborty et al. 2015). As such, it provides the best testing grounds for evolutionary scenarios concerning some important aspects of human language.

The vocal learning literature, especially the line of research pursued by Erich Jarvis and colleagues, already offers interesting scenarios to test. Let us briefly sketch them here, as they will play an important role in the background of the next sections. Vocal learning birds and humans share a number of forebrain structures specialized in song and speech control, respectively (Jarvis 2004). Among them, all three learning avian species

exhibit several brain nuclei that are distributed in two pathways: the anterior, or vocal learning pathway, which is mainly specialized in vocal imitation and malleability, and the posterior, or vocal production pathway, which associates with the intentional production of (learned) vocalizations. Within this posterior pathway, which will be the main focus in the following sections, oscines, parrots, and hummingbirds present three analogous motor regions in the cortex, namely the robust nucleus of the arcopallium (RA), the central nucleus of the anterior arcopallium (AAC), and the vocal nucleus of the arcopallium (VA), respectively, which are in turn analogous to the laryngeal motor cortex (LMC) in humans. In both learning birds and humans, this nucleus makes a direct projection to the brainstem motor neurons (MN) that controls the syrinx in birds and the larynx in humans (Jarvis 2004; Feenders et al. 2008; Pfenning et al. 2014; Simonyan 2014; Chakraborty et al. 2015).

On the basis of such similarities, a motor theory of vocal learning has been proposed (Feenders et al. 2008), arguing that cerebral systems specialized for vocal learning in distantly related lineages are independent evolutions of a motor system inherited from their common ancestor. Analyses in gene expression (Feenders et al. 2008; Shimizu et al. 2010; Wang et al. 2010; Jarvis et al. 2013) certainly point in this direction, further supporting that the posterior pathway, which we will focus on next, must have emerged from a primitive motor system (Feenders et al. 2008; Fitch et al. 2010; Chakraborty et al. 2015). Since several forebrain motor learning pathways with sensory input appear to be formed during early development by successive duplications, thereafter projecting to various brainstem or spinal cord neurons associated with different muscle groups, it has been proposed that the posterior connection appeared similarly as one further duplication that then projected to the brainstem MN in charge of the vocal organs (Fitch 2011; Chakraborty et al. 2015). Pathway duplication unfolds in a manner analogous to gene duplication—with a whole pathway duplicating and the duplicate taking on new function—and actually having gene duplication as one possible underlying mechanism (Chakraborty and Jarvis 2015).

Neuroanatomical research conducted with primates has identified homologous representations of the larynx in the motor cortex (LMC) both in human (Penfield and Boldrey 1937; Rödel et al. 2004) and in nonhuman primates, such as chimpanzees (*Pan troglodytes*) (Leyton and Sherrington 1917), rhesus monkeys (*Macaca mulatta*) (Sugar et al. 1948; Hast et al. 1974), and squirrel monkeys (*Saimiri Sciureus*) (Hast and Milojkvic

1966; Hast et al. 1974). However, although the LMC connectivity network is broadly similar among primates tested, a robust cortico-laryngeal direct projection to the vocal MN in the brainstem has been found only in humans (Simonyan 2014; Belyk and Brown 2017).

There are reasons to believe that the posterior pathway develops gradually, as it is present at a very rudimentary level in the brain of a nonvocal learning suboscine species. Indeed, as Liu et al. (2013) have shown, the eastern phoebe (*Sayornis phoebe*), closely related to songbirds, possesses a specialized forebrain region that seems homologous to the RA in oscines. This region presents descending projections to the brainstem respiratory nucleus and has a singing-associated function. In this regard, eastern phoebes present a long period (8–9 months) of song plasticity before its crystallization. This circuitry seems to be a proto-form of what we find in vocal learning oscines, though not developed enough for vocal learning brain-readiness inasmuch as, unlike in songbirds, there is no direct projection from the arcopallial RA-like nucleus to the tracheosyringeal neurons.

Once this critical neural pathway is established, it is quite likely to undergo further elaborations, giving rise to more complex forms of vocal learning. A case in point that can serve as an example for such specializations can be found in parrots, known to be able to imitate vocalizations not only of conspecifics, but also sounds produced by other species. A study involving the three superfamilies of parrots (*Strigopoidea*, *Cacatuoidea*, and *Psittacoidea*) (Joseph et al. 2012) has revealed an internal subdivision in their song cortical nuclei, wherein a core region shows different gene expression from the surrounding shell area, while both exhibit in turn different expression from the surrounding motor cortical region. Interestingly, the posterior connection to the brainstem MN associated with the syrinx, along with other connections with different forebrain vocal regions, is projected exclusively from the core region and not from the shell (Chakraborty et al. 2015). Chakraborty and Jarvis (2015) suggest that the core region in the parrot AAC evolved convergently in all three avian vocal learning species via duplication from the surrounding motor regions, and subsequently the shell area was developed in parrots, allowing for their more complex vocal proficiency.

As we just saw, critical neural stuctures such as the posterior pathway, taken as a reference point for the origin of the vocal learning capacity, likely emerge in proto-form, and, once present, can be subject to further elaboration, under the influence of several factors.

We believe that the same could be true for the emergence of language in our lineage (Boeckx 2017).

1.2 The *SRGAP2* gene suite and the timing of critical evolutionary steps in *Homo*

Although *SRGAP2* is highly conserved among mammals (Dennis et al. 2012) and has remained unchanged at least in the last 6 million years of our evolution (its F-BARx domain is identical in humans, chimpanzees, bonobos, and orangutans) (Sporny et al. 2017), it has given rise to three human-specific duplications, two of which underwent subsequent mutations. The sequence of events, identified by Dennis et al. (2012), illustrated in Fig. 1, happened as follows (the chronological ranges have been calculated assuming the timing of divergence between chimpanzee and human lineages within a span of 5–7 million years ago (mya), based on fossil records (Brunet et al. 2002; Vignaud et al. 2002; Brunet et al. 2005) and genetic analyses (Patterson et al. 2006): the first duplication took place 2.8–3.9 mya, when the promoter and first nine exons of the original gene—designated *SRGAP2A* to distinguish it from its derivatives—were duplicated from the locus 1q32.1 to 1q21.1, thus giving rise to the primitive *SRGAP2B* (*P-SRGAP2B*). A second duplication occurred 2.0–2.8 mya, when *P-SRGAP2B* was copied from 1q21.1 to 1p12, leading to the primitive *SRGAP2C* (*P-SRGAP2C*). In the aftermath of this event (Dennis et al. 2012; Sporny et al. 2017), the two primitive duplicated copies, *P-SRGAP2B* and *P-SRGAP2C*, accumulated nonsynonymous mutations which resulted in the contemporary *SRGAP2B* and *SRGAP2C* forms, carrying five (R73H, R108W, R205C, R235H, R250Q) and two (R79C, V366L) aminoacid replacements, respectively. Finally, the third and last duplication, which occurred 0.4–1.3 mya, copied the modern *SRGAP2B* within 1q21.1, thus giving rise to *SRGAP2D* (Dennis et al. 2012). Consistent with the timing of their appearances, all three human paralogs, *SRGAP2B*, *SRGAP2C*, and *SRGAP2D*, have been found also in the genomes of Neanderthals and Denisovans (Hillert 2015).

Importantly, the timing of the *SRGAP2* duplications appears to correspond fairly closely to some landmarks in our lineage in terms of brain size and use of stone tools in the transition from *Australopithecus* to *Homo*, raising the possibility that the relevant duplications contributed to these phenotypic changes (Buckner and Krienen 2013; Hillert 2015; Boeckx 2017). Thus, the time of the first duplication (*P-SRGAP2B*) matches the appearance of *Australopithecus*, which had an average brain size of ca. 475 cc, similar to that of genus *Pan*. The second duplication span (*P-SRGAPC*) corresponds
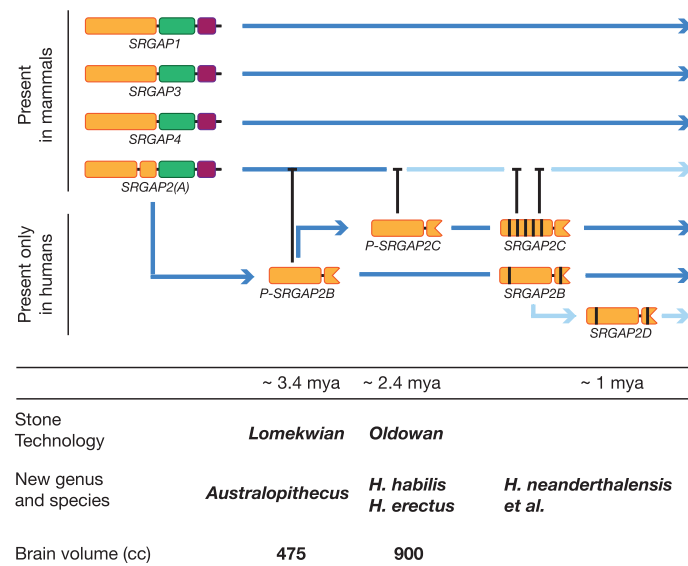
Figure 1. Evolutionary history of SRGAPs and chronological correspondence with human landmarks. On top, the colored figures represent each of the *SRGAP* genes. In orange, the F-BAR domains, with an F-BAR extension in the case of *SRGAP2(A)*. The human duplicate copies are devoid of RhoGAP (green) and SH3 (violet) domains, but conserve the most part of the F-BARx domain. Darker arrows symbolize functional continuity of the gene; the reduced activity of *SRGAP2(A)* by *SRGAP2C*, along reduced activity of *SRGAP2D*, are represented by arrows in a lighter shade. The dates in the central horizontal fringe correspond to the emergence of the primitive form of *SRGAP2B* (*P-SRGAP2B*; ~3.4 mya) and *SRGAP2C* (*P-SRGAP2C*; ~2.4 mya), which parallels the first (Lomekwian) and second (Oldowan) known generations of stone technology. The aminoacid replacements that *P-SRGAP2B* and *P-SRGAP2C* underwent to reach their modern forms (two in *SRGAP2B*, five in *SRGAP2C*) are represented by black bars. Around ~1 mya, *SRGAP2D* emerged as a copy of *SRGAP2B* and carries the same two substitutions. The penultimate row in the figure gives account of the chronological correspondence between the duplication events that led to *P-SRGAP2B* and *P-SRGAP2C*, and the appearances of the genus *Australopithecus* and *Homo* (*H. habilis*; *H. erectus*), respectively; similarly, the appearance of *H. neanderthalensis*, likewise that of other sister *Homo* species, parallels the emergence of *SRGAP2D*. The last row depicts the differences between the estimated brain size of *Australopithecus* (475 cc) and those of *H. habilis* and *H. erectus* (900 cc).

to the appearance of *Homo habilis* and *Homo erectus*, having an average brain size of ca. 900 cc. Finally, the last duplication (*SRGAP2D*) is associated with the emergence of late *H. erectus*, of Neanderthals and of other sister species (Hillert 2015). In addition, the timing of the first and the second duplications, *P-SRGAP2B* (~3.4 mya) and *P-SRGAP2C* (~2.4 mya), shows a fairly close correspondence with the first and second generations of the use of stone tool technology, Lomekwian and Oldowan (Sporny et al. 2017).

In light of claims that total number of neocortical neurons is shown to be a better correlate of cognitive complexity than brain size *per se* (both absolute or relative) (Herculano-Houzel 2012; 2016), it is also interesting to point out that the evolutionary rate of the *SRGAP2* gene has been claimed to positively correlate with an increase in the number of cortical neurons in mammals (Tiwary 2016).

Not surprisingly, several authors suggested that *SRGAP2* duplications may underlie some of the changes that led to human cognition. The most explicit suggestion along these lines that we are aware of was made in (Chakraborty et al. 2015). Building on the existing

literature on the functional effects of the relevant duplications, Chakraborty and Jarvis (2015) write:

The duplicated copies act as competitive inhibitors to slow cortical dendritic development of already existing brain pathways, which in turn allow greater neural plasticity into adulthood. SRGAP2 *modulates activity of the* ROBO *axon guidance receptors, which are in turn activated by the* SLIT *family of protein ligands to modulate axonal/dendritic migration and branching in various brain regions. Intriguingly, the* SLIT1 *ligand is uniquely downregulated in the song production nucleus RA analogue of vocal learning birds (songbird RA, parrot AAC and hummingbird VA) and the analogous human LMC, which would mean that there could be a synergistic effect of the duplicated* SRGAP2 *GTPase and lower* SLIT1 *levels in the duplicated vocal motor pathways in humans.* [references omitted]

We find this suggestion very insightful, and what follows is meant to provide support for it. Doing so requires spelling out some of the assumptions and findings that are alluded to in this quote. We turn to this next.

2. *SRGAP2* genes, filopodia, and axon guidance

The first thing to point out in the context of Chakraborty and Jarvis' suggestion is that the existing literature on *SRGAP2* does not immediately support it. Despite their names (*SRGAP* genes—*SLIT-ROBO GTPase activating protein* coding genes), the nature of the interactions between *SLIT* genes, *ROBO* genes, and *SRGAP* genes does not always go in the desired direction for vocal learning, by which we mean the axon guidance role, for reasons we discuss briefly in the next subsection.

2.1 *SLIT* and *ROBO* axon guidance genes and the vocal learning posterior pathway

As has been said above, a direct neural projection from a cortical/pallial motor nucleus and the brainstem MN controlling the larynx/syrinx appears to be a key component in the evolution of the vocal learning ability. To form this structure during the early development of the brain, the axonal extensions of the neurons in the cortical region must be sent and guided along pathways to eventually reach their synaptic targets in the brainstem through a process which requires the action of axon guidance genes (Dickson 2002).

In this regard, as alluded to in the quote from Chakraborty et al. (2015), studies conducted with birds from the three groups of species of avian vocal learners (Pfenning et al. 2014; Wang et al. 2015) have shown that axon guidance genes of the SLIT-ROBO families present a convergent differential regulation in the pallial motor nucleus of the learning species.

Summarizing briefly these results, we can say that *SLIT1*, a gene belonging to the SLIT family of repulsive axon guidance genes (Dickson 2002), shows a differential downregulation precisely in the songbird RA and in the analog regions in parrots (AAC) and hummingbirds (VA), that is, the arcopallial nuclei making the direct projection to the brainstem MN. The expression of *SLIT1* in these nuclei is remarkably low compared to the surrounding arcopallium. More precisely, in the case of the parrot AAC, which has a subdivision between core and shell we had already expounded, the downregulation of *SLIT1* occurs only in the core region, which is the one sending the projection to the brainstem MN. In contrast, no such regulation of *SLIT1* was observed either in the arcopallium of nonvocal learning birds tested (quails and ring doves) or in a recently discovered putative LMC of mice, thus highlighting the specificity of this expression pattern to vocal learning lineages (Wang et al. 2015). All in all, the particular pattern of expression of *SLIT1* strongly suggests a functional relation between the downregulation of the axon guidance factor and the formation of the neural projection from the cortical nucleus to the brainstem MN in charge for the syrinx, a relation which would be consistent with the similar downregulation of *SLIT1* that has been found in the human LMC (Pfenning et al. 2014).

ROBO1 belongs to the Roundabout (ROBO) family of axon guidance genes, whose encoded proteins act as receptors of SLIT ligands to transduce the repulsive cue into the intracellular domain (Brose et al. 1999; Dickson 2002; Long et al. 2004). Similarly to *SLIT1*, *ROBO1* also shows a differential expression in relation to the posterior pathway: upregulated in the parrot AAC core and in the hummingbird VA, compared to the surrounding arcopallium, whereas in the songbird RA it is downregulated. Despite the divergence in songbirds with respect to the other two groups, *ROBO1* has been observed to be temporarily upregulated in male zebra finches (endowed with a higher capacity for song compared to females) between posthatch days 35 and 65, a period deemed critical for vocal learning (Wang et al. 2015).

2.2 SRGAPs, SLITs, and ROBOs

In mammals, the SRGAP family of genes consists of four members: *SRGAP1*, *SRGAP2*, *SRGAP3*, and the distantly related *SRGAP4* (Pontus Aspenström 2008). The first three were uncovered in 2001 by Wong et al. (2001) in a yeast two-hybrid experiment in which the SRGAPs were found to interact with the C-terminal region of rat *ROBO1*. After their identification, the researchers further analyzed, through different *in vitro* experiments in human embryonic kidney (HEK) cells, various aspects of the interaction between *SRGAP1* and *ROBO1*, including the effect of extracellular *SLIT2* in such binding. Among other results, they found that extracellular *SLIT2* upregulated *ROBO1-SRGAP1* binding in a dose-dependent manner, thus leading to the inactivation of *CDC42*, a member of the Rho GTPase family, which has a well-documented role in the regulation of the cytoskeletal dynamics (Hall 1998). In the light of these findings, the authors proposed that the newly discovered SRGAPs are intracellular effectors in the downstream of a SLIT-ROBO signaling pathway and play a role in the guidance function of SLITs. This approach would make it possible, therefore, for *SRGAP2* to interact with *ROBO1* downstream of an axon guidance cue, which are part of the mechanism leading to the constitution of the aforementioned posterior pathway.

However, and disappointingly for our purposes, subsequent research did not provide support for this initial proposal concerning ROBO1-SRGAP2 binding. Building on the suggestion in Wong et al. (2001), Yao et al. (2008) investigated the SRGAPs messenger RNA (mRNA) expression in rat brain, at various developmental stages and could find only a relative coincidence with the localized *ROBO1* expression reported by other scholars (Marillat et al. 2002; Whitford et al. 2002). A subsequent study (Bacon et al. 2009) on SRGAPs expression in several embryonic and postnatal stages noted similarities of *SRGAP2* pattern with that of *ROBO2*, but did not report any interaction with *ROBO1*. Li et al. focused on the CC3 motif of *ROBO1* that Wong et al. (2001) had found in interaction with the SH3 domain of *SRGAP1*, and then assessed their binding with the SH3 domains of *SRGAP1*, *SRGAP2*, and *SRGAP3* (Li et al. 2006). The result was that most of the recreated peptides did not bind, and only one showed a feeble and transient interaction. Similarly, Okada et al. (2011) did not identify *ROBO1* as a ligand for *SRGAP2*. (Below we return to these unsuccessful attempts, as a recent study (Guez-Haddad et al. 2015) provides a possible reason for these results.)

On a more positive note, SRGAPs, and specifically *SRGAP2* on which we focus here, have been reported to serve various functions regarding cortical development at early stages. First, *SRGAP2* has been shown to regulate axon–dendrite morphogenesis and neuronal migration through its ability to induce protrusions at the plasma membrane. A study of cortical neurons in mice showed that the knockdown of *SRGAP2* significantly decreased both dendritic and axonal branching, while, on the other hand, neurons with short hairpin (shRNA)-silenced expression of *SRGAP2* migrated roughly 25% faster than the control group, thus showing an inhibitory effect (Guerrier et al. 2009). These results support the suggestion in Wong et al. (2001) (based on experiments on *SRGAP1*) that SRGAPs can regulate cell migration. A subsequent study (Charrier et al. 2012) showed the same effects *in vivo*, and demonstrated, in addition, that the expression of *SRGAP2C* in mouse cortical neurons had a similar effect to that caused by bi-ancestral *SRGAP2* knockdown, viz. an increase in the rate of cell migration. In the knockdown condition, Charrier et al. (2012) added another function of *SRGAP2* to those already established: it promotes the maturation of the dendritic spines and limits their density. Indeed, an experiment *in vivo* carried out with heterozygous *SRGAP2*-knockout mice revealed a substantially higher density of dendritic spines in comparison with the control group, with thinner and longer spines. Charrier et al. (2012) also found that the expression of SRGAP2C in mouse pyramidal neurons inhibited the function of *SRGAP2A* and extended the period of development of the spines (spinal 'neoteny'), thus evoking an increase in their number per unit area and in their length. Interestingly, this last trait is considered characteristic of the human neocortex (Benavides-Piccione et al. 2002), and led to claims linking *SRGAP2* duplication with this particular property of the human neocortex.

As a final remark on the function of SRGAPs, we report their ability to co-regulate the ratio between excitatory and inhibitory synapses at their early development to reach the correct equilibrium at the mature stage. A recent *in vivo* study (Fossati et al. 2016) in mouse cortical pyramidal neurons has shown that *SRGAP2A* increases the growth of inhibitory synapses and restricts their density. Curiously, in a way similar to the one mentioned earlier for dendritic spines, *SRGAP2C* antagonizes functions of *SRGAP2A* during synaptic development, prolonging their maturation period and increasing their final density.

As a result, *SRGAP2* duplication has not figured prominently in the literature on the evolution of vocal learning, since to the best of our knowledge neotenous spines are not (yet) considered a central property of vocal learners. Other more established neural traits associated with vocal learning appear not to be directly connected with the role of *SRGAP2*. Nevertheless, in the following sections we show how the well-documented function of *SRGAP2*, namely its ability to regulate protrusions at the plasma membrane of the neuron (Guerrier et al. 2009; Coutinho-Budd et al. 2012; Dominik Fritz et al. 2015; Sporny et al. 2017), can be related to more canonical properties of vocal learning brain-readiness, specifically axon guidance.

2.3 *SRGAP2* and axon guidance: an indirect link

In the process of axon guidance, a series of secreted proteins, such as the SLIT family, act as extracellular biochemical guiding effectors by evoking a signaling cascade that ultimately changes the cytoskeletal dynamics of the axon and directs its outgrowth either toward or away from the signaling source. These directional changes take place at the growth cone, a motile structure located at the distal end of the axon which is endowed with two types of F-actin–based structures: filopodia, which are narrow cylindrical protrusions based in unbranched parallel bundles of actin filaments (F-actin) formed by Ena/VASP and formin proteins, and lamellipodia, sheet-like protrusions based in a network of branched actin which is formed by the Arp2/3 complex. Axon guidance can be understood as a directed,

recurrent process of enlargement and maturation of the growth cone, starting with the formation and extension of filopodia and lamellipodia at its leading edge, through the polymerization of actin filaments, followed by the flow of filopodia along the sides of the growth cone. The final step of the process is their eventual retraction at the base of the growth cone caused by the depolymerization of the F-actin. This last retraction allows the membrane to contract, thus forming a cylindrical consolidated axon shaft (Dickson 2002; Dent and Gertler 2003). Although the mechanisms whereby axons manage to find the correct pathways across the nervous system remain to be fully characterized, the two actin-supported structures that are characteristic of the axon growth cone, filopodia and lamellipodia, are considered to play a crucial role (Dent and Gertler 2003).

In relation to filopodia and axon guidance, a recent study *in vivo* in mouse dorsal root ganglia cells (McConnell et al. 2016) has investigated the dynamics of the growth cone specifically during the axonal repulsion evoked through the SLIT-ROBO molecular pathway. Crucially for us, it has reached an unexpected conclusion: despite the classic view whereby a repulsive signal entails actin depolymerization at the side of the growth cone facing the guidance source, the amino-terminal fragment of *SLIT2* that contains the domain responsible for binding to *ROBO1* and *ROBO2* induced the formation and elongation of actin-based filopodia at the axon growth cone via SLIT-ROBO molecular pathway. Importantly, these SLIT-induced filopodia, which are longer and elongate distinctively toward the sources of the repulsive cue, are indispensable to elicit the guiding signal in the downstream of SLIT-ROBO. We think that these results are essential to understand how *SRGAP2A*, and perhaps some of its human-specific paralogs, can be related to axon guidance (see Fig. 2), thus supporting Chakraborty and Jarvis' (2015) suggestion, and enabling us to provide novel support for the claim that vocal learning was established fairly early in our lineage.

2.4 SRGAP2A and SRGAP2C

SRGAP2A has a singular threefold composition: an F-BAR domain, which has an amino-terminal extension; a RhoGAP domain, and an SH3 domain (Sporny et al. 2017). Remarkably, the extended F-BARx domain allows the protein to explore the geometry of the membrane and to bind selectively to bulging sites or protrusions (Guerrier et al. 2009; Coutinho-Budd et al. 2012; Dominik Fritz et al. 2015). Once in place, *SRGAP2A* can regulate the dynamics of the actin-based

cytoskeleton through its RhoGAP domain, thus evoking different effects in these protrusions. As examples of this, Guerrier et al. (2009) showed that the overexpression of the *SRGAP2A* F-BAR in cortical neurons induced filopodia-like membrane protrusions, whereas Fritz et al. (2015) have shown that it evoked a retraction of the membrane protrusions in a cell–cell overlap context by inactivating local pools of *Rac1* and *CDC42*, which, in turn, caused a breakdown of the actin-supported cytoskeleton and the subsequent retraction. There may be several factors conditioning the specific result of the protrusion regulation that *SRGAP2A* evokes, but, as Fritz et al. (2015) note, one of them must be the upstream input that it receives, most likely from the SLIT-ROBO pathways. In fact, they show that the detected effect of *SRGAP2A* is elicited in the downstream of the *SLIT2-ROBO4* signaling. It is in the context of binding axon guidance molecules that the SH3 domain has shown to be indispensable, although not exclusive, since all three domains (F-BARx, Rho-GAP, and SH3) have been proven to exert a cooperative participation in binding *ROBO1* (Guez-Haddad et al. 2015). As Guez-Haddad et al. (2015) point out, this must be the reason why previous attempts to attest a significant interaction between *ROBO1* and the isolated SH3 domain of *SRGAP2A* (summarized above) had failed. Summing up then, the particular threefold composition of *SRGAP2A* endows it with the ability to regulate membrane protrusions likely in the downstream of the axon guidance SLIT-ROBO pathway.

SRGAP2A molecules are homodimers in solution. Prototypically, F-BAR domains form anti-parallel dimers that bind the plasma membrane through their concave N-surface, thus associating with membrane invaginations. However, the SRGAP2A homodimerization is not only mediated by the F-BAR domain, as typically could be expected, but rather by a large interface that includes the F-BAR, its Fx extensions, the RhoGAP, and the SH3 domains. This particular cooperative dimerization, which additionally increases the ability of the dimer to bind the membrane, evokes an inverted, convex N-surface that associates with protrusions instead of invaginations. The potential of *SRGAP2A* to regulate membrane protrusions likely depends on this particular form of homodimerization (Sporny et al. 2017).

The duplicated copy *SRGAP2C* consists of a truncated form of *SRGAP2A* containing nearly all of the F-BARx with three modifications, two of which occurred in the first duplication event (~3.4 mya), thus being present in the primitive forms, *P-SRGAP2B* and *P-SRGAP2C*. As Sporny et al. (2017) have recently shown, SRGAP2C has the ability to heterodimerize with
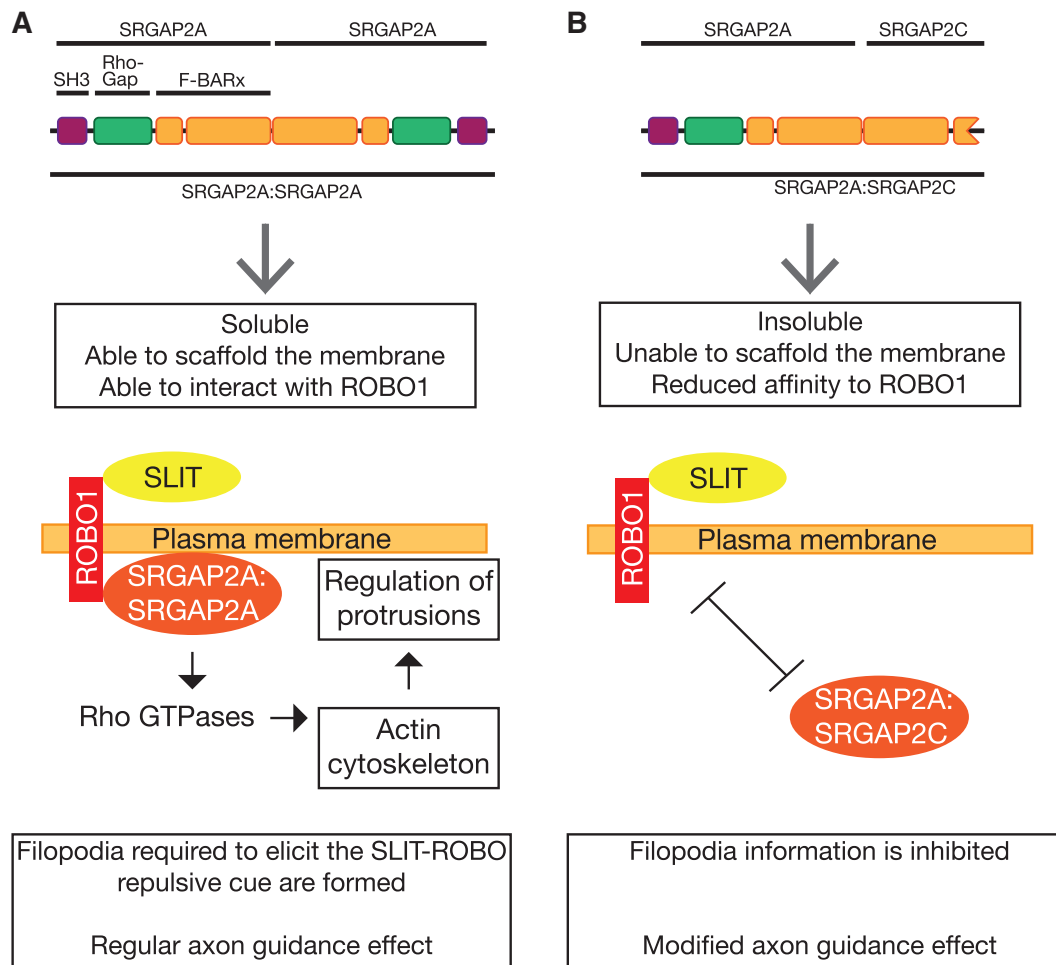
Figure 2. Proposed model for the implication of SRGAP2A and SRGAP2C in an axon guidance signaling pathway. (A) SRGAP2A molecules homodimerize through their F-BARx domains, thus forming soluble dimers. These dimers have a singular inverse geometry which allows them to colocalize at the membrane at sites of protrusions. Once in place, these molecules are able to transduce a SLIT-ROBO axon guidance cue by interacting with Rho GTPases through their RhoGAP domains, thus regulating the actin cytoskeleton and scaffolding protrusions. The chain of interactions leads to the constitution of filopodia which extend toward the sources of SLIT. These filopodia are crucial to elicit the repulsive axon guidance cue. (B) SRGAP2C heterodimerizes with SRGAP2A. The resulting molecule is insoluble, unable to scaffold the membrane, and has a limited affinity for ROBO1. Thus, SRGAP2C inactivates SRGAP2A's ability to regulate filopodia, ultimately resulting in a modified effect in axon guidance.

SRGAP2A, a property which was already present in the primal form *P-SRGAP2C*, which appeared ~2.4 mya. Crucially, unlike SRGAP2A homodimers, SRGAP2A: SRGAP2C heterodimers are insoluble, thus being unable to reach the proper sites in the plasma membrane and consequently being rendered inactive. An experimental quantification of the effect of *P-SRGAP2C* and *SRGAP2C* in compromising *SRGAP2A* solubility has been carried out by Sporny et al. (2017) reflecting that, when coexpressed with recreated P-SRGAP2C and with SRGAP2C in Sf9 cells, 60% and 40% of *SRGAP2A*, respectively were insoluble. In light of these data, it is clear that SRGAP2C acts as an inhibitor of SRGAP2A by cancelling its ability to bind to the membrane and regulate protrusions. Relevantly, this capacity of

SRGAP2C to form stable heterodimers with SRGAP2A and its consequent efficiency at antagonizing the original gene was evolutionarily refined over the mutagenesis phase which took place after the duplication event (about 2.4 mya). In addition, but independently from their insolubility, the SRGAP2A: SRGAP2C heterodimers present a significantly reduced ability to bind *ROBO1* (Sporny et al. 2017).

SRGAP2A mRNA has been shown to be expressed in different regions of the central nervous system at early developmental stages. It was found to be expressed at embryonic and postnatal days in many tissues in mice, including the dorsal and ventral thalamus, the ventrolateral thalamic nucleus, the superior and inferior colliculi, the cerebellum, and the spinal cord (Bacon et al. 2009).

Also, in mice, Guerrier et al. (2009) detected that it follows an increasing pattern of expression during early development in the cortex, reaching its maximum level at postnatal day 1 (P1), then stabilizing until P15, and gradually decreasing although still being expressed in adult stages. Charrier et al. (2012) compare its expression with that of *SRGAP2C* and reach the conclusion that both are expressed in embryonic and adult human brain (though not always in exactly the same way). Various human brain expression databases we consulted generally agree that SRGAPs are expressed in frontal parts of the neocortex early in development (data on SRGAP2C specifically tend to be too sparse to draw any firm conclusion at this point.) The resources we consulted include: Brainspan (http://www.brainspan.org), Human Brain Transcriptome (http://hbatlas.org), Bgee (http: //bgee.org), Proteomics DB (https://proteomicsdb. org), Human Protein Atlas (http://www.proteinatlas. org), Gene Enrichment Profiler (http://xavierlab2.mgh. harvard.edu/EnrichmentProfiler/index.html), and GTex (http://www.gtexportal.org) generally agree that SRGAPs are expressed in frontal parts of the neocortex early in development. (Data on *SRGAP2C* specifically tend to be too sparse to draw any firm conclusion at this point.)

3. Concluding remarks

SRGAP2C may have had other functional consequences (Guerrier et al. 2009; Charrier et al. 2012; Fossati et al. 2016), but we have provided evidence that mechanistically we can expect *SRGAP2C* to have had an effect on the SLIT-ROBO axon guidance pathway, and—no doubt together with other genetic changes—may have contributed to the establishment of a critical aspect of the vocal learning circuit, as first suggested in Chakraborty et al. (2015). We have shown that until very recently studies focusing on *SRGAP2* failed to provide evidence in this direction. It is only thanks to the results in Guez-Haddad et al. (2015) and Sporny et al. (2017) and the link between filopodia and axon guidance made precise in McConnell et al. (2016) that we can adduce a greater degree of plausibility to the claim in Chakraborty et al. (2015) that *SRGAP2* duplications may have contributed to the emergence of aspects of our language faculty, a claim made at a time when the relevant results we rely on had not yet been obtained. Since paleogenomic work has shown that the relevant mutation that led to this effect is not specific to *Homo sapiens*, we are led to conclude that core ingredients of the vocal learning pathway predated the emergence of our species.

In a certain sense, *SRGAP2C* acts like the member of the SRGAP family that most closely interacts with *ROBO1*: *SRGAP1*. Unlike *SRGAP2A*, which as we saw, induces filopodia-like membrane protrusions, *SRGAP1*'s F-BAR domain prevents filopodia (Coutinho-Budd et al. 2012). By inhibiting the ability of *SRGAP2A* to induce filopodia, *SRGAP2C* makes *SRGAP2* function like *SRGAP1*. In light of this, it is noteworthy that a gene expression study (Ip et al. 2011) carried out in human developing neocortical neurons has shown a relation between *ROBO1* and *SRGAP1*. Both genes were found to be co-expressed in human corticospinal axons at various fetal periods during the formation of the corticospinal tract, which is the main descending sensorimotor projection, an elaboration of which could have given rise to the critical connection of the posterior vocal learning circuit.

As pointed out in Wang et al. (2015), *SLIT1* is a direct target of *FOXP2* (Vernes et al. 2007; Konopka et al. 2009). Although human *FOXP2* has been reported to modulate stronger upregulation of *SLIT1* than chimpanzee *FOXP2* (Konopka et al. 2009), which does not fit well with the relevant convergent downregulation of *SLIT1* in vocal learning birds found in Wang et al. (2015), *SLIT1* is among the *FOXP2* targets found to be significantly downregulated in response to *FOXP2* expression in Devanna et al. (2014). So, there could be another synergistic effect here between the effect of *FOXP2* on *SLIT1* and the action of *SRGAP2C* on the SLIT-ROBO pathway.

Incidentally, just like *SRGAP2C* works its effect on the SLIT-ROBO pathway by inhibiting an inhibitor (in this case, *SRGAP2A*), *FOXP2* also appears to work its effects by inhibiting inhibitors, such as *MEF2C*. As reported in Chen et al. (2016), (mouse) *Foxp2* controls synaptic wiring of corticostriatal circuits, critical for vocal learning, by opposing *Mef2c*, which itself suppresses corticostriatal synapse formation and striatal spinogenesis. So, achieving a positive effect (establishment of a vocal learning circuit) by inhibiting inhibitors or suppressing the activity of suppressors, appears to have been a common strategy in the evolution of our lineage and our cognitive phenotype.

We still do not know exactly when the relevant *FOXP2* mutations emerged in our lineage, so we cannot know for sure if the emergence of modern *SRGAP2C* coincided with the two *FOXP2* mutations thought to be critical for vocal learning. Evidence for a selective sweep associated with *FOXP2* yields ambiguous results (assuming that the relevant mutations were the actual selection targets): there is not only evidence for a recent *H. sapiens*-specific partial selective sweep (Maricic et al. 2013; Mallick et al. 2016), but also evidence for another, much earlier sweep (Mallick et al. 2016; Supplementary Table S12.1).

It remains to be seen if these sweeps correspond to landmarks in the establishment of the human vocal learning circuit, possibly corresponding to the stages that can be derived from the work on vocal learning birds (e.g., suboscine/proto-vocal learning stage (Liu et al. 2013), core vocal learning circuit stage (Wang et al. 2015), shell vocal learning circuit stage (Chakraborty et al. 2015)).

Though modest, we think that our contribution is of a kind that is necessary to make claims about when components of our language faculty mosaic emerged. It is not enough to simply identify changes on potentially relevant genes. It is necessary to show that the changes have functional effects of the right kind. We hope to have taken a small step in this direction.

Authors' contributions

C.B. formulated the hypothesis and directed the study. P.T.M., M.M., and C.B. reviewed the literature, and wrote the article.

Conflict of interest statement. None declared.

Funding

References

Bacon, C., Endris, V., and Rappold, G. (2009) 'Dynamic Expression of the Slit-Robo GTPase Activating Protein Genes during Development of the Murine Nervous System', *Journal of Comparative Neurology*, 513/2: 224–36.

Belyk, M., and Brown, S. (2017) 'The Origins of the Vocal Brain in Humans', *Neuroscience & Biobehavioral Reviews*, 77: 177–93.

Benavides-Piccione, R. et al. (2002) 'Cortical Area and Species Differences in Dendritic Spine Morphology', *Journal of Neurocytology*, 31/3: 337–46.

Bickerton D. (2002) *From Protolanguage to Language: The Speciation of Modern Homo sapiens*. Oxford: Oxford University Press.

Boeckx, C. (2013) 'Biolinguistics: Forays into Human Cognitive Biology', *Journal of Anthropological Sciences*, 91: 63–89.

—— (2017) 'Language evolution', in Kaas J. (ed.) *Evolution of Nervous Systems*, 2nd edn., vol. 4, pp. 325–39. London: Elsevier.

Brose, K. et al. (1999) 'Slit Proteins Bind Robo Receptors and Have an Evolutionarily Conserved Role in Repulsive Axon Guidance', *Cell*, 96/6: 795–806.

Brunet, M. et al. (2002) 'A New Hominid from the Upper Miocene of Chad, Central Africa', *Nature*, 418/6894: 145–51.

—— et al. (2005) 'New Material of the Earliest Hominid from the Upper Miocene of Chad', *Nature*, 434/7034: 752–5.

Buckner, R. L., and Krienen, F. M. (2013) 'The Evolution of Distributed Association Networks in the Human Brain', *Trends in Cognitive Sciences*, 17/12: 648–65.

Chakraborty, M. et al. (2015) 'Core and Shell Song Systems Unique to the Parrot Brain', *PLoS One*, 10/6: e0118496.

——, and Jarvis, E. D. (2015) 'Brain Evolution by Brain Pathway Duplication', *Philosophical Transactions of the Royal Society B*, 370/1684: 20150056.

Charrier, C. et al. (2012) 'Inhibition of SRGAP2 Function by Its Human-specific Paralogs Induces Neoteny during Spine Maturation', *Cell*, 149/4: 923–35.

Chomsky N. (2010) 'Some Simple Evo devo Theses: How True Might They Be for Language?' in *The Evolution of Human Language: Biolinguistic Perspectives*. Cambridge University Press.

Coutinho-Budd, J. et al. (2012) 'The F-BAR Domains from srGAP1, srGAP2 and srGAP3 Regulate Membrane Deformation Differently', *Journal of Cell Science*, 125/14: 3390–401

Dediu, D., and Levinson, S. C. (2013) 'On the Antiquity of Language: The Reinterpretation of Neandertal Linguistic Capacities and Its Consequences', *Frontiers in Psychology*, 4: 397.

Dennis, M. Y. et al. (2012) 'Evolution of Human-specific Neural SRGAP2 Genes by Incomplete Segmental Duplication', *Cell*, 149/4: 912–22.

Dent, E. W., and Gertler, F. B. (2003) 'Cytoskeletal Dynamics and Transport in Growth Cone Motility and Axon Guidance', *Neuron*, 40/2: 209–27

Devanna, P., Middelbeek, J., and Vernes, S. C. (2014) 'FOXP2 Drives Neuronal Differentiation by Interacting with Retinoic Acid Signaling Pathways', *Frontiers in Cellular Neuroscience*, 8: 305.

Dickson, B. J. (2002) 'Molecular Mechanisms of Axon Guidance', *Science*, 298/5600: 1959–64

Dominik Fritz, R. et al. (2015) 'SrGAP2-dependent Integration of Membrane Geometry and Slit-Robo-repulsive Cues Regulates Fibroblast Contact Inhibition of Locomotion', *Developmental Cell*, 35/1: 78–92

Enard, W. et al. (2009) 'A Humanized Version of Foxp2 Affects Cortico-basal Ganglia Circuits in Mice', *Cell*, 137/5: 961–71.

Feenders, G. et al. (2008) 'Molecular Mapping of Movement-associated Areas in the Avian Brain: A Motor Theory for Vocal Learning Origin', *PLoS One*, 3/3: e1768

Fitch, W. T. (2011) 'The Evolution of Syntax: An Exaptationist Perspective', *Frontiers in Evolutionary Neuroscience*, 3/9: 1–12.

—— (2017) 'Empirical Approaches to the Study of Language Evolution', *Psychonomic Bulletin & Review*, 24/1: 3–33

—— (2017) Special Issue on the Biology and Evolution of Language, *Psychonomic Bulletin & Review*, 24/1.

——, Huber, L., and Bugnyar, T. (2010) 'Social Cognition and the Evolution of Language: Constructing Cognitive Phylogenies', *Neuron*, 65/6: 795–814

Fossati, M. et al. (2016) 'SRGAP2 and its Human-specific Paralog Co-regulate the Development of Excitatory and Inhibitory Synapses', *Neuron*, 91/2: 356–69

Guerrier, S. et al. (2009) 'The F-BAR Domain of srGAP2 Induces Membrane Protrusions Required for Neuronal Migration and Morphogenesis', *Cell*, 138/5: 990–1004.

Guez-Haddad, J. et al. (2015) 'The Neuronal Migration Factor srGAP2 Achieves Specificity in ligand binding through a two-component molecular mechanism', *Structure*, 23/11: 1989–2000.

Hall, A. (1998) 'Rho GTPases and the Actin Cytoskeleton', *Science*, 279/5350: 509–14

Hast, M. H. et al. (1974) 'Cortical Motor Representation of the Laryngeal Muscles in *Macaca mulatta*', *Brain Research*, 73/2: 229–40.

——, and Milojkvic, R. (1966) 'The Response of the Vocal Folds to Electrical Stimulation of the Inferior Frontal Cortex of the Squirrel Monkey', *Acta Oto-Laryngologica*, 61/1–6: 196–204.

Herculano-Houzel, S. (2012) 'The Remarkable, Yet Not Extraordinary, Human Brain as a Scaled-up Primate Brain and Its Associated Cost', *Proceedings of the National Academy of Sciences USA*, 109(Suppl 1): 10661–8.

—— (2016) *The Human Advantage*. Cambridge, MA: MIT Press.

Hillert D. (2015) *The Premodern Language-ready Brain of our Common Ancestor*. San Diego: University of California.

Hillert, D. G. (2015) 'On the Evolving Biology of Language', *Frontiers in Psychology*, 6: 1796.

Ip, B. K. et al. (2011) 'The Corticofugal Neuron-associated Genes ROBO1, SRGAP1, and CTIP2 Exhibit an Anterior to Posterior Gradient of Expression in Early Fetal Human Neocortex Development', *Cerebral Cortex*, 21/6: 1395–407.

Jarvis, E. D. (2004) 'Learned Birdsong and the Neurobiology of Human Language', *Annals of the New York Academy of Sciences*, 1016/1: 749–77.

—— et al. (2013) 'Global View of the Functional Molecular Organization of the Avian Cerebrum: Mirror Images and Functional Columns', *Journal of Comparative Neurology*, 521/16: 3614–65.

——, and Mello, C. V. (2000) 'Molecular Mapping of Brain Areas Involved in Parrot Vocal Communication', *Journal of Comparative Neurology*, 419/1: 1–31.

Joseph, L. et al. (2012) 'A Revised Nomenclature and Classification for Family-group Taxa of Parrots (psittaciformes)', *Zootaxa*, 3205/2: 26–40.

Konopka, G. et al. (2009) 'Human-specific Transcriptional Regulation of CNS Development Genes by FOXP2', *Nature*, 462/7270: 213–17.

Krause, J. et al. (2007) 'The Derived *foxp2* Variant of Modern Humans was Shared with Neandertals', *Current Biology*, 17/21: 1908–12.

Lai, C. S. L. et al. (2001) 'A Forkhead-domain Gene is Mutated in a Severe Speech and Language Disorder', *Nature*, 413/6855: 519–23.

Leyton, A. S. F., and Sherrington, C. S. (1917) 'Observations on the Excitable Cortex of the Chimpanzee, Orang-utan, and Gorilla', *Experimental Physiology*, 11/2: 135–222

Li, X. et al. (2006) 'Structural Basis of Robo Proline-rich Motif Recognition by the srGAP1 Src Homology 3 Domain in the Slit-Robo Signaling Pathway', *Journal of Biological Chemistry*, 281/38: 28430–7.

Liu, W-c. et al. (2013) 'Rudimentary Substrates for Vocal Learning in a Suboscine', *Nature Communications*, 4: 1–12.

Long, H. et al. (2004) 'Conserved Roles for Slit and Robo Proteins in Midline Commissural Axon Guidance', *Neuron*, 42/2: 213–23.

Mallick, S. et al. (2016) 'The Simons Genome Diversity Project: 300 Genomes from 142 Diverse Populations', *Nature*, 538/7624: 201–6.

Maricic, T. et al. (2013) 'A Recent Evolutionary Change Affects a Regulatory Element in the Human FOXP2 Gene', *Molecular Biology and Evolution*, 30/4: 844–52.

Marillat, V. et al. (2002) 'Spatiotemporal Expression Patterns of *slit* and *robo* Genes in the Rat Brain', *Journal of Comparative Neurology*, 442/2: 130–55

McConnell, R. E. et al. (2016) 'A Requirement for Filopodia Extension Toward Slit during Robo-mediated Axon Repulsion', *Journal of Cell Biology*, 213/2: 261–74.

Mendez, F. L. et al. (2016) 'The Divergence of Neandertal and Modern Human Y Chromosomes', *The American Journal of Human Genetics*, 98/4: 728–34.

Mithen, S. (2005) *The Singing Neanderthals: The Origin of Language, Music, Mind and Body*. London: Weidenfeld and Nicolson.

Okada, H. et al. (2011) 'SH3 Domain–Based Phototrapping in Living Cells Reveals Rho Family GAP Signaling Complexes', *Science Signaling*, 4/201: 1–17.

Pääbo, S. (2014) *Neanderthal Man: In Search of Lost Genomes*. New York, NY: Basic Books.

Patterson, N. et al. (2006) 'Genetic Evidence for Complex Speciation of Humans and Chimpanzees', *Nature*, 441/7097: 1103–8.

Penfield, W., and Boldrey, E. (1937) 'Somatic Motor and Sensory Representation in the Cerebral Cortex of Man as Studied by Electrical Stimulation', *Brain: A Journal of Neurology*, 60/4: 389–443.

Petkov, C. I., and Jarvis, E. (2012) 'Birds, Primates, and Spoken Language Origins: Behavioral Phenotypes and Neurobiological Substrates', *Frontiers in Evolutionary Neuroscience*, 4: 1–24.

Pfenning, A. R. et al. (2014) 'Convergent Transcriptional Specializations in the Brains of Humans and Song-learning Birds', *Science*, 346/6215: 1256846.

Pontus Aspenström (2008) 'Roles of f-bar/pch Proteins in the Regulation of Membrane Dynamics and Actin Reorganization', *International Review of Cell and Molecular Biology*, 272: 1–31.

Rödel, R. M. W. et al. (2004) 'Human Cortical Motor Representation of the Larynx as Assessed by Transcranial Magnetic Stimulation (TMS)', *The Laryngoscope*, 114/5: 918–22.

56

Schreiweis, C. et al. (2014) 'Humanized Foxp2 Accelerates Learning by Enhancing Transitions from Declarative to Procedural Performance', *Proceedings of the National Academy of Sciences USA*, 111/39: 14253–8.

Shen, H. H. (2017) 'News Feature: Singing in the Brain', *Proceedings of the National Academy of Sciences USA*, 114/36: 9490–3.

Shimizu, T., Patton, T. B., and Husband, S. A. (2010) 'Avian Visual Behavior and the Organization of the Telencephalon', *Brain, Behavior and Evolution*, 75/3: 204–17.

Simonyan, K. (2014) 'The Laryngeal Motor Cortex: Its Organization and Connectivity', *Current Opinion in Neurobiology*, 28: 15–21.

Sporny, M. et al. (2017) 'Structural History of Human SRGAP2 Proteins', *Molecular Biology and Evolution*, 34: 1463–78.

Sugar, O., Chusid, J. G., and French, J. O. H. N. D. (1948) 'A Second Motor Cortex in the Monkey (*Macaca mulatta*)', *Journal of Neuropathology and Experimental Neurology*, 7/2: 182–9.

Tiwary, B. K. (2016) 'Evolution of the SRGAP2 Gene is Linked to Intelligence in Mammals', *Biomedicine Hub*, 1/1: 443947–443947

Vernes, S. C. et al. (2007) 'High-throughput Analysis of Promoter Occupancy Reveals Direct Neural Targets of FOXP2, a Gene Mutated in Speech and Language Disorders', *The American Journal of Human Genetics*, 81/6: 1232–50.

Vignaud, P. et al. (2002) 'Geology and Palaeontology of the Upper Miocene Toros-Menalla Hominid Locality, Chad', *Nature*, 418/6894: 152–5.

Wang, R. et al. (2015) 'Convergent Differential Regulation of SLIT-ROBO Axon Guidance Genes in the Brains of Vocal Learners', *Journal of Comparative Neurology*, 523/6: 892–906

Wang, Y., Brzozowska-Prechtl, A., and Karten, H. J. (2010) 'Laminar and Columnar Auditory Cortex in Avian Brain', *Proceedings of the National Academy of Sciences USA*, 107/28: 12676–81

West-Eberhard, M. J. (2003) *Developmental Plasticity and Evolution*. Oxford: Oxford University Press.

Whitford, K. L. et al. (2002) 'Regulation of cortical dendrite development by Slit-Robo interactions', *Neuron*, 33/1: 47–61.

Wong, K. et al. (2001) 'Signal Transduction in Neuronal Migration: Roles of GTPase Activating Proteins and the Small GTPase Cdc42 in the Slit-Robo Pathway', *Cell*, 107/2: 209–21.

Yao, Q. et al. (2008) 'Regulated Shuttling of Slit-Robo-GTPase Activating Proteins between Nucleus and Cytoplasm during Brain Development', *Cellular and Molecular Neurobiology*, 28/2: 205–21.

Chen, Y. C. et al. (2016) 'Foxp2 Controls Synaptic Wiring of Corticostriatal Circuits and Vocal Communication by Opposing Mef2c', *Nature Neuroscience*, 19: 1513–22.

Appendices

Appendix A

What we talk about when we talk about biolinguistics

Pedro Tiago Martins* and Cedric Boeckx

What we talk about when we talk about biolinguistics

Abstract: The study of the biological foundations of language is sometimes called *biolinguistics*. This particular term finds its historical origins in the 1950s, and for various reasons it has also gained considerable traction in recent years. While its increasing use apparently signals an equally increasing interest in biology, apart from a few exceptions not much is added to and beyond standard linguistic theorizing by those linguists who use it, resulting in a complex and confusing literature. This state of affairs has led, on the one hand, to the perpetuation of proposals that are hard to relate to the biological literature and, on the other, to ill-placed criticism on the progress and even the very legitimacy of a biologically-informed study of language. By reviewing different ways in which research under the biolinguistics label has been carried out, as well as some common criticisms, we hope to dispel some misconceptions about what constitutes a biolinguistic approach, as well as point out what we contend is real progress in the study of the biological bases and evolution of the human language faculty, to which the term is better and rightly applied.

Keywords: biolinguistics, generative grammar, interdisciplinarity, language faculty

1 Introduction

The study of the biological foundations of language is sometimes called "biolinguistics", a term that has gained considerable traction in recent years. Those who agree that something like a language faculty exists are inclined to use it as a means of emphasizing that their object of inquiry is ultimately a biological one. There are, however, two different factions that prominently adopted the term:

  i)   those who use it as a rebranding of theoretical linguistics of the generativist persuasion
  ii) those who use it as a departure from the common practices of theoretical linguistics, firmly oriented towards biology

For those in i), this divide might not exist,. Its members will believe that there is no difference between generative linguistics and studying the biological nature of language, and perhaps rely on the Chomskyan origin of the term when holding that position. For those in ii), the divide is clear, and they believe that even though points of commonality can be found, they are scarce when compared to the points of departure. From an outsider's point of view, who may or may not be terribly interested in the discussion, this divide might also not exist, and all uses of the term "biolinguistics" are lumped together under the first kind. Similarly, more often than not those who call themselves "biolinguists" are thought to have roughly similar conceptions concerning (i) what the faculty of language is, (ii) what degree of biological uniqueness it has, (iii) what the primitives that constitute it look like, (iv) the extent to which real engagement with other fields, especially biology, is needed, and even (v) what work they are supposed to oppose or not care for. This impression, which we will show is not

---

*Corresponding author: Pedro Tiago Martins, Section of General Linguistics, Universitat de Barcelona, Barcelona, Spain,
E-mail: ptsgmartins@gmail.com
Cedric Boeckx, Section of General Linguistics, Universitat de Barcelona, Barcelona, Spain; ICREA, Barcelona, Spain; UB Institute
for Complex Systems, Barcelona Spain, E-mail: cedric.boeckx@ub.edu

totally correct, can perhaps be attributed to the strong relation between the histories of biolinguistics and of generative linguistics, which many – opponents and advocates alike – indeed see as indissociable.[1]

Biolinguistics, also called the "biolinguistic enterprise" or "biolinguistic approach", has its origins in the work of Noam Chomsky and Eric Lenneberg, among a few others, beginning in the 1950s, as a reaction to the behaviorist paradigm that at the time dominated psychology and linguistics. Chomsky's *Syntactic Structures* (1957) and his review of B. F. Skinner's *Verbal Behavior* (1959) arguably put an end to behaviorism, emphasizing the importance and necessity of internal mechanisms – what one could call cognition – in the acquisition and use of human language. Chomsky further developed important concepts in his *Aspects of the Theory of Syntax* (1965), at the same time that Lenneberg was drawing evidence from a host of domains and fields for a biological basis for human language, which he put together in his *Biological Foundations of Language* (1967). These works are stepping stones in the shaping of biolinguistic thought, in what was in effect a change of paradigm in the study of language, though in embryonic form. Interdisciplinary meetings were later organized with the intention of further exploring and developing what seemed like an emergent discipline, and eventually the term "biolinguistics" was coined in 1974, by Massimo Piattelli-Palmarini. However, after these meetings both the term and the interest in interdisciplinary dialogue quickly faded away, giving way to sophisticated theories of linguistic structure, and remaining in obscurity for decades. Around the turn of the century, for several reasons the interest in the biological foundations resurfaced. These reasons have to do mostly with advancements in the biological sciences and a push for simpler machinery in linguistics and cognitive science (for details, see Boeckx 2013).

As can be witnessed by the volume of publications and meetings bearing the name, biolinguistics seems to have attained unprecedented support, apparently signaling increasing interest in biology as a crucial gateway for the unveiling of the nature of language. We argue that this is in fact not the case, and that most uses of the term "biolinguistics" and most research under the biolinguistics label – including in

---

1 It would not be fair to say that all linguists who have adopted the term at some point are generativist (e. g. Givón 2002, though this author only uses the term as the title of a book, and does not discuss or even mention it in the work itself).

It would also not be fair to say that all generative linguists are sympathetic to biolinguistics as a term or as what they feel it represents. Generative linguist Gillian Ramchand says the following in her statement for a recently held round table on the future of generative syntax:

> There are parts of 'generative' grammar that I do not feel a particularly strong part of. For example, I am not sympathetic to recent trends in Biolinguistics, which to my mind is guilty of extreme Overreach in attempting to connect linguistics to Biology. I think it gives the whole field a bad name. The granularity gap and the terminology gap (to put it in Poeppel's terms) are still too great to sustain the specific kinds of proposals that are being taken seriously in this sub-group.
>
> (Ramchand 2015)

We cannot determine the kind of biolinguistics Ramchand has in mind – as we intend to show in the present paper, the literature is quite confusing –, but it seems it might be that of i), since she considers biolinguistics to be a 'sub-group' of generative linguistics, and most likely is more familiar with 'biolinguistic' proposals made those who use term as just another name for the kind of linguistic theorizing that has been common practice for decades. Also, one of the points of departure of group ii) is precisely to take Poeppel's problems (see Poeppel and Embick 2005; Embick and Poeppel 2015) seriously and work towards plausible bridging theories, so that the gaps become less and less great. It must be someone's task to diminish the ontological and methodological gaps that plague interdisciplinary research, and use them not as cop-outs, but instead as research directions. A good biolinguistic will attempt to contribute to that task. We do not know what Ramchand refers to when she mentions a "terminology gap", which she attributes to Poeppel, which we could not find anywhere in his work. We think it could be a sign of the lack of interest in the problems Poeppel brings attention to and thus in the importance of interdisciplinarity. Anyone familiar with Poeppel's reflections will quickly recognize that terminology is *not* the issue. Poeppel's problems will not be solved by a good dictionary. It may be worth noting in the context of Ramchand's assessment that a similar, pessimistic statement linguists often make – and it is far too common for us to pinpoint a particular case – is that we don't know much about the brain or its relation to cognition in general. This has been said for decades now, and while it still holds true, the amount we know is far vaster than what we used to know. Again, it is up to those interested in language as a biological object of study to make the gap smaller.

venues dedicated exclusively to 'biolinguistic' work – ignore biology for the most part and add little to no substance to linguistic theorizing. We contend its use has done more service to the perpetuation of proposals that are hard to relate to the biological literature than to the rapprochement of linguistics to the biological sciences, in effect impeding the kind of interdisciplinarity that a biology of language demands if it is to be successful.

This is no fault of the term "biolinguistics" – it's hard to think of a more appropriate one if the goal is to study language in light of biology – but rather of how confusingly it has been applied. Concomitantly, most criticisms of biolinguistics or even of the very legitimacy of a biologically-informed study of language are made with unfortunate, albeit more frequent, senses of biolinguistics in mind. This results in the repetition of known debates in the history of linguistics or in theory internal disputes that do not really concern or directly bear on the biological study of language, the latter being the only sense in which the term "biolinguistics" should in our view be used, and to which we will return and clarify in § 3.

The overwhelming occurrence of what we see as less fruitful uses of the term compared to what we defend as the more fecund kind of biolinguistics would lead one to conclude that ours is simply a competing usage, no better deserving than the other usages of the term. However, we believe it's important to rescue the term for what it really means – the biological study of language –, rather than concede that widespread misuse has rendered "biolinguistics" an accepted synonym of enterprises that already bear other names. We see no point in bringing back the term if no changes in the study of language come along. The assumption that our language capacity hinges on our biology is not enough for the 'bio-' prefix to be used. If it were, by the same token all fields – except perhaps mathematics – should bear a 'physico-' prefix, since ultimately everything in the natural world falls out from physics. Thus, we see our discussion not merely as one of terminology, but as a call for bringing biological issues and perspectives to the fore, where they can be addressed and tested. In the case of those linguists who are not interested in biological issues, certainly far greater in number than those who are interested, we fail to see the point of insisting on such a biologically charged term.

2 Different senses of biolinguistics

It is symptomatic of 'biolinguistic' discussions that biological issues and literature are rarely dealt with. When biology is even mentioned, it is usually through metaphors or quoted passages from linguists who use biological jargon, or rephrasings thereof. It is therefore no surprise to find that most support and criticism of biolinguistics does not bear at all on issues that serious biolinguists care about or work on. Again, just the assumption that language has biological underpinnings is not enough for biolinguistics; the main goal is trying to understand what these are.

When they criticize biolinguistics, most authors are thus actually criticizing the implausibility of certain biological proposals in linguistics, using "biolinguistics" in the same way the authors of said proposals do. This is not to say that the biolinguistics we think worthy of the name is immune to criticism; like any other approach, it is obviously not.

In what follows we will give some examples of what we see as unfortunate senses of the term 'biolinguistics' in the literature (summarized in Table 1). We cannot hope to provide an exhaustive list, but we think that the following offers a representative picture.[2]

---

2 We will exclude the earliest uses of the term, which precede and are not related to the earliest work of Chomsky and Lenneberg. Its earlier use, as far as we have been able to determine, goes back to at least the 1920s, when it appeared in catalogs and bulletins of the University of Michigan and also in a book on the psychology on language (Pillsbury and Meader 1928). In fact, during this time a "Laboratory of Biolinguistics" was created at that university, devoted to the study of "various forms of articulation, intonation, intensity, pitch, quantity, and the other factors which contribute to the effectiveness of discourse." (QJS 1925).It later and more notably appeared as the main topic of a handbook (Meader and Muyskens 1950), which came to be the first major publication bearing "biolinguistics" in its title. Meader and Muyskens (1950) already contains ideas in favor of the

Table 1: Different senses of "biolinguistics" in the literature.

Sense	Description
biolinguistics as generative linguistics (§ 2.1)	biolinguistics is synonymous to generative linguistics
biolinguistics as minimalism (§ 2.2)	biolinguistics is synonymous to minimalist accounts of language
biolinguistics as the genetics of language (§ 2.3)	biolinguistics assumes that (at least some) linguistic properties are encoded in the genes, which in turn it must discover
biolinguistics as the study of the uniquely human and linguistic (§ 2.4)	biolinguistics assumes there is something biologically unique to language and humans
biolinguistics as analogy to biology (§ 2.5)	biolinguistics is the description of linguistic processes or mechanisms through perceived similarities with or resemblance to biological ones
biolinguistics as a Chomskyan enterprise (§ 2.6)	biolinguistics is whatever represents Noam Chomsky's views at any given moment

We also do not claim that any one author keeps to the same sense across his or her work. It does vary, be it because of inconsistency or development of one's own conception. For example, in what we see as a particularly confusing case, the following passage seems to encapsulate most of these senses:

> In the prevalent intellectual scenario, it is of considerable interest that the contemporary discipline of generative linguistics – also called "biolinguistics" – has raised the prospects for developing a form of inquiry achieved only in some of the basic sciences. Biolinguistics is arguably the only attempt in the history of ideas in which, according to Noam Chomsky, the study of an aspect of the human mind – language – is beginning to have the "feel of scientific inquiry." Biolinguistics is currently suggesting that structure of language may be 'perfect" in design, not unlike the arrangement of petals in the sunflower and the double helix of the DNA. Yet these advances have been accomplished essentially independently of the natural sciences, especially biology. In that sense, biolinguistics has initiated a (basic) science in its own terms.
>
> (Mukherji 2010: xv)

Our goal is to highlight that "nothing 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" (Lenneberg 1964).

2.1 Biolinguistics as generative linguistics

Many linguists who assume that the language faculty exists see their work as essentially biolinguistic: they see no difference whatsoever between generative linguistics and biolinguistics, and sometimes use the latter as a means of emphasizing that position.[3]

For a somewhat surprising instance, witness the glossary entry for "Biolinguistics" in Chomsky's 2012 *Science of Language*, presumably attributable to the editor of the volume:

> [Biolinguistics:] The current name for the internalist and naturalistic study of language undertaken by Chomsky and others. Earlier descriptions for this methodology include "Cartesian Linguistics", "naturalistic approach" and "methodological monism". As these names suggest, those who adopt this methodology assume that language is a system in the head that is

---

dialogue between linguistics and biology, but the overall assumptions and focus differ quite markedly from the biolinguistics that started later, with the so-called Cognitive Revolution of the 1950s, and as such it bears no historical relation to the current uses. We do however concede that this earlier, pre-1950s use of the term might also be a source of confusion for those who stumble upon it in search of clarification.

3 The first author has conducted several, as-yet-unpublished interviews with prominent researchers on biolinguistic topics. In one such instance, a famous generative linguist confessed to seeing no difference between 'generative linguistics' and 'biolinguistics', remarking that they are "one and the same", and questioning whether "there [were] such people who did not think they were."

innate in some sense (grows/develops as other mental systems do) and is to be studied in the same ways as any other phenomenon, according to the usual desiderata for naturalistic scientific research. The assumptions appear to be reasonable ones: they and the methodology yield good theories of the language faculty.

Surely, all of these "earlier descriptions" may share some assumptions, and they certainly do, but assumptions are not enough to distinguish between fields or approaches; they are just a starting point, and arguably the only part of any field that can be left untouched anyway. The claim that "language is a system in the head that is innate in some sense [...]" is not enough to equate all approaches that consider it to be an axiom. We all know – or at least find it extremely reasonable – that the language faculty is biological, but if we do not address its biological foundations we are not, in effect, doing anything bio-. The bar cannot be set as low as simply assuming a biological character somewhere down the line.

2.2 Biolinguistics as minimalism

Speaking on the distinction between the 'weak' and 'strong' senses of biolinguistics (after Boeckx and Grohmann 2007), Derek Bickerton says:

> It is difficult to avoid the conclusion that adhering to the latest version of generative grammar [=minimalism] is indeed a prerequisite, not perhaps for simply attempting to engage in biolinguistics, but certainly for being taken seriously by serious biolinguists.
>
> (Bickerton 2014: 73)

In order to make the point that Bickerton's conclusion is misguided, we shall first reproduce here the original distinction:

> *The weak sense of the term* refers to "business as usual" for linguists, so to speak, to the extent they are seriously engaged in discovering the properties of grammar, in effect carrying out the research program Chomsky initiated in Syntactic Structures.
>
> *The strong sense of the term 'biolinguistics'* refers to attempts to provide explicit answers to questions that necessarily require the combination of linguistic insights and insights from related disciplines (evolutionary biology, genetics, neurology, psychology, etc.).
>
> (Boeckx and Grohmann 2007: 2, our emphasis)

We believe this distinction should no longer be maintained. The reason is that for linguists to be seriously engaged in discovering the properties of grammar they must combine linguistic insights with insights from other fields, and there is no way around it. There are of course different ways of going about this task, but any route will require dealing with different disciplines. This 'weak sense' of biolinguistics, inclusive as it was, now almost 10 years ago, has not yielded any findings that could not have been reached by a linguistics without mention of biology. The 'weak sense' was perhaps useful at the onset of the biolinguistics revival (and also of the Biolinguistics journal, where it was introduced) as a way of prompting reflection by those who understandably do not want to get their hands 'wet', but still think biology is worth considering when coming up with theories of language. This has not happened, as far as we can tell, as linguistic theories and the entities they posit are still of the same, *sui generis* kind. In other words, the "extent [to which linguists] are seriously engaged in discovering the properties of grammar" has been null in the context of biology. The biggest motivation behind most linguistic theorizing is still to account for attested linguistic variation, and to do so in such a way that the attested range fits. This can be done completely independently from biology, and in fact it is. To actually understand how the child acquires a language, we need to understand what goes on in the child's brain and how it got there, and not just derive an "explanation" from our preferred formalizations of language. A logical argument is not necessarily a plausible biological argument. We see this as a clear indication that 'business as usual' will not cut it for studying the biology of language. Linguistic description is a fine research topic, and it becomes more interesting when the theoretical tools are more sophisticated than

before, but it is now clear that there is no need to call that enterprise 'biolinguistics in the weak sense' if biology is left at the door and used just rhetorically.

Bickerton himself, just before his quote above, seems suspicious of the success of this distinction:

> Fears widespread among both linguists and non-linguists that 'biolinguistics' may turn out to be merely a more scientific-sounding term for generative minimalism are reinforced by the way the distinction is made between 'strong' and 'weak' senses of biolinguistics by Boeckx and Grohmann (2007: 2).
>
> (Bickerton 2014, 73)

Bickerton justifies his suspicion by conflating the two senses into the 'weak' one, and goes on to criticize the practices of minimalists when dealing with biology and more specifically Evo-Devo, with the general idea being that they cherry-pick their biology in order to support their minimalist claims. This is very clear from the following quotes from Jackendoff (2011):

> In recognition of the goal of interfacing linguistic theory with biology, practicioners of the minimalist program have begin calling the enterprise (e.g. Jenkins 2000, Larson et al. 2010, Di Sciullo & Boeckx 2011, and the online journal *Biolinguistics* (www.biolinguistics.eu)).
>
> (Jackendoff 2011: 589)

> In each case the proposed answer is also motivated on grounds internal to language, and in each case it differs from the proposals of the minimalist program and biolinguistics, which are based on criteria of 'perfection', optimal design, and efficient computation.
>
> (Jackendoff 2011: 617)

But the conflation we are focusing on here should go the other way: if there is a failure to properly deal with biology on the part of minimalists, it is the strong sense that should be preserved as the biolinguistics worthy of the name, and the weak sense demoted.

2.3 Biolinguistics as the genetics of language

In generative grammar, the fact that the very specific properties of UG cannot by definition be inferred from the environment has quickly led to the assumption that they are to be found in the genes. This idea has been expressed in several ways throughout the recent history of the field, with terms like 'linguistic genotype' or 'genetic endowment' becoming staples in the UG literature, as the following quotes illustrate.

> The first section [of the book] includes contributions from Howard Lasnik, Tim Hunter and Paul Pietroski. The three chapters address issues regulating the derivations of FLN [Faculty of Language in Narrow sense], as well as more general issues pertaining to the mapping of syntactic expressions onto logical forms. By doing so they raise fundamental questions on the nature of FLN and its interfaces with the semantic system, addressing the central biolinguistic questions on the properties of the genetic endowment of human language. The discussions on the properties of the form of the operations and the derivations on FLN contribute to theoretical biolinguistics.
>
> (Di Sciullo 2012: 1)

> The theoretical proposals of results from experimental studies present genuine contributions to the Biolinguistic program. They contribute to our understanding of the properties of the interfaces derived by the computations of FLN, viewed as an organ in human biology. The chapters add to the usual notion of interfaces, which is generally understood as the connection between the semantic system on the one hand and the sensorimotor system on the other. They raise novel interface questions on how these connections are possible at all. They anchor the formal properties of grammar at the interfaces between language and biology, language and experience, as well as factors reducing complexity.
>
> (Di Sciullo 2012: 1)

> [...] language emerges through an interaction between our genetic inheritance and the linguistic environment to which we happen to be exposed. English-speaking children learn from their environment that the verb *is* may be pronounced [iz] or [z], and native principles prevent the reduced form from occurring in the wrong place. Children learn from their environment that *he*, *his*, etc. are pronouns, while native principles entail where pronouns may not refer to a preceding noun. The interaction of the environmental information and the native principles accounts for how the relevant properties emerge in an English-speaking child.
>
> (Anderson and Lightfoot 2000: 6).

Numerous practicioners in biology know that this gene-centric view is far too simplistic. There is no direct route from a linguistic entity – or a native principle [that prevents] the reduced form from occurring in the wrong place –, and a gene or genes. For this reason, we think that biolinguists should refrain from narrowing the domain of the field to genetics, and should avoid conflating nativism and geneticism.

2.4 Biolinguistics as the study of the uniquely human and linguistic

The idea that whatever constitutes the faculty of language is uniquely linguistic and uniquely human has received a great deal of attention. This has become very obvious especially after the work of Hauser et al. (2002) who make the distinction between Faculty of Language – Broad Sense (FLB), that is, that which contributes to language but is not unique to it or humans, and therefore to be found in other domains and species, and Faculty of Language – Narrow, sense (FLN), that which is unique to language and to humans. It seems that most researchers in the field of generative linguistics went for the latter. The result has been that many expect FLN to be the main focus of inquiry of biolinguistics, and this goes for proponents of generative grammar and more specifically FLN, who do not recognize the importance or even the legitimacy of studying broader, shared mechanisms, and for critics, who mistakenly think FLN is what biolinguists must care about and therefore incredulously await for evidence of such biological uniqueness. It is then no surprise to find 'laments' like the following:

> But if language is a biological object, as anyone holding a biolinguistics view must assume, then language must have at least some biological, that is physical, properties. Furthermore, specifically biological linguistic properties must be at least in some aspect different from biological non-linguistic properties. Yet, the biolinguistic enterprise has not specified a single biological property that is uniquely linguistic.

> (Behme 2015: 38)

It is sometimes hard to distinguish between critics' own assumptions about biology and their emulation of what they perceive the biolinguists' assumptions to be, for the sake of their argument. Whatever the case, there is no reason to hold that just because language is due to biology, that it should have 'specifically biological linguistic properties [that] must be at least in some aspect different from biological non-linguistic properties'. Behme's (2015) criticism[4] is thus addressed at those who think that some form FLN – that which is defined as biologically unique to language and humans – is the only thing worth studying, a group which not only is not representative of all biolinguists, and certainly not of those who pay attention to evolutionary biology, but is also most likely wrong, for reasons we discuss elsewhere (see Boeckx 2013), but also hint at in §3. The lack of evidence for a biologically special linguistic structure is a failure only for defendants of the FLN/FLB distinction (and even more so for those who focus on FLN or take it to be the interesting half of the distinction).

This kind of criticism will most likely continue to occur for as long as linguists presuppose or insist on the idea that language is a well-delineated, well-confined, *sui generis* biological 'organ', and definitely for as long as they proclaim that to be a (or the) central tenet and focus of biolinguistics.

2.5 Biolinguistics as analogy to biology

Yet another sense in which biolinguistics is used is as an analogy between language and biology, that is, drawing commonalities between perceived structural and evolutionary principles of language and those of

---

4 It is worth noticing that Behme (2015) is a continuation of an ongoing discussion with another author, which revolves around what is called 'Chomsky's biolinguistic ontology' – despite the far more ambitious and generic title –, and as such qualifies for what we have identified as 'biolinguistics as a Chomskyan enterprise'. Ultimately, the discussion does not bear on what we believe are genuine biolinguistic concerns.

biological structures in general. Analogies and metaphors can be used to one's advantage, to convey a particular point, but there is nothing to be gained from saying that the evolution and nature of language is akin to that of organisms in general if nothing more than superficial (and not very clear) resemblance connects them. Instead, it is important to recognize that the mechanisms that subserve language evolution and development are *the same ones* that are valid for all of biology. In other words, it is not the case that there is one set of mechanisms that subserve biology *and* one set of mechanisms that subserve language exclusively (and therefore no reason to make the latter resemble the former while keeping them separate). The faculty of language – whatever one thinks it is – is part of nature, and as such it is subject to biology and ultimately physics, as one other aspect of living things. As an illustration of the sense discussed in this section, recall Mukherji's passage, which we quote on page 5, where he says that "Biolinguistics is currently suggesting that structure of language may be "perfect" in design, not unlike the arrangement of petals in the sunflower and the double helix of the DNA". The allusion to petals in the sunflower or the double helix of the DNA, though poetic, does not serve any purpose other than somehow – in the mind of the author – offering further validation of the *bio-* prefix. A demonstration of the *same principles* in action would be more interesting.

Witness, for example, the following passage, from a recent chapter entitled "Object pronouns in the evolution of Romanian: A biolinguistic Perspective":

> The notion of language evolution goes beyond the classical notion of language change and grammaticalization (Roberts & Roussou 2003) by incorporating recent results from evolutionary developmental biology. This incorporation has both descriptive and explanatory advantages over classical notions of language change and grammaticalization. The descriptive advantage is that fluctuating stages are predicted to occur and can be described systematically. The explanatory advantage is that questions such as why languages change and why grammaticalization exists can be addressed on the basis of the existence of general laws governing the development and evolution of biological form.
>
> (Di Sciullo and Somesfalean 2015: 272–273)

This instance of "evolutionary developmental biology" above, which we assume motivates the title chosen by Di Sciullo and Somesfalean (2015), constitutes the only mention of biology in the whole text. The reader is then left wondering about what "recent results from evolutionary developmental biology" the authors have in mind, and what "general laws governing the development and evolution of biological form" support their work. The authors point to what they call the 'Directional Asymmetry Principle', already in Di Sciullo (2011), according to which "language evolution is symmetry breaking", an analogy to the notions of symmetry breaking in physics in biology. But no rationale is given for the choice of analyzing the properties of Romanian in such a way, as no discussion of how one gets from the underlying, micro scale mechanisms of symmetry breaking to the faculty of language and then to variation in languages. And much less for why this would be the key factor. It seems to us that many important steps are being skipped if one can jump directly from the impressionistic appropriation of biophysical mechanisms to diachronic variation of any one language, specially while ignoring what should be the main topic of any biolinguistic inquiry: language as a cognitive faculty. It is interesting to observe that proposals inspired by other work – as opposed to proposals *built on* on other work – tend to ignore everything else around it and assume it as all-explanatory.

2.6 Biolinguistics as a Chomskyan enterprise

All of the major senses of biolinguistics we have distilled from the literature (as generative linguistics, minimalism, genetics of language, study of FLN, and analogy to biology) have one thing in common: they are all equated with Chomsky's work or work heavily inspired by it. From there, a more generic, blanket-sense becomes clear: biolinguistics as a (or the) Chomskyan enterprise. Chomsky's work is undeniably polarizing: it seems that one is supposed to be either for it or against it as a whole (of course, some variation on the details is allowed). We find it fair to say that both supporters and critics of Chomsky's take his role in the genesis and development of biolinguistic thought to mean that all of biolinguistics must rest

upon and be confined by his and his followers' work and opinions, and as a result endorse or dismiss a field that is far broader.

It is then not surprising to find passages such as the following:

> For decades *Chomskyan biolinguistics* have [sic] failed to respond to arguments showing that their research program rests on an internally incoherent ontology. [...]I conclude that six decades after the Chomskyan revolution, the foundations of biolinguistics remain internally incoherent.
>
> (Behme 2015: abstract, our emphasis)

The idea that Chomsky's work and biolinguistics are equivalent is also found in non-linguistic literature:

> The idea that language can be approached in these terms is stressed in some recent work under the heading of Biolinguistics (see e.g. Chomsky (2005)). While we are sympathetic to many of the (mostly programmatic) suggestions in *Chomsky's work*, in practice much of the work that falls under that particular heading differs markedly in focus from the programme that we advance here.
>
> (Embick and Poeppel 2015: note 3, our emphasis)

The fact that these authors refer to biolinguistics "in the third-person" also goes to show that the way in which the term has been used despite its core meaning is indeed confusing. Something has gone wrong if someone who actively studies the biological properties of language (and who frequently points out important methodological problems that redefine this endeavor, e. g., Poeppel and Embick 2005) does not identify with biolinguistics. A quick glance through Poeppel's work will reveal a much closer approximation to serious biolingistic work – studying the biological properties of language – than what is the case in the work of those who most frequently use the term.

Incidentally, we think that another valuable insight can be gained from this passage. There is indeed a strong impression from outside of linguistics that Chomsky's work offers mostly programmatic proposals, suggesting that failing to go beyond it will prevent progress from actually taking place. It is the role of biolinguists to go beyond programs and towards proposals.

Returning to the notion of biolinguistics as Chomsky's work in the way it is used in the linguistic literature, we think that the sort of pluralism that a genuine biolinguistics demands forces us to cast a much wider net and gather insights not only from the generative tradition in linguistics, but also from approaches that are often taken as incompatible with the orthodox Chomskyan viewpoint, for instance work in cognitive linguistics, Construction Grammar, etc.

3 A biologically informed study of language

It is not reasonable to expect one researcher to be an expert in all fields that form part of the attempt to elucidate the biological foundations of language. But they must care. As there is still no degree in biolinguistics, any one who calls himself a biolinguist will have a specific background that might dictate more or less the way in which research is carried out, but they must nevertheless pay attention to the information and solutions from other fields, as a compass of sorts. This is now a central concern in biology at large, where it has become apparent that progress lies on being aware and ideally comfortable with neighboring fields, and this is something that must be fostered already at the undergraduate level or even before, in an effort to ensure that real interdisciplinarity becomes not only possible but also the norm. The modern field of Systems Biology arose out of concerns of this kind, and efforts are being put in place to ensure that the research methods of biological properties go hand in hand with their eclectic nature. The same way that it is now understood that biological properties are not controlled by single genes, but instead are the result of complex, dynamic interactions between genes, the environment, and everything in between, the fields that are necessary for the study of all these aspects are diverse, and must come together in a dynamic interaction as well (Cvijovic et al. 2016). The lesson from Systems Biology that we ought to pay

attention to is that compartmentalization should be fought against: it is when borders are crossed that real progress takes place.[5]

In practice, being a biolinguist means both collaborating with researchers in other disciplines and becoming well versed in their methods and literature. The findings and proposals ranging from genetics to neuroscience to cultural transmission must inform and constrain biolinguistic proposals. And a good biolinguistic proposal will have ramifications that make sense in others fields as well. This does not mean forcing hot topics into every proposal (e. g. whenever a gene is discovered to have played a crucial role in human evolution or development, this should not be taken as *carte blanche* for proposing that gene as the one for language). Instead, it means bearing in mind that for a theory to be good, it must make sense in light of what is known independently. Thus, a good theory of the language faculty must be constrained by biological theory, that is, what doesn't make sense in biology cannot make sense in a theory of language.[6]

An important step towards achieving a biologically plausible account of language is thus to rethink and, if need be, do away with notions used traditionally in linguistics. A fair reaction to this call would be to ask "why, and why not have the biologists adapt to linguistics?". The answer is very simple: biologists are not doing linguistics or studying what are ultimately linguistic objects. Because the converse is claimed to be true, linguists have to be prepared for the primitives they posit and the theories they formulate to be scrutinized by the biologist who comes across them, and they should actually foster this kind of interaction. We find it remarkable how many linguists set themselves apart from the 'traditional linguist' by making clear what their assumptions are, and by actively seeking theoretical sophistication, but at the same time play down concrete biological insights and resist interdisciplinary efforts, taking cover behind the 'linguistics is biology at a suitable level of abstraction' mantra. This only contributes to an insular and basically unfruitful state of affairs. Linguistics must come to grips with biological information if dialogue is to take place. Developing its own ontology, though a guarantee of circumscribed success, is not the way to do it. The advantages of making dialogue and interaction possible are not only on the side of the linguists interested in language in a biological context, but also on the side of the biologists who may profit from insights coming from linguists, something which so far has not really happened either. If linguists do not worry about the evolution and brain implementation of what they posit, no one else will do that for them.

The exact notions that warrant serious rethinking should become apparent as interdisciplinary research proceeds. If a notion cannot be accommodated by – or, more importantly, if it is shown to be completely at odds with – current biological theory and data, that's a red flag. Such a notion will need to be reframed or even abandoned. This is a customary procedure in all sciences, and it need not be taken as radical. Anyone

---

5 In the specific case of biolinguistics, the need for change is perhaps even more urgent, given that a large portion of what makes up the study of language typically finds a home in departments of humanities, whose academic and administrative practices are far removed from those of the hard sciences. Opening up or facilitating the possibility of attending introductory courses in different subjects, even if it means going to different departments or faculties, would be a good way of encouraging interested students to adopt an interdisciplinary mindset. At the graduate level, taking in a co-advisor from a different field should be seen not as something exotic and out of the ordinary, but instead as something to be fostered. Ideally, the study of language as a capacity would not be tied to humanities departments, but instead become a topic of interest that can be studied at different departments depending on the specific problem to be studied. This is not to say that such fertile environments for the study of language do not exist at all; they do, but they are few – too few given the complexity of the problem but also the inherent interest this field attracts even in the general population. Making these proposals a reality will also depend on the perceived virtues of the interdisciplinary study of language, i. e., genuine results. It is the role of those already in a position to make a difference to highlight the benefits of interdisciplinary approaches, both as instructors and as researchers.

6 It goes without saying, as we hope is apparent throughout this paper, that we are not addressing linguists who explicitly do not care about the nature of language, whose work is not affected in the slightest by biological considerations (say, if biology changed overnight beyond recognition, the work of the descriptive linguist would remain unscathed; what they do is devise formal descriptions of patterns in linguistic corpora).

whose goal is to get closer to a good biological theory of language should be glad to proceed this way: whittle away whatever isn't biologically plausible (or recontextualize it in a biologically plausible manner). As an example that merits this treatment given current biological knowledge, we'll discuss the notion of uniqueness.

No one doubts that only humans have language, and that this fact makes humans special. One way – perhaps the way that prevailed until recently – to get to the bottom of this fact is to look for something biologically unique in humans, which by process of elimination makes them, indeed, special. This is also what led Hauser et al. (2002) to propose the FLN/FLB distinction – they recognize that the faculty of language must have evolved, yet they still want to account for its uniqueness. The focus should, however, be on FLB (to the point where the distinction stops making sense). We should abandon a top-down, anthropocentric perspective (why humans are different from other species) in favor of a bottom-up perspective (what humans have in common with other species). This kind of bottom-up approach is in fact one of the traits of the minimalist program in linguistics ("approaching UG from below"; Chomsky 2007), but the important motivations for it are not always kept in mind. The goal is not to have a minimalist theory (that's a consequence); it is rather to assign biological concerns the role they undoubtedly deserve. We should then go further: We should look for basic underlying mechanisms in the animal kingdom, as opposed to looking for "human" cognitive capacities, which anyway are only human in hindsight; evolution is not goal-oriented. The way to best interpret this message is to follow the wide-ranging bottom-up approach in comparative psychology, which de Waal and Ferrari (2010) herald in a recent opinion article. They ask the following question: "What if we were to replace our obsession with complex cognition with an exploration of basic processes? Instead of asking which species can do X, the question would become how does X actually work?".

There's another, related reason for rethinking uniqueness: When one looks at it from an evolutionary point of view, it becomes *novelty*. The bar is set very high for what constitutes a novel trait:

> A morphological novelty is a structure that is neither homologous to any structure in the ancestral species nor homonomous to any other structure of the same organism.
>
> (Muller and Wagner 1991: 243)

This looks like a generalization of Hauser et al.'s (2002) FLN. But current biology, especially under the impetus of Evo-Devo, finds it very hard to be certain about real examples of novel traits, even if we are just looking at non-cognitive, morphological ones (Moczek 2008; Prud'homme et al. 2011). In order words, in biology things are extremely unlikely to be truly unprecedented; they are built on top of previous structures, different from the end result that we see but extremely important in its history and nature. This is not a problem for the researcher with little interest in the real biological implications of his work or the theories that support it. But for the serious biolinguist it forces a rethinking of what it means to deem language a novel, *biologically unique* trait, on the one hand, and what it means to study it from a biological point of view, on the other: as a problem (Martins and Boeckx 2016; Boeckx 2016), and not as a mystery (cf. Hauser et al. 2014).

Thus, in more practical terms, a biolinguist should not be surprised with – and should in fact welcome – the study of phenomena that apparently have nothing to do with language or with humans. This goes for behavioral phenomena across domains and species, in an attempt to disentangle and understand basic mechanisms that subserve linguistic abilities, but also for different levels of analysis of these phenomena. By different levels of analysis we are of course not referring here to the traditional divisions in linguistics (syntax, semantics, phonology, etc.), but instead to the different dimensions one has to travel through to get from the phenome to, ultimately, the genome. It is essentially a "translation" task. If we do our job the right way, that is, if we build "linguistic models that are explicit about the computational primitives (structures and operations) they require, and that attempt to define linguistic problems at a fine enough grain that one can discuss algorithmic and implementational approaches to their solution" (Fitch 2009), we can reconcile what we know about language with what the geneticists,

the neuroscientists, and the neurobiologists know, instead of resisting to do so and asserting and/or complaining that so little is known in these areas.

No one doubts that language is special, but maybe none of its components are, and what is remarkable is to find them all in the same organism, reshaped and put to new use. Chomsky has actually hinted at this point before:

> Now a question that could be asked is whether whatever is innate about language is specific to the language faculty or whether it is just some combination of the other aspects of the mind. That is an empirical question and there is no reason to be dogmatic about it; you look and you see. What we seem to find is that it is specific.

(Chomsky 2000)

In the context of the FLN/FLB distinction, this point was also brought up:

> Something about the faculty of language must be unique in order to explain the differences between humans and the other animals – if only the particular combination of mechanisms in FLB.

(Fitch et al. 2005: 182)

The claim above actually renders the whole notion of FLN (and thus the FNL/FLB dichotomy) useless. In fact, perhaps in recognition of that very state of affairs, [Fitch 2011: 384], in a departure from the work he did with Chomsky and Hauser, seems to have abandoned the distinction altogether, and we think he is right:

> [...] the distinction between general and linguistically specialized mechanisms is hard to draw, even is those cases where the mechanisms seem fairly clearly defined. Most areas of language are not, and will not soon be, so clearly defined, and the distinction itself is of little use in furthering our understanding of the mechanisms.

One thing that the biolinguist must do in order to capture the dynamic way in which complex traits come to be – they do not come out of nowhere – is to explore the role of development. This approach is embodied in Evolutionary Developmental Biology (Evo-Devo) (Figure 1). This is no easy task to undertake, and caution is needed: Evo-Devo is not a straightforward, well-delimited field, but rather an approach that seeks to reconcile evolution and development – an aspect lacking in (neo-)Darwinistic versions of biology, prevalent at the time contemplation of language in a biological context started.
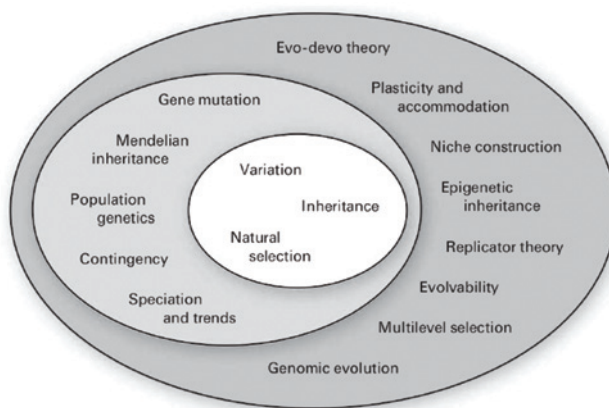

Figure 1: The current Evo-Devo landscape [Pigliucci and Muller, 2010].

As an example of how taking development into account forces the rethinking of long-held assumptions, recent genetic research indicates that one quality that the language faculty *doesn't* display is homogeneity – there are individual developmental differences at the neuronal level that are influenced

by genetic factors (Kos et al. 2012; Le Floch et al. 2012). Results such as this pave the way for contextualizing language in a broader picture.

There are many examples of work which in our view is on the right track, and could be taken as illustrative of the direction that seems most fruitful for the study of the biology of language. Poeppel and colleagues' work on a model of processing based on neuronal oscillations is one such clear example (e. g. Giraud and Poeppel 2012; Ding et al. 2016). There is also very important work that seeks to explicitly explore and characterize the genetic underpinnings of language, which shows the inadequacy of the single-mutation story that many linguists naïvely insist on. It shows us that language, like other aspects of human biology and behavior, is a product of the intricate workings of neuronal circuits, which are shaped by a complex network of genes in interaction with environment. Figuring out the molecular biology of language is not a small detail that researchers of language should not care about and leave to whoever feels like picking it up: it is a necessary condition for understanding how language came about and why in humans. An easily digestible review of work of this kind is offered by Fisher and Vernes (2015), who themselves along with many colleagues have contributed immensely to it over the years. Animal models are also essential for this endeavor, particularly in the identification of further genes that affect language related disorders (such as *FOXP2* or *CNTNAP2*), a crucial window into how language works (see Vernes and Fisher 2013). In this context, but also in the pursuit of neural mechanisms that may underlie human speech and language, much valuable work has been done and seems to us to be of increasing importance, namely in the case of vocal learners among birds and bats, which lend themselves to fruitful comparative research (for reviews of important work by some of its main representatives, see, for example, Fitch 2010; Fitch and Jarvis 2013; Chakraborty and Jarvis 2015; Okanoya 2015; Rodenas-Cuadrado et al. 2015). There is of course much that could mentioned, but we believe that looking into the nature and insights of the work of the kind we have alluded to here paints a realistic picture of what we see as the way forward.

4 Conclusions

We do not own the term 'biolinguistics', and obviously people are free to use it in any way they want. But our impression is that work under the rubric of biolinguistics often tends to be dismissed by researchers from other disciplines because they associate the term with one or more of the senses we have reviewed in this paper.

We cannot disagree with Eric Raimy when he writes that:

> *Biolinguistics* is the current buzzword adopted by linguists pursuing a particular branch of generative grammar closely associated with Noam Chomsky's Minimalist Program (Chomsky 2005). This word refers to a combination of aspects of biology and linguistics that varies depending on the particular researcher and the particular moment.
>
> (Raimy 2012: 926)

Indeed, this seems to be an accurate description of those we included in group i), in our introduction (§ 1). We do not see how using the term as a more modern sounding buzzword for (generative) linguistics will do any service to the study of language from a biological point of view. If anything, the contrast between the ambitious *bio-* prefix and the naïve biological notions that usually come along with it is more likely to further the gap between linguistics and the biological sciences than it is to bring them closer together.

But, like him, we think that there is a far more promising use of the term:

> If we interpret biolinguistics as a research program that is dedicated to understanding the relationship between the genotypes and phenotypes responsible for explaining human language, the benefits of collaborations among biologists, linguists, psychologists, cognitive scientists, and zoologists are clear. Each scientific field provides a unique perspective of the content within this expansive definition.
>
> (Raimy 2012: 927)

We hope that integrative works along these lines will prosper.

References

Anderson, S. R. & D. W. Lightfoot. 2000. The human language faculty as an organ. *Annual Review of Physiology* 62(1). 697–722.

Behme, C. 2015. Is the ontology of biolinguistics coherent?. *Language Sciences* 47. 32–42.

Bickerton, D. 2014. Some problems for biolinguistics. *Biolinguistics* 8. 73–96.

Boeckx, C. 2013. Biolinguistics: Forays into human cognitive biology. *Journal of Anthropological Sciences* 91. 63–89.

Boeckx, C. 2016. The language-ready head: Evolutionary considerations. *Psychonomic Bulletin and Review*, doi: 10.3758/s13423-016-1087-5.

Boeckx, C. & K. K. Grohmann. 2007. The biolinguistics manifesto. *Biolinguistics* 1(1). 1–8.

Chakraborty, M. & E. D. Jarvis. 2015. Brain evolution by brain pathway duplication. *Philosophical Transactions of the Royal Society B* 370(1684). 20150056.

Chomsky, N. 1957. *Syntactic structures*. The Hague: Mouton.

Chomsky, N. 1959. Review of *Verbal Behavior*, by B. F. Skinner. *Language* 35(1). 26–57.

Chomsky, N. 1965. *Aspects of the theory of syntax*. Camdribdge, MA: MIT Press.

Chomsky, N. 2000. *The architecture of language*. Oxford: Oxford University Press.

Chomsky, N. 2007. Approaching UG from below. In U. Sauerland & H.-M. Gärtner (eds.), *Interfaces+recursion=language? Chomsky's minimalism and the view from semantics*, 1–30. Berlin: Mouton de Gruyter.

Chomsky, N. 2012. *The science of language: Interviews with James McGilvray*. Cambridge, MA: Cambridge University Press.

Cvijovic, M., T. Höfer, J. Acimovic, L. Alberghina, E. Almaas, D. Besozzi, A. Blomberg, T. Bretschneider, M. Cascante, O. Collin et al. 2016. Strategies for structuring interdisciplinary education in systems biology: An European perspective. *NPJ Systems Biology and Applications* 2. 16011.

de Waal, F. & P. F. Ferrari. 2010. Towards a bottom-up perspective on animal and human cognition. *Trends in Cognitive Sciences* 14(5). 201–207.

Di Sciullo, A. M. 2011. *A biolinguistic approach to variation*, 305–326. Oxford: Oxford University Press.

Di Sciullo, A. M. 2012. Interfaces in a biolinguistic perspective. In A. M. Di Sciullo (ed.), *Towards a biolinguistic understanding of grammar: Essays on interfaces*. Amsterdam: John Benjamins Publishing.

Di Sciullo, A. M. & S. Somesfalean. 2015. *Object pronouns in the evolution of Romanian: A biolinguistic perspective*, 269–289. Leiden: Brill.

Ding, N., L. Melloni, H. Zhang, X. Tian & D. Poeppel. 2016. Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience* 19(1). 158–164.

Embick, D. & D. Poeppel. 2015. Towards a computational (IST) neurobiology of language: Correlational, integrated and explanatory neurolinguistics. *Language, Cognition and Neuroscience*, 30(5). 357–366.

Fisher, S. E. & S. C. Vernes. 2015. Genetics and the language sciences. *Annual Review of Linguistics* 1(1). 289–310.

Fitch, W. T. 2009. Prolegomena to a future science of biolinguistics. *Biolinguistics* 3(4). 283–320.

Fitch, W. T. 2010. *The evolution of language*. Cambridge: Cambridge University Press.

Fitch, W. T. 2011. Unity and diversity in human language. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 366(1563). 376–388.

Fitch, W. T., M. D. Hauser & N. Chomsky. 2005. The evolution of the language faculty: Clarifications and implications. *Cognition* 97. 179–210.

Fitch, W. T. & E. D. Jarvis. 2013. Birdsong and other animal models for human speech, song, and vocal learning. In M. Arbib (ed.), *Language, music, and the brain*, 499–539. Cambridge, MA: MIT Press.

Giraud, A.-L. & D. Poeppel. 2012. Cortical oscillations and speech processing: Emerging computational principles and operations . *Nature Neuroscience* 15. 511–517.

Givón, T. 2002. *Bio-linguistics: The Santa Barbara lectures*. Amsterdam/Philadelphia: John Benjamins.

Hauser, M. D., N. Chomsky & W. T. Fitch. 2002. The faculty of language: What is it, who has it and how did it evolve?. *Science* 298. 1569–1579.

Hauser, M. D., C. Yang, R. C. Berwick, I. Tattersall, M. Ryan, J. Watumull, N. Chomsky & R. Lewontin. 2014. The mystery of language evolution. *Frontiers in Psychology* 5. 401).

Jackendoff, R. 2011. What is the human language faculty?: Two views. *Language* 87(3). 586–624.

Kos, M., D. van den Brink, T. M. Snijders, M. Rijpkema, B. Franke, G. Fernandez & P. Hagoort. 2012. Cntnap2 and language processing in healthy individuals as measured with erps. *PloS one* 7(10). e46995.

Le Floch, É., C. Lalanne, V. Frouin, P. Pinel, L. Trinchera, A. Tenenhaus, A. Moreno, M. Zilbovicius, T. Bourgeron, S. Dehaene et al. 2012. Significant correlation between a set of genetic polymorphisms and a functional brain network revealed by feature selection and sparse partial least squares. *Neuroimage*. 63(1). 11–24.

Lenneberg, E. H. 1964. A biological perspective of language. In E. H. Lenneberg (ed.), *New directions in the study of language*. Cambridge, MA: MIT Press.

Lenneberg, E. H. 1967. *Biological foundations of language*. New York: Wiley.

Martins, P. T. & C. Boeckx. 2016. Language evolution: Insisting on making it a mystery or turning it into a problem?. In L. Dupuy, A. Grabizna, N. Foudon & P. Saint-Germier (eds.), *Papers dedicated to Anne Reboul*, 1–8. Lyon: Institut des Sciences Cognitives.

Meader, C. L. & J. H. Muyskens. 1950. *Handbook of biolinguistics*., volume 1 Toledo: HC Weller.

Moczek, A. P. 2008. On the origins of novelty in development and evolution. *BioEssays* 30(5). 432–447.

Mukherji, N. 2010. *The primacy of grammar*. Cambridge, MA: MIT Press.

Muller, G. B. & G. P. Wagner. 1991. Novelty in evolution: Restructuring the concept. *Annual Review of Ecology and Systematics*, 22, 229–256.

Okanoya, K. 2015. Evolution of song complexity in bengalese finches could mirror the emergence of human language. *Journal of Ornithology* 156(1). 65–72.

Pigliucci, M. & G. B. Müller. 2010. *Evolution: The extended synthesis*. Cambridge, MA: MIT Press.

Pillsbury, W. B. & C. L. Meader. 1928. *The psychology of language*. New York: D. Appleton & Co.

Poeppel, D. & D. Embick. 2005. Defining the relation between linguistics and neuroscience. In A. Cutler (ed.), *Twenty-First Century Psycholinguistics: Four Cornerstones*, 103–118. Mahwah, NJ/London: Lawrence Erlbaum Associates.

Prud'homme, B., C. Minervino, M. Hocine, J. D. Cande, A. Aouane, H. D. Dufour, V. A. Kassner & N. Gompel. 2011. Body plan innovation in treehoppers through the evolution of an extra wing-like appendage. *Nature* 473(7345). 83–86.

QJS. 1925. Laboratory and research. *Quarterly Journal of Speech* 11(3). 274–285.

Raimy, E. 2012. Phonological architecture: A biolinguistic perspective. *BioScience* 62(10). 925–927.

Ramchand, G. 2015. Statement. Roundtable discussion. *Generative Syntax in the Twenty-first Century: The Road Ahead*, Athens, Greece. Also posted as *On being a generative linguist* at http://generativelinguist.blogspot.com./2015/04/i-am-generative-linguist-a.html.

Rodenas-Cuadrado, P., X. S. Chen, L. Wiegrebe, U. Firzlaff & S. C. Vernes. 2015. A novel approach identifies the first transcriptome networks in bats: A new genetic model for vocal communication. *BMC Genomics* 16(1). 836.

Vernes, S. C. & S. E. Fisher. 2013. Genetic pathways implicated in speech and language. In S. A. Helekar (ed.), *Animal models of speech and language disorders*, 13–40. Springer.

Appendix B

Language evolution: Insisting on making it a mystery or turning it into a problem?

Language evolution: Insisting on making it a mystery or turning it into a problem?

Pedro Tiago Martins[1] and Cedric Boeckx[1,2]
[1]Universitat de Barcelona,[1,2]ICREA

In a recent, widely-read paper, Hauser et al. (2014) oer a rather negative view of the state of aairs in language evolution. More specically, the authors believe that little to no progress has been made in the various relevant elds regarding the age-old questions of the origin and evolution of the human capacity for language. We beg to dier.

The authors' strategy is to target some of the elds that have spawned the most activity and hypotheses in recent years (comparative animal behavior studies, achaeology, molecular biology and modelling), and then show what they have done wrong. These elds, they say, have not advanced much more than speculation. Instead, we think that it is the progress in these elds that accounts in large part for the revival of biolinguistic concerns (Boeckx, 2013).

The intention of Hauser et al.'s paper is to point to the damage that has been done during the last decades, by calling attention to the dangers of jumping from simplistic, impoverished data and observations to full-edged accounts. To some extent, we agree. But we nd it curious that linguistics is not of the targets of the paper, even though the eld is rife with speculative and untestable proposals and implications for how language evolved. The implicit but in our view obvious corollary is that — for the authors — linguistic theorizing plays at present a crucial

role in advancing what we know about language evolution, or at the very least does not have much to be criticized (while other elds do have a lot to be criticized for, since they do not match what has been or could be accomplished by linguistic theorizing). We take this absence with a grain of salt, as we nd it hard to explain how a paper on the status of language evolution studies does not even dabble in the shortcomings of what is in e ect the eld of expertise of half of its 8 co-authors.

While it is true that we do not know how language evolved — if we did, no one would be working on it any more —, to diminish the work that has been done recently on various disciplines to the point of irrelevancy is not only dubious (we feel it ends up throwing the baby with the bathwater) but, in the case of this group of authors, confusing (a close look at the literature will reveal that di erent combinations of the authors of Hauser et al. (2014) make arguments of the sort they take issue with, and rely on sources of information that in the paper under discussion are deemed unreliable). In what follows we will brie y touch on the di erent elds targeted by Hauser et al. (2014), and point out both incongruence and unjusti ed pessimism in their arguments. We will not o er here in-depth rebuttals or quali cations of the authors' positions, but instead provide a little glimpse into what we see as more heat than light.

In relation to the archaeological record, which the authors in the abstract say "does not inform our understanding of the computations and representations of our earliest ancestors, leaving details of origin and selective pressure unresolved", Chomsky (2005, p. 3), on the basis of work by Tattersall, writes of the faculty of language as part of a "a complex of capacities that seem to have crystallized fairly recently, perhaps a little over 50,000 years ago, among a small breeding group of which we are all descendants — a complex that sets humans apart rather sharply from other animals, including other hominids, judging by traces they have left in the archaeological record." In the very same page, Chomsky goes on to say that the great

leap forward is "the result of some genetic event that rewired the brain, allowing for the origin of modern language with the rich syntax that provides a multitude of modes of expression of thought, a prerequisite for social development and the sharp changes of behavior that are revealed in the archaeological record [. . .]." The same ideas are echoed, for example, in Chomsky (2010). Yang (2010) claims that we cannot ask too much of Universal Grammar, because "[a] theory of Universal Grammar is a statement of human biology, and one needs to be mindful of the limited structural modi cation that would have been plausible under the extremely brief history of *Homo sapiens* evolution." But how do we know this if language evolution has been a complete mystery for years? Speculation goes both ways, and one should not dismiss one and support the other. It seems that for Hauser et al. (2014) arguments of this sort were ne while the relevant archaeological record was thought to have been left by humans, and only now that we know it is most likely Neanderthal (e.g. Zilhão, 2011), the authors claim we shouldn't try to derive inferences from archaeology. Nevertheless, one needs not look hard to nd resort to archaeological evidence in support of a non-gradualist position as recently as earlier this year (e.g. Berwick and Chomsky, 2016, pp. 37–38), leaving us all the more confused as to what their overall position regarding its reliability as a source of information really is.

Hauser et al. (2014) also take issue with comparative animal work, which "provide[s] virtually no relevant parallels to human linguistic communication, and none of the underlying biological capacity." The problem with this assessment is that it equates the testing of all-or-nothing hypotheses (animal X displays some form of language phenotype property P) with everything that such endeavors might have to o er. We take it that not many people still believe in "talking birds" and "signing apes" (if this was ever the case for serious scientists), but those studies and their scrutiny were important to determine what humans and non-human animals do, and animal studies are becoming increasingly more important in the study of underlying

mechanisms shared by di erent species and formulation of hypothesis concerning humans in particular. Berwick and Chomsky (2013) seem like they would agree, and Berwick et al. (2011), for example, draw a connection between birdsong syntax and underlying mechanisms of human speech, and state that "comparing the structure of human speech and birdsong can be a useful tool for the study of evolution of brain and behavior" (p. 120). This quali es as Hauser et al.'s (2002) FLN, which in the present paper the authors stress as referring not only to the mechanisms for discrete in nity but also to the "mappings to the interfaces with the conceptual-intentional and sensory-motor systems." Hauser et al. (2014) are right to point out that some current techniques used in animal studies fail to capture the animals' actual capacities, which they are more likely to display in their natural habitats, roaming free and devoid of extensive, goal-oriented training, but in doing so they are targeting the lookout for the linguistic phenotype in other species, rather than the bottom-up comparative work of the kind advocated, for example, by de Waal and Ferrari (2010), to which we will return later.

As for molecular biology, Hauser et al. (2014) do not present a critique *per se*, but rather an overview of current work which shows that there is no clear path from genes to linguistic behavior. This is not surprising to molecular biologists, and in fact simplistic proposals of the sort they criticize — coming up with just-so stories out of thin air, or on the basis of impoverished observations — usually come from the eld of linguistics (see Boeckx (2016) for discussion of a recent example). It is for this reason that work in linguistics must provide information that can be used to creating linking hypotheses, which currently and for the most part it cannot. This di culty in creating linking hypothesis between genes and linguistic behavior is ampli ed by this inadequacy of linguistics in providing primitives that other elds can work with (for a discussion of this problem, see Poeppel and Embick, 2005). A logical theory of the language faculty does not necessarily amount to a biologically plausible one, which is

what we should be aiming for. This state of affairs alone would warrant a discussion of linguistics as a source of information in language evolution studies that is absent from Hauser et al. (2014). The way in which authors present the linguistic phenotype — a novel recursion mechanism, unique to humans — is enough to stall or severely hinder the kind of linking hypotheses we would all would like to see, and which Hauser et al. (2014) say we have no hopes of seeing any time soon. The reason for that is that we actually know that novelty doesn't simply "arise". While traits may on the surface seem novel, or *sui generis* (for discussion, see Wagner and Müller (2002); Moczek (2008), among others), their nature is "largely reorganizational, rather than the product of innovative genes" (West-Eberhard, 2005, p. 6547), that is, phenotypic novelties are the result of the combination of different, more generic mechanisms.

Hauser et al.'s case against current work in modelling is the most consistent with each author's practice, but their general disdain for the role of culture in evolution — "In this paper, we are interested in biological *as opposed* to cultural evolution" (p. 2, our emphasis) — overlooks important advancements in evolutionary biology which show that culture and environment might really be crucial. "Culture" is a taboo notion in most generative circles, perhaps because it is usually seen as detrimental to biology in a theory of language. We find this to happen only under a naïve view of biology, along with an axiomatic incompatibility with linguistic approaches that give pride of place to culture. Crucially, one should not ignore the role of environmental factors in the shaping of the genotype, and in turn the shaping of the phenotype. There is no reason to seek explanation of phenotypic variation only in environmental or genetic factors. Instead, one should incorporate the lessons from Evo-Devo, and pay attention to work on the genotype-environment interaction (West-Eberhard, 2003), which shows that the degree to which environmental choices affect the way genetic blueprint is expressed depends on the specific genotype-environment interaction in each case.

In a somewhat more optimistic tone, Hauser et al. (2014) o er some suggestions of "paths forward", both interspersed throughout the paper and as a nal comment. These suggestions, however, are very much con ned, and su er from the same problems that their negative assessment of the various elds does. In a nutshell, the authors insist on gauging the usefulness of theoretical and experimental work by whether or not it "speaks" to Merge, the recursive mechanism they place at the center of the linguistic phenotype. It is not surprising that the presupposition that Merge must be at the center of inquiry into language evolution drastically reduces what can be done in practice, but in doing so it pushes the *mystery* the authors speak of. That is not what parsimony is for. Language evolution thus becomes a mystery only to adherents of this presupposition, and a problem — like many others in the sciences — for those willing to explore further.

In the case of animal studies, the authors put their money on the development of new techniques that could allow the collection of neural data from free-ranging animals, thus revealing their capacities in the absence of reinforcement. We agree that such techniques would work wonders for the eld, but what propels Hauser et al. (2014) is that we would then be able to devise a "set of stimuli that are generated from a recursive operation such as Merge (a recursive operation that combines two objects, such as two lexical items, to construct a new object, such as a phrase, in a process that can be iterated inde nitely), expose animals to a subset of these, and then test them on a wide range of alternatives that extend beyond the initial set in ways that can reveal substantial generalization, and thus comprehension of the underlying generative operation." (pp. 9–10) Presumably, these tests would reveal that animals either fail miserably or are able to generalize by relying on di erent, nite mechanisms, thus showing the uniqueness of Merge and supporting the discontinuity hypothesis. But there are myriad (other) ways in which animal studies can work in favor of a deeper knowledge about the biology of language. In this

context, we find it appropriate to quote a passage by (de Waal and Ferrari, 2010, p. 201):

> Over the last few decades, comparative cognitive research has focused on the pinnacles of mental evolution, asking all-or-nothing questions such as which animals (if any) possess a theory of mind, culture, linguistic abilities, future planning, and so on. Research programs adopting this top-down perspective have often pitted one taxon against another, resulting in sharp dividing lines. Insight into the underlying mechanisms has lagged behind. A change in focus now seems to be under way, however, with increased appreciation that the basic building blocks of cognition might be shared across a wide range of species. We argue that this bottom-up perspective, which focuses on the constituent capacities underlying larger cognitive phenomena, is more in line with both neuroscience and evolutionary biology.

Indeed, looking for a full-fledged ability such as language something that looks close enough to it is bound not to tell us much, but that's not what we should be looking for. Instead, we should decompose it into more generic mechanisms, not unique to neither the language domain nor the human species. This path will inevitably leads us to the study of abilities with little resemblance to language, and mechanisms at levels far deeper than the behavioral and the cognitive. But it's these levels we need to get to in order to arrive at true linking hypotheses.

As for modeling, the authors say that "it must focus on the computations and representations of the core competence for language, recognize the distinction between these internal processes and their potential externalization in communication, and lay out models that can be empirically tested in our own and other species." Again, it must speak to Merge (which is how we must interpret "the core competence for language" when reading Hauser et al. 2014), and a host of other possible modeling work is not even considered. We don't see how this would change the status of the field if all we are allowed to focus on is the core recursive mechanisms the authors equate with the linguistic phenotype (and perhaps the interfaces between and externalization systems, which are usually left vague in any

case). Opening one's mind to the role of the environment (or culture, which we find hard to tease from "environment" in a meaningful way) is likely to prove fruitful, and modeling work pays particular attention to the influence this aspect of the world might have. We agree with Kirby (2013, p. 473) that "the particular learning mechanisms that we bring to bear on the task of acquiring language are assuredly part of our biology. The key questions to ask about this aspect of the biology of language are: what is the nature of our biological endowment that relates to language learning? and to what extent is this endowment specific to language? These questions essentially define the biolinguistics enterprise, and their answer depends on an understanding of the relationships between learning, cultural transmission, and biological evolution."

In sum, Hauser et al. (2014) paint an ugly picture of language evolution that seems to have been caused by other, incautious scientists, while in reality the authors themselves have incurred in the same kind of arguments and assumptions — the kind they deem poor and speculative. This practice has not stopped with this paper: a quick read through the latest book by two of the authors (Berwick and Chomsky, 2016) will reveal discussion of topics that in Hauser et al. (2014) we are advised not to pay much attention to. This kind of incongruous back-and-forth is bound to cause more confusion than resolution. Furthermore, insisting on the idea that the evolution of language is mysterious — and not a problem we can look into right now, with its own difficulties and promising avenues — will deter only those who are stuck with a naïve view of biology and its dynamics that allows for such a simplistic position.

What is clear to us, and not so clear from reading Hauser et al. (2014) is that in order to make language evolution more of a problem and less of a mystery, everyone — linguists included — will have to make the mapping between mind to brain the focus of study. It is this intermediate level between genotype and phenotype that must be the target of intensive investigation. If the mind is what the brain does, it is

imperative to understand how the brain came to do what it does. This will necessarily involve a reconsideration of the nature and fabric of the language faculty, for only those descriptions of linguistic knowledge that can be associated with concrete neural correlates will have a ghting chance of going beyond the limitations of the fossil record, and exploit ndings in paleoneurology, paleogenetics, and comparative cognitive biology.

References

Berwick, Robert C & Noam Chomsky. 2013. "Foreword: A bird's-eye view on human language and evolution." In Johan J Bolhuis & Martin Everaert (eds.), *Birdsong, speech, and language*, ix–xii. Cambridge, MA: MIT Press.

— (2016). *Why only us; Language and Evolution*. MIT Press, Cambridge, MA.

Berwick, Robert C, Kazuo Okanoya, Gabriel JL Beckers & Johan J Bolhuis. 2011. "Songs to syntax: the linguistics of birdsong." *Trends in cognitive sciences* 15(3). 113–121.

Boeckx, Cedric. 2013. "Biolinguistics: Forays into human cognitive biology." *Journal of Antropological Sciences* 91. 63–89.

— Ljiljana Progovac, Evolutionary syntax (Oxford Studies in the Evolution of Language). Oxford: Oxford University Press, 2015. Pp. xv +261. *Journal of Linguistics*, First View:1–5.

Chomsky, Noam. 2005. "Three factors in language design." *Linguistic inquiry* 36(1). 1–22.

— "Some simple evo devo theses: How true might they be for language". In Richard K Larson Larson, Viviane Déprez & Hiroko Yamakido (eds.), *The evolution of language: Biolinguistic perspetives*, 45–62. Cambridge: Cambridge University Press.

de Waal, Frans & Pier Francesco Ferrari. 2010. "Towards a bottom-up perspective on animal and human cognition." *Trends in cognitive sciences* 14(5). 201–207.

Hauser, Marc D, Noam Chomsky & W Tecumseh Fitch. 2002. "The Faculty of Language: What is It, Who has It and How Did It Evolve?" *Science* 298. 1569–1579.

Hauser, Marc D, Charles Yang, Robert C Berwick, Ian Tattersall, Michael Ryan, Je rey Watumull, Noam Chomsky & Richard Lewontin. 2014. "The mystery of language evolution." *Frontiers in Psychology* 5(401).

Kirby, Simon. 2013. "Language, culture, and computation: An adaptive systems approach to biolinguistics." In Cedric Boeckx & Kleanthes K Grohmann (eds.), *The Cambridge Handbook of Biolinguistics*, 460–477. Cambridge: Cambridge University Press.

Moczek, A. P. (2008). "On the origins of novelty in development and evolution". *BioEssays*, 30(5):432–447.

Poeppel, D. and Embick, D. (2005). "De ning the relation between linguistics and neuroscience". In Cutler, A., editor, *Twenty-First Century Psycholinguistics*, pages 103–118. Lawrence Erlbaum, Mahwah, NJ/London.

Wagner, G. P. and Müller, G. B. (2002). "Evolutionary innovations overcome ancestral constraints: a re-examination of character evolution in male sepsid ies". *Evolution & Development*, 4(1):1–6.

West-Eberhard, Mary Jane. 2003. *Developmental plasticity and evolution*. Oxford: Oxford University Press.

— "Developmental plasticity and the origin of species di erences". *Proceedings of the National Academy of Sciences of the United States of America*, 102(Suppl 1):6543–6549.

Yang, Charles. 2010. "Three factors in language acquisition." *Lingua* 1160–1777.

Zilhão, João. 2011. "The emergence of language. art and symbolic thinking: A neandertal test of competing hypotheses." In Christopher S Henshilwood & Francesco d'Errico (eds.), *Homo symbolicus. The dawn of language, imagination and spirituality*, 111–131. Amsterdam: John Benjamins.

Appendix C

Do life histories shape vocal production learning differences in birds? An exploratory approach

DO LIFE HISTORIES SHAPE VOCAL PRODUCTION LEARNING DIFFERENCES IN BIRDS? AN EXPLORATORY APPROACH

SARA SILVENTE I FONT[*1], PEDRO TIAGO MARTINS[2,3], and CEDRIC BOECKX[2,3,4]

[*]Corresponding Author: sara.silvente@uclouvain.be
[1] Psychological Sciences Research Institute, University of Louvain, Louvain-la-Neuve, Belgium
[2]Section of General Linguistics, University of Barcelona, Barcelona, Spain
[3]University of Barcelona Institute of Complex Systems, Barcelona, Spain
[4]Catalan Institute for Research and Advanced Studies, Barcelona, Spain

1. Introduction

Vocal production learning (VPL) is the ability to change vocal output as a result of experience (e.g. auditory, be it through modification of spectral and/or temporal aspects of vocalizations or completely novel calls (see (Janik & Slater, 2000) for an influential definition). Despite the structural differences between the human cortex and the avian pallium, similar neural pathways have been proposed for VPL (Fitch, 2017; Jarvis, 2007), involving a direct forebrain (primary motor cortex/arcopallium) projection to the phonatory muscles (larynx/syrinx). Higher encephalization has been associated with complex behavioral traits (Isler & Schaik, 2009), including VPL (Liu, Wada, Jarvis, & Nottebohm, 2013). Charvet and Striedter (2011) proposed that the telenchephalic expasion undergone by VPL birds is the result of delayed and protracted neurogenesis in this brain region, with a major effect in some song nuclei during song learning (Liu et al., 2013). In turn, this form of post-hatching maturation of the telencephalin would be promoted by an altricial developmental mode (Charvet & Striedter, 2011). In summary, prolonged brain maturation is likely to lead to an extended developmental period that would benefit the development of complex behavioral traits including VPL. Defining VPL is a non-trivial issue, which we put aside in the present work, which is of an exploratory nature. With this in mind, we opt for a bottom-up approach in which we aim to explore whether differences between birds considered to be VPL and non-VPL naturally emerge from other data. For example, life histories are affected by body mass, such that larger species mature more slowly (Minias & Podlaszczuk, 2017). Avian species that learn their vocalizations tend to have relatively small body sizes (Liu et al., 2013). These species would be expected to group closer together by virtue of their small body size. The goals of the present work are the following: i) to explore whther groups reflecting divisions based on VPL/Non-VPL profiles emerge from avian life history traits (i.e, development

trajectories) and two biometric measures, brain mass and body mass; ii) to test whether differences in (adult) relative brain mass (measure here by the ration of brain-to-body mass) are related to life history traits and contribute to shaping the differences between the groups obtained above.

2. Methods

Data. Six developmental variables representative of life history were used: incubation operiod, fledging age, period of post-fledging parental care, age of sexual maturity for males and for females, and maximum lifespan. Absolute body size and brain mass were used as biometric variables, as well as the brain-to-body mass ratio (as an approximate measure of relative brain mass). **Sample.** These data were extracted for 1498 species (34 orders) from the literature and public databases (Hoyo et al., 1992; Iwaniuk & Nelson, 2003; Striedter & Charvet, 2008; Myhrvold et al., 2015). After excluding species with missing or unclear data, the final sample consists of 179 species (96 thought to be vocal learners) with several orders and families represented. **Analyses.** *clustering* In order to explore how VPL and non-VPL birds grouped as a function of their life histoies and biometric measures, we ran a hierarchical clustering analysis. Since developmental ctraits are correlated with one another, we specified Spearman correlation coefficients as the distance metric between clusters. Two hierarchical clusters were built by using the average and Ward's methods (average method: 0.674, Ward's method: 0.657). Non-hierarchical methods were also used to determine whether the groups obtained above emerged naturally from the data (e.g. k-metoid clustering using Partitioning Around Medoids (PAM). Different statistics were calculated over a range of 2 to 7 possuble k values. *Permutation*. In order to extend the descriptive results, we tested the different groups obtained for significant differences using a permutation test (n = 19999, no replacement). When statistically significant differences were found, Monto Carlo-based permutation was used (n = 9999) to assess the effect of random shuffling. Since avian life history traits are correlated with relative brain size, we tested whether the groups obtained in the hierarchical clustering analysis differed in life history traits when controlling for brain-to-body mass ratio (n = 10000).

3. Results & discussion

Our analyses yielded the following 3 meaningful clusters: VL (vocal leaners, 30 species), NVL (non-VL, 31 species), and a mix of both (MIX, 118 species, 66 vocal learners) (see supplementary materials). VL emerged as the most consistent group throughout our different analyses. Our findings suggest body size to range too widely to be a predictor of bird vocal VPL abilities. We found brain-to-body mass ratio to be related to all life history traits, but it has a significant effect in distinguishing the three groups when related to period of incubation and age of

sexual maturity. We discuss the implications of our results for the study of VPL in birds, and perhaps other species, as well as some cautionary notes.

References

Charvet, C. J., & Striedter, G. F. (2011). Developmental Modes and Developmental Mechanisms can Channel Brain Evolution. *Frontiers in Neuroanatomy*, *5*.

Fitch, W. T. (2017). Empirical approaches to the study of language evolution. *Psychonomic Bulletin & Review*, *24*(1), 3-33.

Hoyo, J. del, Elliott, A., Sargatal, J., Christie, D., International Council for Bird Preservation, & BirdLife International. (1992). *Handbook of the birds of the world.* Barcelona: Lynx [for] ICBP. (OCLC: 225496301)

Isler, K., & Schaik, C. P. van. (2009). The expensive brain: A framework for explaining evolutionary changes in brain size. *Journal of Human Evolution*, *57*(4), 392–400.

Iwaniuk, A. N., & Nelson, J. E. (2003). Developmental differences are correlated with relative brain size in birds: A comparative analysis. *Canadian Journal of Zoology*, *81*(12), 1913-1928.

Janik, V. M., & Slater, P. J. (2000). The different roles of social learning in vocal communication. *Animal Behaviour*, *60*(1), 1–11.

Jarvis, E. D. (2007). Neural systems for vocal learning in birds and humans: A synopsis. *Journal of Ornithology*, *148*(1), 35–44.

Liu, W.-c., Wada, K., Jarvis, E. D., & Nottebohm, F. (2013). Rudimentary substrates for vocal learning in a suboscine. *Nature Communications*, *4*, 2082.

Minias, P., & Podlaszczuk, P. (2017). Longevity is associated with relative brain size in birds. *Ecology and Evolution*, *7*(10), 3558–3566.

Myhrvold, N. P., Baldridge, E., Chan, B., Sivam, D., Freeman, D. L., & Ernest, S. K. M. (2015). An amniote life-history database to perform comparative analyses with birds, mammals, and reptiles: *Ecological Archives* E096-269. *Ecology*, *96*(11), 3109-000.

Striedter, G. F., & Charvet, C. J. (2008). Developmental origins of species differences in telencephalon and tectum size: Morphometric comparisons between a parakeet (*Melopsittacus undulatus*) and a quail (*Colinus virgianus*). *The Journal of Comparative Neurology*, *507*(5), 1663-1675.

Appendix D

Clarifications on the no half-Merge fallacy

Pedro Tiago Martins & Cedric Boeckx

Abstract: Berwick and Chomsky (2019) raise a number of issues with Martins and Boeckx (2019). In what follows, we will briefly outline and dispel these issues, as well as clarify some points which have been the object of misrepresentation.

D.1 Introduction

In a comment to our paper (Martins and Boeckx, 2019), Berwick and Chomsky (2019) make a number of claims that grossly misrepresent both our paper and our points of view on the the topic of language evolution more generally. In this brief formal comment we identify such claims and clarify why they are misleading.

D.2 The no half-Merge fallacy

Our paper points out what we call the "no half-Merge fallacy". This fallacy consists of deriving from the formal complexity of Merge the number of evolutionary steps it took for it to arise. Even though our motivation is to show that it is not biologically plausible or at the very least not warranted to claim that Merge arose in a single step (by means of single mutation), the fallacy is independent of the actual number of steps it took for Merge to arise. The fallacy is also independent of the actual existence of Merge. We use "fallacy" in the conventional sense meaning an argument which does not work either because of its logical structure or its content.

The first misguided assumption by Berwick and Chomsky (2019) is what they understand by "no half-MERGE fallacy". They take it to be a claim about whether such a thing as "half-MERGE" exists, which it is not. Instead, what we mean by it, as we state multiple times in our article (Martins and Boeckx, 2019) is that "from the formal simplicity of MERGE one cannot infer the evolutionary steps that led to it".

The premises that MERGE exists and that there can be no such thing as "half-MERGE" are in the domain of formal analysis, and the conclusion is about how it must have *evolved*, a completely different domain, biology.

1. There is such a thing as MERGE.

2. There cannot be such a thing as half-MERGE (MERGE is formally so simple is cannot be reduced)

∴. Therefore, MERGE had to evolve in a single step.

One hopes there are connections between the two (or else the formal analysis of mental processes would be in vain), but there is no necessary or warranted connection between formal simplicity and steps leading to biological emergence (the latter being furthermore difficult to define). It is fallacious to draw that connection as conceptually necessary. It is fallacious to claim that a mental operation that cannot be reduced formally definitely arose suddenly as the result of a single mutation. We used fairly conventional terminology in naming the fallacy the way we did. To take a similarly named fallacy (though different in its structure, but that is beside the point here), the "no True Scotsman fallacy" (Flew, 1975) says nothing about the existence of true Scotsmen, or Scotsmen for that matter. Or more generically, the Historian's fallacy (Fischer, 1970) says nothing about the existence of historians. The way informal fallacies are named is an allusion to the way arguments are formulated. Fallacy names are not claims, but rather shorthand descriptions or allusions to argumentative structures that do not hold for one reason or another. That's the approach we take here as well. A claim is made: "MERGE arose in a single step", others question it "why is that?", and the fallacy arises with the answer "well, there is no half-MERGE". To sum up: the argument would work if there were a correspondence between the formal structure of a computational operation and the biological changes that would lead to it. Since there isn't such a connection (as we defend in Martins and Boeckx 2019), it is fallacious to claim that the state of affairs of there not being "half-MERGE" shows that MERGE arose suddenly. Any criticism from Berwick and Chomsky (2019) of the "no half-MERGE fallacy" on the basis of the premises and not the unwarranted conclusion is seriously misguided. Much can be said about the premises too, of course, but that is outside of the scope of the fallacy.

Berwick and Chomsky (2019) also seem to assume that we attempt to replace their fallacious argumentation by claiming that MERGE evolved in two steps instead of one.

They do so already in their abstract: "Furthermore, their specific evolutionary scenario counterproposal for a "two-step" evolution of MERGE does not work." But as we say quite clearly in our paper, pre-empting this interpretation (Martins and Boeckx, 2019, 4): "To be clear, we are not suggesting that it actually took exactly two steps for MERGE to arise. We simply use Berwick and Chomsky's methodological approach to try to derive evolutionary steps by looking only at formal properties and conclude that these don't entail a single mutation." We simply show that there would be possibilities beyond a single-step scenario, highlighting the fallibility of their approach.

D.3   Agree to disagree

Berwick and Chomsky (2019) list five issues on which they supposedly agree with us, or us with them. But all of these are loaded statements attributing to us more than what we have said or let through with our paper (and in many cases indeed with our work in general).

D.3.1   Core properties of language and MERGE

Berwick and Chomsky (2019) say that we:

> "do not question [their] assumption that the core properties of language are based on the combinatorial operation MERGE".

This cannot be concluded from reading our paper, or at the very least is ambiguous. The point in our paper is independent of such sweeping statements about the "core properties of language", and if one is allowed to infer things that are not said, we can point to several articles of ours where we question precisely this point (refs), though we think that none of this is warranted, since we reason to defend ourselves from objections about things we haven't said, beyond pointing out that we indeed haven't said them.

D.3.2   Implementation of MERGE in the brain

Berwick and Chomsky (2019) say:

> "we both agree that it is important to determine how MERGE is implemented in the brain. [...] we advance a specific proposal about this neural "wiring," grounded on recent explicit neurological and comparative primate findings. [Martins & Boeckx] do not challenge this proposal. We therefore put the matter of neural implementation aside here."

While we think it's relevant to understand the brain implementation of anything that enters the realm of cognition, and particularly language, that is not specifically

what our paper is about, and we definitely do not agree with Berwick & Chomsky that a single rewiring of the brain yields Merge. Again, nothing in our paper is a claim in favor of Berwick & Chomsky's view on the implementation of Merge, much less Merge *qua* central property of language.

D.3.3   How a Merge-based system is used

Berwick and Chomsky (2019) say:

> "we both agree that it is important to determine how a Merge-based system is used, that is, how it is externalized in the sensory-motor system (typically, though not necessarily, sound) and then actually used in performance". They then list several ways in they discussed the architecture of such a Merge-based system, which they do in a chapter their book (Berwick and Chomsky, 2016), and how we fail to criticize or indeed mention any of this.

It is a reasonable expectation that we would not wish or be able to do an in-depth review of the contents of Berwick and Chomsky (2016) in our paper, which is about a problem of argumentation, and not an exhaustive evaluation of the body of work of Berwick and Chomsky or any others. Fulfilling this expectation cannot be used to argue that we agree with said contents.

Moreover, and again this is beside the point, we are of course aware that such matters are discussed in Berwick and Chomsky (2016) but it still stands that, whatever the algorithm, and whatever must be physically in place for that algorithm to be used, a multitude of structures and connections are required, which had to evolve, and cannot be attributed to a small change.

D.3.4   Fallacy or no fallacy?

The fourth claim of agreement is confusing. Berwick and Chomsky (2019) say:

> "We agree that there need not be, as [Martins & Boeckx] notes in its abstract, a "parallelism between the formal complexity of the operation at the computational level and the number of evolutionary steps it must imply." [...] We too regard it as "problematic" and, indeed, a "fallacy.""

This statement is confusing, given that what they defend in their paper and in their other work they point us to incurs this very fallacy.

They then go on to say:

> "What is under discussion is not operations in general but rather a specific one, the simplest combinatorial operation, binary set formation, called

> MERGE. Crucially, as we discuss next, MB's own proposal adopts our
> account of the evolution of MERGE unchanged, thus tacitly recognizing
> that binary set formation (MERGE) cannot be decomposed and emerges in
> a single step. MB then add new proposals about immediate precursors to
> our shared account of the evolution of MERGE. The justification for the
> added complexities that they propose about precursors to MERGE is the
> sole point at issue."

First, there is an assumption, left unexplained, that MERGE is somehow a special operation and somehow above the fallacy we describe. This, we do not understand. Simple operations such as MERGE are precisely the kind of entity that expose the "dangers" of simple and simplistic evolutionary scenarios. Secondly, they immediately incur the fallacy again, when they say we tacitly recognize that "binary set formation (MERGE) cannot be decomposed and emerges in a single step". The "and" conjunction in this statement is a huge leap, and it what our paper is about. That's our real issue; that's the fallacy.

The latter part of the quote is again taking our exercise of deriving more than one step for the evolution of MERGE as an actual proposal for its evolution, which we already deny in the original paper, earlier in this very same paper, and now again.

D.3.5 Long evolutionary history

Berwick and Chomsky (2019) say:

> "we both agree that it would be important to discover the long evolution-
> ary history that preceded the appearance of MERGE. [...] In this case,
> although both we and [1] agree that there were multiple steps that pre-
> ceded the appearance of MERGE, neither we nor [1] present any explicit
> proposals about these previous steps, so we can put this matter aside too."

Again, it would be far beyond the scope of our paper to present explicit proposals about what preceded MERGE, and the same applies to Berwick and Chomsky (2019). One would, however, be hard-pressed to identify this concern in the work they point us to (e.g. Berwick and Chomsky, 2016), which insists on MERGE as the core part of language: "Any residue of principles of language not reducible to MERGE and optimal computation will have to be accounted for by some other evolutionary process—one that we are unlikely to learn much about, at least by presently understood methods, as Lewontin (1998) notes." (Berwick and Chomsky, 2016, 72).

This is effectively the position that studying language evolution in a meaningful way can be reduced to studying MERGE. It is also not entirely persuasive to allude to work by Lewontin from 1998, more than two decades ago, as casting doubt on "presently understood methods" (which are in any case left unidentified).

D.3.6 A secondary issue

After going through our exercise in deriving an alternative analysis from the formal properties of Merge, Berwick and Chomsky (2019) conclude:

> The errors in Martins and Boeckx (2019) concerning emergence of EM [External Merge] and IM [Internal Merge] are, however, secondary. The crucial point is that the sole proposal in Martins and Boeckx (2019) about evolution of language is untenable. The "no half-Merge fallacy" analysis in Martins and Boeckx (2019) collapses because there is no such fallacy.

Our point would stand without our having ventured into the task of deriving more than one step from the formal properties of Merge, and this is what Berwick and Chomsky (2019) spend most of their paper on. We of course will have our bones to pick regarding how "correct" our analysis is as a formal exercise (that goes beyond the scope of this reply), but we do indeed wholeheartedly agree: all of this is secondary. There is no aspiration of tenability in our proposal, since we argue precisely that such proposals are untenable. One can get one step for the evolution of Merge by looking at its formal properties, or one can get two like we did. Most likely, other analyses could get different numbers of steps.

Finally, after at some point recognizing that the fallacy we point to exists (see section D.3.4), they now say there is no fallacy. But given their practice of repeatedly committing the fallacy, which is the real problem (and not what is said about it), this too is secondary.

D.4   Final Remarks

To conclude, the reply by Berwick and Chomsky (2019) does not counter the points in our paper, and we maintain our conclusions that 1) from the simplicity of a formal operation one cannot derive the evolutionary steps that led to it, and 2) doing so is not a productive way of understanding the evolution of language.

Bibliography

Anderson, S. R. and Lightfoot, D. W. (2000). The Human Language Faculty as an Organ. *Annual Review of Physiology*, 62(1):697–722.

Arriaga, G. and Jarvis, E. D. (2013). Mouse vocal communication system: Are ultrasounds learned or innate? *Brain and Language*, 124(1).

Arriaga, G., Zhou, E. P., and Jarvis, E. D. (2012). Of Mice, Birds, and Men: The Mouse Ultrasonic Song System Has Some Features Similar to Humans and Song-Learning Birds. *PLoS ONE*, 7(10):e46610.

Aspenström, P. (2008). Chapter 1 Roles of F-BAR/PCH Proteins in the Regulation of Membrane Dynamics and Actin Reorganization. *International Review of Cell and Molecular Biology*, 272:1–31.

Bacon, C., Endris, V., and Rappold, G. (2009). Dynamic expression of the Slit-Robo GTPase activating protein genes during development of the murine nervous system. *The Journal of Comparative Neurology*, 513(2):224–236.

Bae, C. J., Douka, K., and Petraglia, M. D. (2017). On the origin of modern humans: Asian perspectives. *Science*, 358(6368):eaai9067.

Baker, A., editor (2016). *The Linguistics of Sign Languages: An Introduction*. John Benjamins Pub. Company, Amsterdam ; Philadelphia.

Behme, C. (2015). Is the ontology of biolinguistics coherent? *Language Sciences*, 47(Part A):32–42.

Belyk, M. and Brown, S. (2017). The origins of the vocal brain in humans. *Neuroscience & Biobehavioral Reviews*, 77(Supplement C):177–193.

Benavides-Piccione, R., Ballesteros-Yáñez, I., DeFelipe, J., and Yuste, R. (2002 Mar-Jun). Cortical area and species differences in dendritic spine morphology. *Journal of Neurocytology*, 31(3-5):337–346.

Berwick, R. C. and Chomsky, N. (2016). *Why Only Us; Language and Evolution*. MIT Press, Cambridge, MA.

Berwick, R. C. and Chomsky, N. (2019). All or nothing: No half-Merge and the evolution of syntax. *PLoS biology*, 17(11):e3000539.

Berwick, R. C., Friederici, A. D., Chomsky, N., and Bolhuis, J. J. (2013). Evolution,

brain, and the nature of language. *Trends in Cognitive Sciences*, 17(2):89–98.

Berwick, R. C., Okanoya, K., Beckers, G. J. L., and Bolhuis, J. J. (2011). Songs to syntax: The linguistics of birdsong. *Trends in Cognitive Sciences*, 15(3):113–121.

Bianchi, S., Reyes, L. D., Hopkins, W. D., Taglialatela, J. P., and Sherwood, C. C. (2016). Neocortical grey matter distribution underlying voluntary, flexible vocalizations in chimpanzees. *Scientific Reports*, 6:34733.

Bickerton, D. (2002). *From Protolanguage to Language: The Speciation of Modern Homo Sapiens*. Oxford University Press, Oxford.

Bickerton, D. (2014). Some Problems for Biolinguistics. *Biolinguistics*, 8(0):073–096.

Boeckx, C. (2013). Biolinguistics: Forays into human cognitive biology. *Journal of anthropological sciences*, 91:63–89.

Boeckx, C. (2017a). Language Evolution. In Kaas, J., editor, *Evolution of Nervous Systems, 2nd Ed., Vol. 4*, pages 325–339. Elsevier, London.

Boeckx, C. (2017b). The language-ready head: Evolutionary considerations. *Psychonomic Bulletin & Review*, 24(1):194–199.

Boeckx, C. and Grohmann, K. K. (2007). The Biolinguistics Manifesto. *Biolinguistics*, 1(0):001–008.

Bolhuis, J. J., Tattersall, I., Chomsky, N., and Berwick, R. C. (2014). How Could Language Have Evolved? *PLOS Biology*, 12(8):e1001934.

Bolhuis, J. J., Tattersall, I., Chomsky, N., and Berwick, R. C. (2015). Language: UG or Not to Be, That Is the Question. *PLOS Biology*, 13(2):e1002063.

Brenowitz, E. A. and Beecher, M. D. (2005). Song learning in birds: Diversity and plasticity, opportunities and challenges. *Trends in Neurosciences*, 28(3):127–132.

Brose, K., Bland, K. S., Wang, K. H., Arnott, D., Henzel, W., Goodman, C. S., Tessier-Lavigne, M., and Kidd, T. (1999). Slit Proteins Bind Robo Receptors and Have an Evolutionarily Conserved Role in Repulsive Axon Guidance. *Cell*, 96(6):795–806.

Brunet, M., Guy, F., Pilbeam, D., Lieberman, D. E., Likius, A., Mackaye, H. T., de León, M. S. P., Zollikofer, C. P. E., and Vignaud, P. (2005). New material of the earliest hominid from the Upper Miocene of Chad. *Nature*, 434(7034):752.

Brunet, M., Guy, F., Pilbeam, D., Mackaye, H. T., Likius, A., Ahounta, D., Beauvilain, A., Blondel, C., Bocherens, H., Boisserie, J.-R., Bonis, L. D., Coppens, Y., Dejax, J., Denys, C., Duringer, P., Eisenmann, V., Fanone, G., Fronty, P., Geraads, D., Lehmann, T., Lihoreau, F., Louchart, A., Mahamat, A., Merceron, G., Mouchelin, G., Otero, O., Campomanes, P. P., Leon, M. P. D., Rage, J.-C., Sapanet, M., Schuster, M., Sudre, J., Tassy, P., Valentin, X., Vignaud, P., Viriot, L., Zazzo, A., and Zollikofer, C. (2002). A new hominid from the Upper Miocene of Chad, Central Africa. *Nature*, 418(6894):145.

Buckner, R. L. and Krienen, F. M. (2013). The evolution of distributed association networks in the human brain. *Trends in Cognitive Sciences*, 17(12):648–665.

Candiotti, A., Zuberbuhler, K., and Lemasson, A. (2012). Convergence and divergence in Diana monkey vocalizations. *Biology Letters*, 8(3):382–385.

Chakraborty, M. and Jarvis, E. D. (2015). Brain evolution by brain pathway duplication. *Phil. Trans. R. Soc. B*, 370(1684):20150056.

Chakraborty, M., Wallø e, S., Nedergaard, S., Fridel, E. E., Dabelsteen, T., Pakkenberg, B., Bertelsen, M. F., Dorrestein, G. M., Brauth, S. E., Durand, S. E., and Jarvis, E. D. (2015). Core and Shell Song Systems Unique to the Parrot Brain. *PLOS ONE*, 10(6):e0118496.

Chamberlain, S. A. and Szöcs, E. (2013). Taxize: Taxonomic search and retrieval in R. *F1000Research*, 2:191.

Charrier, C., Joshi, K., Coutinho-Budd, J., Kim, J.-E., Lambert, N., de Marchena, J., Jin, W.-L., Vanderhaeghen, P., Ghosh, A., Sassa, T., and Polleux, F. (2012). Inhibition of SRGAP2 function by its human-specific paralogs induces neoteny during spine maturation. *Cell*, 149(4):923–935.

Charvet, C. J. and Striedter, G. F. (2011). Developmental Modes and Developmental Mechanisms can Channel Brain Evolution. *Frontiers in Neuroanatomy*, 5.

Chen, Y.-C., Kuo, H.-Y., Bornschein, U., Takahashi, H., Chen, S.-Y., Lu, K.-M., Yang, H.-Y., Chen, G.-M., Lin, J.-R., Lee, Y.-H., Chou, Y.-C., Cheng, S.-J., Chien, C.-T., Enard, W., Hevers, W., Pääbo, S., Graybiel, A. M., and Liu, F.-C. (2016). *Foxp2* controls synaptic wiring of corticostriatal circuits and vocal communication by opposing *Mef2c*. *Nature Neuroscience*, 19(11):1513.

Cheney, D. L. and Seyfarth, R. M. (2018). Flexible usage and social function in primate vocalizations. *Proceedings of the National Academy of Sciences*, page 201717572.

Choi, J. Y., Takahashi, D. Y., and Ghazanfar, A. A. (2015). Cooperative vocal control in marmoset monkeys via vocal feedback. *Journal of Neurophysiology*, 114(1):274–283.

Chomsky, N. (1956). Three models for the description of language. *IRE Transactions on Information Theory*, 2(3):113–124.

Chomsky, N. (1957). *Syntactic Structures*. Mouton, The Hague.

Chomsky, N. (1959). Review of Verbal behavior. *Language*, 35(1):26–58.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. MIT Press, Cambridge, MA.

Chomsky, N. (1995). *The Minimalist Program*. MIT Press, Cambridge, MA.

Chomsky, N. (2000). *The Architecture of Language*. Oxford University Press, Oxford.

Chomsky, N. (2007). Approaching UG from below. In Sauerland, U. and Gärtner, H.-M., editors, *Interfaces + Recursion = Language? Chomsky's Minimalism and the View from Semantics*, pages 1–30. Mouton de Gruyter, Berlin.

Chomsky, N. (2010). Some simple evo devo theses: How true might they be for language. In Larson, R. K. L., Déprez, V., and Yamakido, H., editors, *The Evolution of Language: Biolinguistic Perspetives*, pages 45–62. Cambridge University Press, Cambridge.

Chomsky, N. (2012). *The Science of Language: Interviews with James McGilvray*. Cambridge University Press, Cambridge.

Coutinho-Budd, J., Ghukasyan, V., Zylka, M. J., and Polleux, F. (2012). The F-BAR domains from srGAP1, srGAP2 and srGAP3 regulate membrane deformation differently. *Journal of Cell Science*, 125(14):3390–3401.

Crockford, C., Herbinger, I., Vigilant, L., and Boesch, C. (2004). Wild Chimpanzees Produce Group-Specific Calls: A Case for Vocal Learning? *Ethology*, 110(3):221–243.

Cvijovic, M., Höfer, T., Aćimović, J., Alberghina, L., Almaas, E., Besozzi, D., Blomberg, A., Bretschneider, T., Cascante, M., Collin, O., de Atauri, P., Depner, C., Dickinson, R., Dobrzynski, M., Fleck, C., Garcia-Ojalvo, J., Gonze, D., Hahn, J., Hess, H. M., Hollmann, S., Krantz, M., Kummer, U., Lundh, T., Martial, G., dos Santos, V. M., Mauer-Oberthür, A., Regierer, B., Skene, B., Stalidzans, E., Stelling, J., Teusink, B., Workman, C. T., and Hohmann, S. (2016). Strategies for structuring interdisciplinary education in Systems Biology: An European perspective. *npj Systems Biology and Applications*, 2:16011.

Dawkins, R. (2015). *Brief Candle in the Dark: My Life in Science*. Ecco, New York, NY.

de Boer, B., Thompson, B., Ravignani, A., and Boeckx, C. (2020). Evolutionary Dynamics Do Not Motivate a Single-Mutant Theory of Human Language. *Scientific Reports*, 10(1):451.

de la Torre, S. and Snowdon, C. T. (2009). Dialects in pygmy marmosets? Population variation in call structure. *American Journal of Primatology*, 71(4):333–342.

de Waal, F. and Ferrari, P. F. (2010). Towards a bottom-up perspective on animal and human cognition. *Trends in Cognitive Sciences*, 14(5):201–207.

Deacon, T. W. (1997). *The Symbolic Species*. Norton, New York.

Dediu, D. and Levinson, S. C. (2013). On the antiquity of language: The reinterpretation of Neandertal linguistic capacities and its consequences. *Frontiers in Psychology*, 4.

del Hoyo, J., Elliott, A., Sargatal, J., Christie, D., International Council for Bird Preservation, and BirdLife International (1992). *Handbook of the birds of the world*. Lynx [for] ICBP, Barcelona.

Dennis, M. Y., Nuttle, X., Sudmant, P. H., Antonacci, F., Graves, T. A., Nefedov, M., Rosenfeld, J. A., Sajjadian, S., Malig, M., Kotkiewicz, H., Curry, C. J., Shafer, S., Shaffer, L. G., de Jong, P. J., Wilson, R. K., and Eichler, E. E. (2012). Evolution of human-specific neural SRGAP2 genes by incomplete segmental duplication. *Cell*, 149(4):912–922.

Dent, E. W. and Gertler, F. B. (2003). Cytoskeletal Dynamics and Transport in Growth Cone Motility and Axon Guidance. *Neuron*, 40(2):209–227.

D'Errico, F. (2003). The invisible frontier. A multiple species model for the origin

of behavioral modernity. *Evolutionary Anthropology: Issues, News, and Reviews*, 12(4):188–202.

Devanna, P., Middelbeek, J., and Vernes, S. C. (2014). FOXP2 drives neuronal differentiation by interacting with retinoic acid signaling pathways. *Frontiers in Cellular Neuroscience*, 8.

Di Sciullo, A. M. (2011). A Biolinguistic Approach to Variation. In Di Sciullo, A. M. and Boeckx, C., editors, *The Biolinguistic Enterprise: New Perspectives on the Evolution and Nature of the Human Language Faculty*, pages 305–326. Oxford University Press, Oxford.

Di Sciullo, A. M. (2012). Interfaces in a biolinguistic perspective. In Di Sciullo, A. M., editor, *Towards a Biolinguistic Understanding of Grammar: Essays on Interfaces*. John Benjamins Publishing, Amsterdam.

Di Sciullo, A. M. and Somesfalean, S. (2015). Object Pronouns in the Evolution of Romanian: A Biolinguistic Perspective. In Hill, V., editor, *Formal Approaches to DPs in Old Romanian*, pages 269–289. Brill, Leiden.

Dickson, B. J. (2002). Molecular Mechanisms of Axon Guidance. *Science*, 298(5600):1959–1964.

Ding, N., Melloni, L., Zhang, H., Tian, X., and Poeppel, D. (2016). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience*, 19(1):158.

Elowson, A. M. and Snowdon, C. T. (1994). Pygmy marmosets, Cebuella pygmaea, modify vocal structure in response to changed social environment. *Animal Behaviour*, 47(6):1267–1277.

Embick, D. and Poeppel, D. (2015). Towards a computational(ist) neurobiology of language: Correlational, integrated, and explanatory neurolinguistics. *Language, Cognition and Neuroscience*, 30(4):357–366.

Enard, W., Gehre, S., Hammerschmidt, K., Hölter, S. M., Blass, T., Somel, M., Brückner, M. K., Schreiweis, C., Winter, C., Sohr, R., Becker, L., Wiebe, V., Nickel, B., Giger, T., Müller, U., Groszer, M., Adler, T., Aguilar, A., Bolle, I., Calzada-Wack, J., Dalke, C., Ehrhardt, N., Favor, J., Fuchs, H., Gailus-Durner, V., Hans, W., Hölzlwimmer, G., Javaheri, A., Kalaydjiev, S., Kallnik, M., Kling, E., Kunder, S., Mossbrugger, I., Naton, B., Racz, I., Rathkolb, B., Rozman, J., Schrewe, A., Busch, D. H., Graw, J., Ivandic, B., Klingenspor, M., Klopstock, T., Ollert, M., Quintanilla-Martinez, L., Schulz, H., Wolf, E., Wurst, W., Zimmer, A., Fisher, S. E., Morgenstern, R., Arendt, T., de Angelis, M. H., Fischer, J., Schwarz, J., and Pääbo, S. (2009). A humanized version of Foxp2 affects cortico-basal ganglia circuits in mice. *Cell*, 137(5):961–971.

Everaert, M. B., Huybregts, M. A., Berwick, R. C., Chomsky, N., Tattersall, I., Moro, A., and Bolhuis, J. J. (2017). What is Language and How Could it Have Evolved? *Trends in Cognitive Sciences*, 21(8):569–571.

Everaert, M. B. H., Huybregts, M. A. C., Chomsky, N., Berwick, R. C., and Bolhuis, J. J. (2015). Structures, Not Strings: Linguistics as Part of the Cognitive Sciences. *Trends in Cognitive Sciences*, 19(12):729–743.

Ey, E., Rahn, C., Hammerschmidt, K., and Fischer, J. (2009). Wild Female Olive Baboons Adapt their Grunt Vocalizations to Environmental Conditions. *Ethology*, 115(5):493–503.

Feenders, G., Liedvogel, M., Rivas, M., Zapka, M., Horita, H., Hara, E., Wada, K., Mouritsen, H., and Jarvis, E. D. (2008). Molecular Mapping of Movement-Associated Areas in the Avian Brain: A Motor Theory for Vocal Learning Origin. *PLOS ONE*, 3(3):e1768.

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(7246):564–568.

Ferhat, A.-T., Torquet, N., Le Sourd, A.-M., de Chaumont, F., Olivo-Marin, J.-C., Faure, P., Bourgeron, T., and Ey, E. (2016). Recording Mouse Ultrasonic Vocalizations to Evaluate Social Communication. *Journal of Visualized Experiments*, 112:53871.

Fischer, D. H. (1970). *Historians' Fallacies.* Harper & Row, New York, NY.

Fisher, S. E. (2017). Evolution of language: Lessons from the genome. *Psychonomic Bulletin & Review*, 24(1):34–40.

Fisher, S. E. and Vernes, S. C. (2015). Genetics and the Language Sciences. *Annual Review of Linguistics*, 1(1):289–310.

Fitch, W. T. (2009). Prolegomena to a Future Science of Biolinguistics. *Biolinguistics*, 3(4):283–320.

Fitch, W. T. (2010). *The Evolution of Language.* Cambridge University Press, Cambridge.

Fitch, W. T. (2011a). The Evolution of Syntax: An Exaptationist Perspective. *Frontiers in Evolutionary Neuroscience*, 3.

Fitch, W. T. (2011b). Unity and diversity in human language. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1563):376–388.

Fitch, W. T. (2017a). Empirical approaches to the study of language evolution. *Psychonomic Bulletin & Review*, 24(1):3–33.

Fitch, W. T. (2017b). Special Issue on the Biology and Evolution of Language. *Psychonomic Bulletin & Review*, 24(21).

Fitch, W. T. (2018). The Biology and Evolution of Speech: A Comparative Analysis. *Annual Review of Linguistics*, 4(1):255–279.

Fitch, W. T., de Boer, B., Mathur, N., and Ghazanfar, A. A. (2016). Monkey vocal tracts are speech-ready. *Science Advances*, 2(12):e1600723.

Fitch, W. T., Huber, L., and Bugnyar, T. (2010). Social Cognition and the Evolution of Language: Constructing Cognitive Phylogenies. *Neuron*, 65(6):795–814.

Fitch, W. T. and Jarvis, E. D. (2013). Birdsong and other animal models for human speech, song, and vocal learning. In Arbib, M., editor, *Language, Music, and the Brain*, pages 499–539. MIT Press, Cambridge, MA.

Fitch, W. T. and Zuberbühler, K. (2013). Primate precursors to human language: Beyond discontinuity. In Altenmüller, E., Schmidt, S., and Zimmermann, E., editors, *Evolution of Emotional Communication: From Sounds in Nonhuman Mammals to Speech and Music in Man*, pages 26–48. Oxford University Press.

Flew, A. (1975). *Thinking about Thinking*. Frontina Press, Glasgow, UK.

Fossati, M., Pizzarelli, R., Schmidt, E. R., Kupferman, J. V., Stroebel, D., Polleux, F., and Charrier, C. (2016). SRGAP2 and Its Human-Specific Paralog Co-Regulate the Development of Excitatory and Inhibitory Synapses. *Neuron*, 91(2):356–369.

Friederici, A. D., Chomsky, N., Berwick, R. C., Moro, A., and Bolhuis, J. J. (2017). Language, mind and brain. *Nature Human Behaviour*, 1(10):713.

Gardner, T. J., Naef, F., and Nottebohm, F. (2005). Freedom and rules: The acquisition and reprogramming of a bird's learned song. *Science (New York, N.Y.)*, 308(5724):1046–1049.

Gaub, S., Groszer, M., Fisher, S. E., and Ehret, G. (2010). The structure of innate vocalizations in Foxp2-deficient mouse pups. *Genes, Brain and Behavior*, 9(4):390–401.

Ghazanfar, A. A., Liao, D. A., and Takahashi, D. Y. (2019). Volition and learning in primate vocal behaviour. *Animal Behaviour*.

Gill, F. and Donsker, D., editors (2019). *IOC World Bird List (v9.2)*.

Giraud, A.-L. and Poeppel, D. (2012). Cortical oscillations and speech processing: Emerging computational principles and operations. *Nature Neuroscience*, 15(4):511.

Givón, T. (2002). *Bio-Linguistics: The Santa Barbara Lectures*. John Benjamins, Amsterdam/Philadelphia.

Grimsley, J. M. S., Monaghan, J. J. M., and Wenstrup, J. J. (2011). Development of Social Vocalizations in Mice. *PLoS ONE*, 6(3):e17460.

Guerrier, S., Coutinho-Budd, J., Sassa, T., Gresset, A., Jordan, N. V., Chen, K., Jin, W.-L., Frost, A., and Polleux, F. (2009). The F-BAR domain of srGAP2 induces membrane protrusions required for neuronal migration and morphogenesis. *Cell*, 138(5):990–1004.

Guez-Haddad, J., Sporny, M., Sasson, Y., Gevorkyan-Airapetov, L., Lahav-Mankovski, N., Margulies, D., Radzimanowski, J., and Opatowsky, Y. (2015). The neuronal migration factor srGAP2 achieves specificity in ligand binding through a two-component molecular mechanism. *Structure*, 23(11):1989–2000.

Gultekin, Y. B. and Hage, S. R. (2017). Limiting parental feedback disrupts vocal development in marmoset monkeys. *Nature Communications*, 8:14046.

Hage, S. R., Gavrilov, N., and Nieder, A. (2013). Cognitive control of distinct vocalizations in rhesus monkeys. *Journal of Cognitive Neuroscience*, 25(10):1692–1701.

Hage, S. R., Gavrilov, N., and Nieder, A. (2016). Developmental changes of cognitive vocal control in monkeys. *The Journal of Experimental Biology*, 219(11):1744–1749.

Hall, A. (1998). Rho GTPases and the Actin Cytoskeleton. *Science*, 279(5350):509–514.

Hardus, M. E., Lameira, A. R., Schaik, C. P. V., and Wich, S. A. (2009). Tool use in wild orang-utans modifies sound production: A functionally deceptive innovation? *Proceedings of the Royal Society of London B: Biological Sciences*, 276(1673):3689–3694.

Hasson, U. and Frith, C. D. (2016). Mirroring and beyond: Coupled dynamics as a generalized framework for modelling social interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*.

Hasson, U., Ghazanfar, A. A., Galantucci, B., Garrod, S., and Keysers, C. (2012). Brain-to-brain coupling: A mechanism for creating and sharing a social world. *Trends in Cognitive Sciences*, 16(2):114–121.

Hast, M., Fischer, J., Wetzel, A., and Thompson, V. (1974). Cortical motor representation of the laryngeal muscles in Macaca mulatta. *Brain research*, 73(2):229–240.

Hast, M. H. and Milojkvic, R. (1966). The Response of the Vocal Folds to Electrical Stimulation of the Inferior Frontal Cortex of the Squirrel Monkey. *Acta Oto-Laryngologica*, 61(1-6):196–204.

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(5598):1569–1579.

Hauser, M. D., Yang, C., Berwick, R. C., Tattersall, I., Ryan, M. J., Watumull, J., Chomsky, N., and Lewontin, R. C. (2014). The mystery of language evolution. *Frontiers in Psychology*, 5.

Hayes, K. J. and Hayes, C. (1951). The Intellectual Development of a Home-Raised Chimpanzee. *Proceedings of the American Philosophical Society*, 95(2):105–109.

Heeren, W. F. L. (2015). Vocalic correlates of pitch in whispered versus normal speech. *The Journal of the Acoustical Society of America*, 138(6):3800–3810.

Henn, B. M., Steele, T. E., and Weaver, T. D. (2018). Clarifying distinct models of modern human origins in Africa. *Current Opinion in Genetics & Development*, 53:148–156.

Henshilwood, C. S. and Marean, C. W. (2003). The Origin of Modern Human Behavior: Critique of the Models and Their Test Implications. *Current Anthropology*, 44(5):627–651.

Herculano-Houzel, S. (2012). The remarkable, yet not extraordinary, human brain as a scaled-up primate brain and its associated cost. *Proceedings of the National Academy of Sciences*, 109(Supplement 1):10661–10668.

Herculano-Houzel, S. (2016). *The Human Advantage*. MIT Press, Cambridge, MA.

Hillert, D. (2015a). *The Premodern Language-Ready Brain of Our Common Ancestor*.

Hillert, D. G. (2015b). On the Evolving Biology of Language. *Frontiers in Psychology*,

6.

Hoffmann, S., Trost, L., Voigt, C., Leitner, S., Lemazina, A., Sagunsky, H., Abels, M., Kollmansperger, S., Maat, A. T., and Gahr, M. (2019). Duets recorded in the wild reveal that interindividually coordinated motor control enables cooperative behavior. *Nature Communications*, 10(1):2577.

Holy, T. E. and Guo, Z. (2005). Ultrasonic Songs of Male Mice. *PLoS Biology*, 3(12):e386.

Hopkins, W. D., Taglialatela, J. P., and Leavens, D. A. (2007). Chimpanzees differentially produce novel vocalizations to capture the attention of a human. *Animal Behaviour*, 73(2):281–286.

Hotchkin, C. F., Parks, S. E., and Weiss, D. J. (2015). Noise-Induced Frequency Modifications of Tamarin Vocalizations: Implications for Noise Compensation in Nonhuman Primates. *PLOS ONE*, 10(6):e0130211.

Ip, B. K., Bayatti, N., Howard, N. J., Lindsay, S., and Clowry, G. J. (2011). The Corticofugal Neuron-Associated Genes ROBO1, SRGAP1, and CTIP2 Exhibit an Anterior to Posterior Gradient of Expression in Early Fetal Human Neocortex Development. *Cerebral Cortex*, 21(6):1395–1407.

Isler, K. and van Schaik, C. P. (2009). The Expensive Brain: A framework for explaining evolutionary changes in brain size. *Journal of Human Evolution*, 57(4):392–400.

Iwaniuk, A. N. and Nelson, J. E. (2003). Developmental differences are correlated with relative brain size in birds: A comparative analysis. *Canadian Journal of Zoology*, 81(12):1913–1928.

Jackendoff, R. (2011). What is the human language faculty?: Two views. *Language*, 87(3):586–624.

Janik, V. M. (2014). Cetacean vocal learning and communication. *Current Opinion in Neurobiology*, 28:60–65.

Janik, V. M. and Slater, P. J. B. (2000). The different roles of social learning in vocal communication. *Animal Behaviour*, 60(1):1–11.

Jarvis, E. D. (2004). Learned Birdsong and the Neurobiology of Human Language. *Annals of the New York Academy of Sciences*, 1016(1):749–777.

Jarvis, E. D. (2019). Evolution of vocal learning and spoken language. *Science*, 366(6461):50–54.

Jarvis, E. D. and Mello, C. V. (2000). Molecular mapping of brain areas involved in parrot vocal communication. *Journal of Comparative Neurology*, 419(1):1–31.

Jarvis, E. D., Yu, J., Rivas, M. V., Horita, H., Feenders, G., Whitney, O., Jarvis, S. C., Jarvis, E. R., Kubikova, L., Puck, A. E., et al. (2013). Global view of the functional molecular organization of the avian cerebrum: Mirror images and functional columns. *Journal of Comparative Neurology*, 521(16):3614–3665.

Johnson, M. (2017). Marr's levels and the minimalist program. *Psychonomic Bulletin & Review*, 24(1):171–174.

Joseph, L., Toon, A., Schirtzinger, E. E., Wright, T. F., and Schodde, R. (2012). A revised nomenclature and classification for family-group taxa of parrots (Psittaciformes). *Zootaxa*, 3205(2).

Jürgens, U. (2002). Neural pathways underlying vocal control. *Neuroscience & Biobehavioral Reviews*, 26(2):235–258.

Jürgens, U. (2009). The Neural Control of Vocalization in Mammals: A Review. *Journal of Voice*, 23(1):1–10.

Jürgens, U. and von Cramon, D. (1982). On the role of the anterior cingulate cortex in phonation: A case report. *Brain and Language*, 15(2):234–248.

Kalan, A. K., Mundry, R., and Boesch, C. (2015). Wild chimpanzees modify food call structure with respect to tree size for a particular fruit species. *Animal Behaviour*, 101:1–9.

Kassen, R. and Bataillon, T. (2006). Distribution of fitness effects among beneficial mutations before selection in experimental populations of bacteria. *Nature Genetics*, 38(4):484–488.

Katz, P. S. (2016). Evolution of central pattern generators and rhythmic behaviours. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1685):20150057.

Kearney, M. G., Warren, T. L., Hisey, E., Qi, J., and Mooney, R. (2019). Discrete Evaluative and Premotor Circuits Enable Vocal Learning in Songbirds. *Neuron*, 104(3):559–575.e6.

Kikusui, T., Nakanishi, K., Nakagawa, R., Nagasawa, M., Mogi, K., and Okanoya, K. (2011). Cross Fostering Experiments Suggest That Mice Songs Are Innate. *PLoS ONE*, 6(3):e17721.

Kimura, M. (1980). Average time until fixation of a mutant allele in a finite population under continued mutation pressure: Studies by analytical, numerical, and pseudo-sampling methods. *Proceedings of the National Academy of Sciences*, 77(1):522–526.

Koda, H., Lemasson, A., Oyakawa, C., Rizaldi, Pamungkas, J., and Masataka, N. (2013). Possible Role of Mother-Daughter Vocal Interactions on the Development of Species-Specific Song in Gibbons. *PLoS ONE*, 8(8):e71432.

Konishi, M. (1965). The role of auditory feedback in the control of vocalization in the white-crowned sparrow. *Zeitschrift Für Tierpsychologie*, 22(7):770–783.

Konopka, G., Bomar, J. M., Winden, K., Coppola, G., Jonsson, Z. O., Gao, F., Peng, S., Preuss, T. M., Wohlschlegel, J. A., and Geschwind, D. H. (2009). Human-specific transcriptional regulation of CNS development genes by FOXP2. *Nature*, 462(7270):213.

Konopka, G., Friedrich, T., Davis-Turak, J., Winden, K., Oldham, M. C., Gao, F., Chen, L., Wang, G.-Z., Luo, R., Preuss, T. M., and Geschwind, D. H. (2012). Human-Specific Transcriptional Networks in the Brain. *Neuron*, 75(4):601–617.

Kos, M., van den Brink, D., Snijders, T. M., Rijpkema, M., Franke, B., Fernandez, G., and Hagoort, P. (2012). CNTNAP2 and Language Processing in Healthy Individuals as Measured with ERPs. *PLOS ONE*, 7(10):e46995.

Krakauer, J. W., Ghazanfar, A. A., Gomez-Marin, A., MacIver, M. A., and Poeppel, D. (2017). Neuroscience Needs Behavior: Correcting a Reductionist Bias. *Neuron*, 93(3):480–490.

Krause, J., Lalueza-Fox, C., Orlando, L., Enard, W., Green, R. E., Burbano, H. A., Hublin, J.-J., Hänni, C., Fortea, J., de la Rasilla, M., Bertranpetit, J., Rosas, A., and Pääbo, S. (2007). The Derived FOXP2 Variant of Modern Humans Was Shared with Neandertals. *Current Biology*, 17(21):1908–1912.

Kuhlwilm, M. and Boeckx, C. (2019). A catalog of single nucleotide changes distinguishing modern humans from archaic hominins. *Scientific Reports*, 9(1):8463.

Kuypers, H. G. J. M. (1958). Some projections from the peri-central cortex to the pons and lower brain stem in monkey and chimpanzee. *Journal of Comparative Neurology*, 110(2):221–255.

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(6855):519.

Lameira, A. R. (2017). Bidding evidence for primate vocal learning and the cultural substrates for speech evolution. *Neuroscience & Biobehavioral Reviews*, 83(Supplement C):429–439.

Lameira, A. R., Hardus, M. E., Bartlett, A. M., Shumaker, R. W., Wich, S. A., and Menken, S. B. J. (2015). Speech-Like Rhythm in a Voiced and Voiceless Orangutan Call. *PLoS ONE*, 10(1):e116136.

Lameira, A. R., Hardus, M. E., Kowalsky, B., de Vries, H., Spruijt, B. M., Sterck, E. H. M., Shumaker, R. W., and Wich, S. A. (2013). Orangutan (*Pongo* spp.) whistling and implications for the emergence of an open-ended call repertoire: A replication and extension. *The Journal of the Acoustical Society of America*, 134(3):2326–2335.

Lameira, A. R., Hardus, M. E., Mielke, A., Wich, S. A., and Shumaker, R. W. (2016). Vocal fold control beyond the species-specific repertoire in an orang-utan. *Scientific Reports*, 6:30315.

Lameira, A. R., Hardus, M. E., and Wich, S. A. (2012). Orangutan Instrumental Gesture-Calls: Reconciling Acoustic and Gestural Speech Evolution Models. *Evolutionary Biology*, 39(3):415–418.

Lameira, A. R., Maddieson, I., and Zuberbühler, K. (2014). Primate feedstock for the evolution of consonants. *Trends in Cognitive Sciences*, 18(2):60–62.

Lameira, A. R. and Shumaker, R. W. (2019). Orangutans show active voicing through a membranophone. *Scientific Reports*, 9(1):1–6.

Lattenkamp, E. Z. and Vernes, S. C. (2018). Vocal learning: A language-relevant trait

in need of a broad cross-species approach. *Current Opinion in Behavioral Sciences*, 21:209–215.

Le Floch, É., Lalanne, C., Frouin, V., Pinel, P., Trinchera, L., Tenenhaus, A., Moreno, A., Zilbovicius, M., Bourgeron, T., Dehaene, S., et al. (2012). Significant correlation between a set of genetic polymorphisms and a functional brain network revealed by feature selection and sparse Partial Least Squares. *Neuroimage*.

Lemasson, A., Glas, L., Barbu, S., Lacroix, A., Guilloux, M., Remeuf, K., and Koda, H. (2011a). Youngsters do not pay attention to conversational rules: Is this so for nonhuman primates? *Scientific Reports*, 1:22.

Lemasson, A., Ouattara, K., Petit, E. J., and Zuberbühler, K. (2011b). Social learning of vocal structure in a nonhuman primate? *BMC Evolutionary Biology*, 11:362.

Lenneberg, E. H. (1964). A Biological Perspective of Language. In Lenneberg, E. H., editor, *New Directions in The Study of Language*. MIT Press, Cambridge, MA.

Lenneberg, E. H. (1967). *Biological Foundations of Language*. Wiley, New York.

Levinson, S. C. (2016). Turn-taking in Human Communication – Origins and Implications for Language Processing. *Trends in Cognitive Sciences*, 20(1):6–14.

Levréro, F., Carrete-Vega, G., Herbert, A., Lawabi, I., Courtiol, A., Willaume, E., Kappeler, P. M., and Charpentier, M. (2015). Social shaping of voices does not impair phenotype matching of kinship in mandrills. *Nature Communications*, 6(1):7609.

Lewontin, R. C. (1998). The evolution of cognition: Questions we will never answer. In Scarborough, D. and Sternberg, S., editors, *An Invitation to Cognitive Science, 2nd Edition*, pages 107–132. MIT Press, Cambridge, MA.

Leyton, A. S. F. and Sherrington, C. S. (1917). Observations on the excitable cortex of the chimpanzee, orang-utan, and gorilla. *Quarterly Journal of Experimental Physiology*, 11(2):135–222.

Li, X., Chen, Y., Liu, Y., Gao, J., Gao, F., Bartlam, M., Wu, J. Y., and Rao, Z. (2006). Structural basis of Robo proline-rich motif recognition by the srGAP1 Src homology 3 domain in the Slit-Robo signaling pathway. *Journal of Biological Chemistry*, 281(38):28430–28437.

Liao, D. A., Zhang, Y. S., Cai, L. X., and Ghazanfar, A. A. (2018). Internal states and extrinsic factors both determine monkey vocal production. *Proceedings of the National Academy of Sciences*, 115(15):3978–3983.

Lidhar, N. K., Insel, N., Dong, J. Y., and Takehara-Nishiuchi, K. (2017). Observational fear learning in degus is correlated with temporal vocalization patterns. *Behavioural Brain Research*, 332:362–371.

Lieberman, P. (2012). Vocal tract anatomy and the neural bases of talking. *Journal of Phonetics*, 40(4):608–622.

Liu, W.-c., Wada, K., Jarvis, E. D., and Nottebohm, F. (2013). Rudimentary substrates for vocal learning in a suboscine. *Nature Communications*, 4:2082.

Long, C. V. (2007). Vocalisations of the defu *OCTODON DEGUS* ,a social caviomorph rodent. *Bioacoustics*, 16(3):223–244.

Long, H., Sabatier, C., Ma, L., Plump, A., Yuan, W., Ornitz, D. M., Tamada, A., Murakami, F., Goodman, C. S., and Tessier-Lavigne, M. (2004). Conserved Roles for Slit and Robo Proteins in Midline Commissural Axon Guidance. *Neuron*, 42(2):213–223.

Mahrt, E. J., Perkel, D. J., Tong, L., Rubel, E. W., and Portfors, C. V. (2013). Engineered Deafness Reveals That Mouse Courtship Vocalizations Do Not Require Auditory Experience. *Journal of Neuroscience*, 33(13):5573–5583.

Mallick, S., Li, H., Lipson, M., Mathieson, I., Gymrek, M., Racimo, F., Zhao, M., Chennagiri, N., Nordenfelt, S., Tandon, A., Skoglund, P., Lazaridis, I., Sankararaman, S., Fu, Q., Rohland, N., Renaud, G., Erlich, Y., Willems, T., Gallo, C., Spence, J. P., Song, Y. S., Poletti, G., Balloux, F., van Driem, G., de Knijff, P., Romero, I. G., Jha, A. R., Behar, D. M., Bravi, C. M., Capelli, C., Hervig, T., Moreno-Estrada, A., Posukh, O. L., Balanovska, E., Balanovsky, O., Karachanak-Yankova, S., Sahakyan, H., Toncheva, D., Yepiskoposyan, L., Tyler-Smith, C., Xue, Y., Abdullah, M. S., Ruiz-Linares, A., Beall, C. M., Rienzo, A. D., Jeong, C., Starikovskaya, E. B., Metspalu, E., Parik, J., Villems, R., Henn, B. M., Hodoglugil, U., Mahley, R., Sajantila, A., Stamatoyannopoulos, G., Wee, J. T. S., Khusainova, R., Khusnutdinova, E., Litvinov, S., Ayodo, G., Comas, D., Hammer, M. F., Kivisild, T., Klitz, W., Winkler, C. A., Labuda, D., Bamshad, M., Jorde, L. B., Tishkoff, S. A., Watkins, W. S., Metspalu, M., Dryomov, S., Sukernik, R., Singh, L., Thangaraj, K., Pääbo, S., Kelso, J., Patterson, N., and Reich, D. (2016). The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. *Nature*, 538(7624):201.

Mampe, B., Friederici, A. D., Christophe, A., and Wermke, K. (2009). Newborns' Cry Melody Is Shaped by Their Native Language. *Current Biology*, 19(23):1994–1997.

Maricic, T., Günther, V., Georgiev, O., Gehre, S., Ćurlin, M., Schreiweis, C., Naumann, R., Burbano, H. A., Meyer, M., Lalueza-Fox, C., de la Rasilla, M., Rosas, A., Gajović, S., Kelso, J., Enard, W., Schaffner, W., and Pääbo, S. (2013). A Recent Evolutionary Change Affects a Regulatory Element in the Human FOXP2 Gene. *Molecular Biology and Evolution*, 30(4):844–852.

Marillat, V., Cases, O., Nguyenf-Ba-Charvet, K. T., Tessier-Lavigne, M., Sotelo, C., and Chédotal, A. (2002). Spatiotemporal expression patterns of *slit* and *robo* genes in the rat brain. *Journal of Comparative Neurology*, 442(2):130–155.

Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. W.H. Freeman, San Francisco.

Martins, P. T. and Boeckx, C. (2016a). Language evolution: Insisting on making it a mystery or turning it into a problem? In Dupuy, L., Grabizna, A., Foudon, N., and Saint-Germier, P., editors, *Papers Dedicated to Anne Reboul*, pages 1–10. Institut des

Sciences Cognitives/CNRS, Lyon.

Martins, P. T. and Boeckx, C. (2016b). What we talk about when we talk about biolinguistics. *Linguistics Vanguard*, 2(1).

Martins, P. T. and Boeckx, C. (2018). Sound Production Learning and the Contiguum Hypothesis. In *Proceedings of the 12th International Conference on the Evolution of Language (Evolang12)*, pages 284–286. Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika.

Martins, P. T. and Boeckx, C. (2019). Language evolution and complexity considerations: The no half-Merge fallacy. *PLOS Biology*, 17(11):e3000389.

Martins, P. T. and Boeckx, C. (2020). Vocal learning: Beyond the continuum. *PLOS Biology*, 18(3):e3000672.

Martins, P. T., Leivada, E., Benítez-Burraco, A., and Boeckx, C. (2016). Biological pluralism in service of biolinguistics. In Fujita, K. and Boeckx, C., editors, *Advances in Biolinguistics: The Human Language Faculty and Its Biological Basis*, pages 153–169. Routledge, London.

Martins, P. T., Marí, M., and Boeckx, C. (2018). SRGAP2 and the gradual evolution of the modern human language faculty. *Journal of Language Evolution*, 3(1):67–78.

Mcbrearty, S. and Brooks, A. S. (2000). The revolution that wasn't: A new interpretation of the origin of modern human behavior. *Journal of Human Evolution*, 39(5):453–563.

McConnell, R. E., van Veen, J. E., Vidaki, M., Kwiatkowski, A. V., Meyer, A. S., and Gertler, F. B. (2016). A requirement for filopodia extension toward Slit during Robo-mediated axon repulsion. *Journal of Cell Biology*, 213(2):261–274.

Meader, C. L. and Muyskens, J. H. (1950). *Handbook of Biolinguistics*. H.C. Weller, Toledo, OH.

Mendez, F. L., Poznik, G. D., Castellano, S., and Bustamante, C. D. (2016). The divergence of Neandertal and modern human Y chromosomes. *The American Journal of Human Genetics*, 98(4):728–734.

Miller, C. T. and Wren Thomas, A. (2012). Individual recognition during bouts of antiphonal calling in common marmosets. *Journal of Comparative Physiology. A, Neuroethology, Sensory, Neural, and Behavioral Physiology*, 198(5):337–346.

Miller, J. R. and Engstrom, M. D. (2007). Vocal Stereotypy and Singing Behavior in Baiomyine Mice. *Journal of Mammalogy*, 88(6):1447–1465.

Minias, P. and Podlaszczuk, P. (2017). Longevity is associated with relative brain size in birds. *Ecology and Evolution*, 7(10):3558–3566.

Mitani, J. and Gros-Louis, J. (1998). Chorusing and Call Convergence in Chimpanzees: Tests of Three Hypotheses. *Behaviour*, 135(8):1041–1064.

Mithen, S. (2005). *The Singing Neanderthals: The Origin of Language, Music, Mind and Body*. Weidenfeld and Nicolson, London.

Moczek, A. P. (2008). On the origins of novelty in development and evolution. *BioEssays*, 30(5):432–447.

Mukherji, N. (2010). *The Primacy of Grammar.* MIT Press, Cambridge, MA.

Muller, G. B. and Wagner, G. P. (1991). Novelty in Evolution: Restructuring the Concept. *Annual Review of Ecology and Systematics*, 22(1):229–256.

Nottebohm, F. (2005). The Neural Basis of Birdsong. *PLoS Biology*, 3(5):e164.

Nottebohm, F., Stokes, T. M., and Leonard, C. M. (1976). Central control of song in the canary, Serinus canarius. *Journal of Comparative Neurology*, 165(4):457–486.

Okada, H., Uezu, A., Mason, F. M., Soderblom, E. J., Moseley, M. A., and Soderling, S. H. (2011). SH3 Domain–Based Phototrapping in Living Cells Reveals Rho Family GAP Signaling Complexes. *Sci. Signal.*, 4(201):rs13–rs13.

Okanoya, K. (2015). Evolution of song complexity in Bengalese finches could mirror the emergence of human language. *Journal of Ornithology*, 156(1):65–72.

Okanoya, K. (2017). Sexual communication and domestication may give rise to the signal complexity necessary for the emergence of language: An indication from songbird studies. *Psychonomic Bulletin & Review*, 24(1):106–110.

Okobi, D. E., Banerjee, A., Matheson, A. M. M., Phelps, S. M., and Long, M. A. (2019). Motor cortical control of vocal interaction in neotropical singing mice. *Science*, 363(6430):983–988.

Pääbo, S. (2014a). The Human Condition—A Molecular Approach. *Cell*, 157(1):216–226.

Pääbo, S. (2014b). *Neanderthal Man: In Search of Lost Genomes.* Basic Books, New York, NY.

Penfield, W. and Boldrey, E. (1937). Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation. *Brain*, 60(4):389–443.

Perfors, A. (2017). On simplicity and emergence. *Psychonomic Bulletin & Review*, 24(1):175–176.

Perlman, M. and Clark, N. (2015). Learned vocal and breathing behavior in an enculturated gorilla. *Animal Cognition*, 18(5):1165–1179.

Petkov, C. I. and Jarvis, E. (2012). Birds, primates, and spoken language origins: Behavioral phenotypes and neurobiological substrates. *Frontiers in Evolutionary Neuroscience*, 4.

Pfenning, A. R., Hara, E., Whitney, O., Rivas, M. V., Wang, R., Roulhac, P. L., Howard, J. T., Wirthlin, M., Lovell, P. V., Ganapathy, G., Mouncastle, J., Moseley, M. A., Thompson, J. W., Soderblom, E. J., Iriki, A., Kato, M., Gilbert, M. T. P., Zhang, G., Bakken, T., Bongaarts, A., Bernard, A., Lein, E., Mello, C. V., Hartemink, A. J., and Jarvis, E. D. (2014). Convergent transcriptional specializations in the brains of humans and song-learning birds. *Science (New York, N.Y.)*, 346(6215):1256846.

Pidoux, L., Le Blanc, P., Levenes, C., and Leblois, A. (2018). A subcortical circuit linking

the cerebellum to the basal ganglia engaged in vocal learning. *eLife*, 7:e32167.

Pigliucci, M. and Müller, G. (2010). *Evolution—The Extended Synthesis*. MIT Press, Cambridge, MA.

Pillsbury, W. B. and Meader, C. L. (1928). *The Psychology of Language*. D. Appleton & Co, New York.

Poeppel, D. and Embick, D. (2005). Defining the relation between linguistics and neuroscience. In Cutler, A., editor, *Twenty-First Century Psycholinguistics*, pages 103–118. Lawrence Erlbaum, Mahwah, NJ/London.

Pomberger, T., Risueno-Segovia, C., Löschner, J., and Hage, S. R. (2018). Precise Motor Control Enables Rapid Flexibility in Vocal Behavior of Marmoset Monkeys. *Current Biology*, 0(0).

Prat, Y., Taub, M., and Yovel, Y. (2015). Vocal learning in a social mammal: Demonstrated by isolation and playback experiments in bats. *Science Advances*, 1(2):e1500019.

Prud'homme, B., Minervino, C., Hocine, M., Cande, J. D., Aouane, A., Dufour, H. D., Kassner, V. A., and Gompel, N. (2011). Body plan innovation in treehoppers through the evolution of an extra wing-like appendage. *Nature*, 473(7345):83.

QJS (1925). Laboratory and research. *Quarterly Journal of Speech*, 11(3):274–285.

Raimy, E. (2012). Phonological Architecture: A Biolinguistic Perspective. *BioScience*, 62(10):925–927.

Ralls, K., Fiorelli, P., and Gish, S. (1985). Vocalizations and vocal mimicry in captive harbor seals, *Phoca vitulina*. *Canadian Journal of Zoology*, 63(5):1050–1056.

Ramchand, G. (2015). *Statement. Roundtable Discussion. \emphGenerative Syntax in the Twenty-First Century: The Road Ahead, Athens, Greece.*

Ravignani, A., Fitch, W. T., Hanke, F. D., Heinrich, T., Hurgitsch, B., Kotz, S. A., Scharff, C., Stoeger, A. S., and de Boer, B. (2016). What Pinnipeds Have to Say about Human Speech, Music, and the Evolution of Rhythm. *Frontiers in Neuroscience*, 10.

Robbins, M. M., Ando, C., Fawcett, K. A., Grueter, C. C., Hedwig, D., Iwata, Y., Lodwick, J. L., Masi, S., Salmi, R., Stoinski, T. S., Todd, A., Vercellio, V., and Yamagiwa, J. (2016). Behavioral Variation in Gorillas: Evidence of Potential Cultural Traits. *PLOS ONE*, 11(9):e0160483.

Rödel, R. M., Olthoff, A., Tergau, F., Simonyan, K., Kraemer, D., Markus, H., and Kruse, E. (2004). Human cortical motor representation of the larynx as assessed by transcranial magnetic stimulation (TMS). *The Laryngoscope*, 114(5):918–922.

Rodenas-Cuadrado, P., Chen, X. S., Wiegrebe, L., Firzlaff, U., and Vernes, S. C. (2015). A novel approach identifies the first transcriptome networks in bats: A new genetic model for vocal communication. *BMC Genomics*, 16(1):836.

Samuels, B. D. (2015). Can a bird brain do phonology? *Frontiers in Psychology*, 6.

Scerri, E. M. L., Thomas, M. G., Manica, A., Gunz, P., Stock, J. T., Stringer, C., Grove,

M., Groucutt, H. S., Timmermann, A., Rightmire, G. P., d'Errico, F., Tryon, C. A., Drake, N. A., Brooks, A. S., Dennell, R. W., Durbin, R., Henn, B. M., Lee-Thorp, J., deMenocal, P., Petraglia, M. D., Thompson, J. C., Scally, A., and Chikhi, L. (2018). Did Our Species Evolve in Subdivided Populations across Africa, and Why Does It Matter? *Trends in Ecology & Evolution*, 33(8):582–594.

Schreiweis, C., Bornschein, U., Burguière, E., Kerimoglu, C., Schreiter, S., Dannemann, M., Goyal, S., Rea, E., French, C. A., Puliyadi, R., Groszer, M., Fisher, S. E., Mundry, R., Winter, C., Hevers, W., Pääbo, S., Enard, W., and Graybiel, A. M. (2014). Humanized Foxp2 accelerates learning by enhancing transitions from declarative to procedural performance. *Proceedings of the National Academy of Sciences*, 111(39):14253–14258.

Shaughnessy, D. W., Hyson, R. L., Bertram, R., Wu, W., and Johnson, F. (2019). Female zebra finches do not sing yet share neural pathways necessary for singing in males. *Journal of Comparative Neurology*, 527(4):843–855.

Shea, J. J. (2011). Homo sapiens Is as Homo sapiens Was. *Current Anthropology*, 52(1):1–35.

Shen, H. H. (2017). News Feature: Singing in the brain. *Proceedings of the National Academy of Sciences*, 114(36):9490–9493.

Shieber, S. M. (1985). Evidence Against the Context-Freeness of Natural Language. In Cooper, R., Engdahl, E., Grandy, R., Savitch, W. J., Bach, E., Marsh, W., and Safran-Naveh, G., editors, *The Formal Complexity of Natural Language*, volume 33, pages 320–334. Springer Netherlands, Dordrecht.

Shimizu, T., Patton, T. B., and Husband, S. A. (2010). Avian visual behavior and the organization of the telencephalon. *Brain, behavior and evolution*, 75(3):204–217.

Silvente i Font, S., Martins, P. T., and Boeckx, C. (2020). Do life histories shape vocal production learning? A bird-based approach. In Ravignani, A., Barbieri, C., Martins, M., Flaherty, M., Jadoul, Y., Lattenkamp, E., Little, H., Mudd, K., and Verhoef, T., editors, *The Evolution of Language: Proceedings of the 13th International Conference (EvoLang13)*.

Simonyan, K. (2014). The laryngeal motor cortex: Its organization and connectivity. *Current opinion in neurobiology*, 28:15–21.

Snowdon, C. T. and Cleveland, J. (1984). "Conversations" among pygmy marmosets. *American Journal of Primatology*, 7(1):15–20.

Snowdon, C. T. and Elowson, A. M. (1999). Pygmy Marmosets Modify Call Structure When Paired. *Ethology*, 105(10):893–908.

Sporny, M., Guez-Haddad, J., Kreusch, A., Shakartzi, S., Neznansky, A., Cross, A., Isupov, M. N., Qualmann, B., Kessels, M. M., and Opatowsky, Y. (2017). Structural History of Human SRGAP2 Proteins. *Molecular Biology and Evolution*, 34(6):1463–1478.

Stoeger, A. S. and Manger, P. (2014). Vocal learning in elephants: Neural bases and adaptive context. *Current Opinion in Neurobiology*, 28:101–107.

Striedter, G. F. and Charvet, C. J. (2008). Developmental origins of species differences in telencephalon and tectum size: Morphometric comparisons between a parakeet (*Melopsittacus undulatus*) and a quail (*Colinus virgianus*). *The Journal of Comparative Neurology*, 507(5):1663–1675.

Stringer, C. (2016). The origin and evolution of Homo sapiens. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1698):20150237.

Sugar, O., Chusid, J. G., and French, J. D. (1948). A Second Motor Cortex in the Monkey (Macaca Mulatta). *Journal of Neuropathology & Experimental Neurology*, 7(2):182–189.

Sugiura, H. (1998). Matching of acoustic features during the vocal exchange of coo calls by Japanese macaques. *Animal Behaviour*, 55(3):673–687.

Takahashi, D. Y. (2019). Vocal Learning: Shaping by Social Reinforcement. *Current Biology*, 29(4):R125–R127.

Takahashi, D. Y., Fenley, A. R., and Ghazanfar, A. A. (2016). Early development of turn-taking with parents shapes vocal acoustics in infant marmoset monkeys. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1693):20150370.

Takahashi, D. Y., Fenley, A. R., Teramoto, Y., Narayanan, D. Z., Borjon, J. I., Holmes, P., and Ghazanfar, A. A. (2015). The developmental dynamics of marmoset monkey vocal production. *Science*, 349(6249):734–738.

Takahashi, D. Y., Liao, D. A., and Ghazanfar, A. A. (2017). Vocal Learning via Social Reinforcement by Infant Marmoset Monkeys. *Current Biology*, 27(12):1844–1852.e6.

Takahashi, D. Y., Narayanan, D. Z., and Ghazanfar, A. A. (2013). Coupled Oscillator Dynamics of Vocal Turn-Taking in Monkeys. *Current Biology*, 23(21):2162–2168.

Takahasi, M. and Okanoya, K. (2010). Song Learning in Wild and Domesticated Strains of White-Rumped Munia, *Lonchura striata* , Compared by Cross-Fostering Procedures: Domestication Increases Song Variability by Decreasing Strain-Specific Bias. *Ethology*, 116(5):396–405.

Tanaka, M., Sun, F., Li, Y., and Mooney, R. (2018). A mesocortical dopamine circuit enables the cultural transmission of vocal behaviour. *Nature*, 563(7729):117–120.

Tchernichovski, O. and Marcus, G. (2014). Vocal learning beyond imitation: Mechanisms of adaptive vocal development in songbirds and human infants. *Current Opinion in Neurobiology*, 28:42–47.

Theofanopoulou, C. and Boeckx, C. (2015). Cognitive phylogenies, the Darwinian logic of descent, and the inadequacy of cladistic thinking. *Frontiers in Cell and Developmental Biology*, 3.

Thomas, J. (2014). *Self-Domestication and Language Evolution*. PhD thesis, University of Edinburgh.

Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift für Tierpsychologie*, 20(4):410–433.

Tiwary, B. K. (2016). Evolution of the SRGAP2 Gene Is Linked to Intelligence in Mammals. *Biomedicine Hub*, 1(1):443947–443947.

Tokimoto, N. and Okanoya, K. (2004). Spontaneous construction of "Chinese boxes" by Degus (*Octodon degu*): A rudiment of recursive intelligence?: Cup nesting by Degus. *Japanese Psychological Research*, 46(3):255–261.

Tyack, P. L. and Sayigh, L. S. (1997). Vocal learning in cetaceans. In Snowdon, C. T. and Hausberger, M., editors, *Social Influences on Vocal Development*, pages 208–233. Cambridge University Press, Cambridge.

Uriagereka, J. (2008). *Syntactic Anchors: On Semantic Structuring*. Cambridge University Press, Cambridge.

van Schaik, C. P., Ancrenaz, M., Borgen, G., Galdikas, B., Knott, C. D., Singleton, I., Suzuki, A., Utami, S. S., and Merrill, M. (2003). Orangutan cultures and the evolution of material culture. *Science (New York, N.Y.)*, 299(5603):102–105.

Vanden Hole, C., Van Daele, P. A., Desmet, N., Devos, P., and Adriaens, D. (2014). Does sociality imply a complex vocal communication system? A case study for *Fukomys micklemi* (Bathyergidae, Rodentia). *Bioacoustics*, 23(2):143–160.

Vernes, S. C. and Fisher, S. E. (2013). Genetic pathways implicated in speech and language. In *Animal Models of Speech and Language Disorders*, pages 13–40. Springer.

Vernes, S. C., Spiteri, E., Nicod, J., Groszer, M., Taylor, J. M., Davies, K. E., Geschwind, D. H., and Fisher, S. E. (2007). High-throughput analysis of promoter occupancy reveals direct neural targets of FOXP2, a gene mutated in speech and language disorders. *American Journal of Human Genetics*, 81(6):1232–1250.

Vernes, S. C. and Wilkinson, G. S. (2020). Behaviour, biology and evolution of vocal learning in bats. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375(1789):20190061.

Vignaud, P., Duringer, P., Mackaye, H. T., Likius, A., Blondel, C., Boisserie, J.-R., de Bonis, L., Eisenmann, V., Etienne, M.-E., Geraads, D., Guy, F., Lehmann, T., Lihoreau, F., Lopez-Martinez, N., Mourer-Chauviré, C., Otero, O., Rage, J.-C., Schuster, M., Viriot, L., Zazzo, A., and Brunet, M. (2002). Geology and palaeontology of the Upper Miocene Toros-Menalla hominid locality, Chad. *Nature*, 418(6894):152.

Wang, R., Chen, C.-C., Hara, E., Rivas, M. V., Roulhac, P. L., Howard, J. T., Chakraborty, M., Audet, J.-N., and Jarvis, E. D. (2015). Convergent differential regulation of SLIT-ROBO axon guidance genes in the brains of vocal learners. *The Journal of Comparative Neurology*, 523(6):892–906.

Wang, Y., Brzozowska-Prechtl, A., and Karten, H. J. (2010). Laminar and columnar auditory cortex in avian brain. *Proceedings of the National Academy of Sciences*, 107(28):12676–12681.

Watson, S. K., Townsend, S. W., Schel, A. M., Wilke, C., Wallace, E. K., Cheng, L., West, V., and Slocombe, K. E. (2015). Vocal Learning in the Functionally Referential Food Grunts of Chimpanzees. *Current Biology*, 25(4):495–499.

Watts, D. P. (2016). Production of grooming-associated sounds by chimpanzees (Pan troglodytes) at Ngogo: Variation, social learning, and possible functions. *Primates*, 57(1):61–72.

West-Eberhard, M. J. (2003). *Developmental Plasticity and Evolution*. Oxford University Press, Oxford.

Whitford, K. L., Marillat, V., Stein, E., Goodman, C. S., Tessier-Lavigne, M., Chédotal, A., and Ghosh, A. (2002). Regulation of Cortical Dendrite Development by Slit-Robo Interactions. *Neuron*, 33(1):47–61.

Wich, S. A., Krützen, M., Lameira, A. R., Nater, A., Arora, N., Bastian, M. L., Meulman, E., Morrogh-Bernard, H. C., Atmoko, S. S. U., Pamungkas, J., Perwitasari-Farajallah, D., Hardus, M. E., van Noordwijk, M., and van Schaik, C. P. (2012). Call Cultures in Orang-Utans? *PLoS ONE*, 7(5):e36180.

Wich, S. A., Swartz, K. B., Hardus, M. E., Lameira, A. R., Stromberg, E., and Shumaker, R. W. (2009). A case of spontaneous acquisition of a human sound by an orangutan. *Primates*, 50(1):56–64.

Wild, J. M., Li, D., and Eagleton, C. (1997). Projections of the dorsomedial nucleus of the intercollicular complex (DM) in relation to respiratory-vocal nuclei in the brainstem of pigeon (Columba livia) and zebra finch (Taeniopygia guttata). *Journal of Comparative Neurology*, 377(3):392–413.

Wirthlin, M., Chang, E. F., Knörnschild, M., Krubitzer, L. A., Mello, C. V., Miller, C. T., Pfenning, A. R., Vernes, S. C., Tchernichovski, O., and Yartsev, M. M. (2019). A Modular Approach to Vocal Learning: Disentangling the Diversity of a Complex Behavioral Trait. *Neuron*, 104(1):87–99.

Wong, K., Ren, X.-R., Huang, Y.-Z., Xie, Y., Liu, G., Saito, H., Tang, H., Wen, L., Brady-Kalnay, S. M., Mei, L., Wu, J. Y., Xiong, W.-C., and Rao, Y. (2001). Signal Transduction in Neuronal Migration:Roles of GTPase Activating Proteins and the Small GTPase Cdc42 in the Slit-Robo Pathway. *Cell*, 107(2):209–221.

Yao, Q., Jin, W.-L., Wang, Y., and Ju, G. (2008). Regulated shuttling of Slit-Robo-GTPase activating proteins between nucleus and cytoplasm during brain development. *Cellular and Molecular Neurobiology*, 28(2):205–221.

Zhang, Y. S. and Ghazanfar, A. A. (2018). Vocal development through morphological computation. *PLOS Biology*, 16(2):e2003933.

Zhao, L., Rad, B. B., and Wang, X. (2019). Long-lasting vocal plasticity in adult marmoset monkeys. *Proceedings. Biological Sciences*, 286(1905):20190817.

Zürcher, Y. and Burkart, J. M. (2017). Evidence for Dialects in Three Captive Populations of Common Marmosets (Callithrix jacchus). *International Journal of Prima-*

tology, 38(4):780–793.

List of publications

The following is a list papers published over the course of my PhD.

Papers

[1] **Martins, P. T.**, & Boeckx, C. (2020). Vocal learning: Beyond the continuum. PLoS Biology, 18(3), e3000672. doi:10.1371/journal.pbio.3000672

[2] **Martins, P. T.** (2019). Comment on "The Siege of Paris" (Berwick and Chomsky, 2019). Inference: International Review of Science, 4(4).

[3] Zanella, M., Vitriolo, A., Andirko, A., **Martins, P. T.**, Sturm, S., O'Rourke, T., Laugsch, M., Malerba, N., Skaros, A., Trattaro, S., Germain, P.-L., Mihailovic, M., Merla, G., Rada-Iglesias, A., Boeckx, C., & Testa, G. (2019). Dosage analysis of the 7q11.23 Williams region identifies BAZ1B as a major human gene patterning the modern human face and underlying self-domestication. Science Advances, 5(12). doi:10.1126/sciadv.aaw7908

[4] **Martins, P. T.**, & Boeckx, C. (2019). Language evolution and complexity considerations: The no half-Merge fallacy. PLoS Biology, 17(11), e3000389. doi:10.1371/journal.pbio.3000389

[5] **Martins, P. T.**, Marí, M., & Boeckx, C. (2018). SRGAP2 and the gradual evolution of the modern human language faculty. Journal of Language Evolution, 3(1), 67–78. doi:10.1093/jole/lzx020

[6] Samuels, B. D., **Martins, P. T.**, & Boeckx, C. (2017). Linguistic knowledge by descent: an evolutionary approach to stress typology. Journal of the Phonetic Society of Japan [Special Issue: Excavating Phonetic/Phonological Fossils in Language: Current Trends in Evolutionary Linguistics], 21(1), 71–78. doi:10.24467/onseikenkyu.21.1_71

[7] Theofanopoulou, C., Gastaldon, S., O'Rourke, T., Samuels, B. D., **Martins, P. T.**, Delogu, F., Alamri, S., & Boeckx, C. (2017). Self-domestication in Homo

sapiens: Insights from comparative genomics. PLoS ONE, 12(10), e0185306. doi:10.1371/journal.pone.0185306

[8] Martins, P. T., & **Boeckx, C.** (2016). What we talk about when we talk about biolinguistics. Linguistics Vanguard, 2(1). doi:10.1515/lingvan-2016-0007

Chapters

[9] **Martins, P. T.**, & Boeckx, C. (2018). Language evolution: Insisting on making it a mystery or turning it into a problem? In P. Saint-Germier (Ed.), Language, Evolution and Mind: Essays in Honour of Anne Reboul (pp. 161–169). London: College Publications.

[10] **Martins, P. T.** (2017). There is no place for markedness in biologically-informed phonology. In B. D. Samuels (Ed.), Beyond markedness in formal phonology (pp. 219–232). Amsterdam: John Benjamins. doi:10.1075/la.241.09mar

[11] Boeckx, C., **Martins, P. T.**, & Leivada, E. (2017). Biolinguistics. In I. Roberts & A. Ledgeway (Eds.), Cambridge Handbook of Historical Syntax (pp. 629–641). Cambridge: Cambridge University Press. doi:10.1017/9781107279070.029

[12] Boeckx, C., & **Martins, P. T.** (2016). Biolinguistics. In M. Aronoff (Ed.), Oxford Research Encyclopedia of Linguistics. doi:10.1093/acrefore/9780199384655.013.20

[13] **Martins, P. T.**, & Boeckx, C. (2016). Language evolution: Insisting on making it a mystery or turning it into a problem? In L. Dupuy, A. Grabizna, N. Foudon, & P. Saint-Germier (Eds.), Papers dedicated to Anne Reboul (pp. 1–10). Lyon: Institut des Sciences Cognitives/CNRS.

[14] **Martins, P. T.**, Leivada, E., Benítez-Burraco, A., & Boeckx, C. (2016). Biological pluralism in service of biolinguistics. In K. Fujita & C. Boeckx (Eds.), Advances in Biolinguistics: The Human Language Faculty and its Biological Basis (pp. 153–169). London: Routledge.

Proceedings

[15] Silvente i Font, S., **Martins, P. T.**, & Boeckx, C. (2020). Do Life Histories Shape Vocal Production Learning? A Bird-Based Approach. In A. Ravignani, C. Barbieri, M. Martins, M. Flaherty, Y. Jadoul, E. Lattenkamp, H. Little, K. Mudd, & T. Verhoef (Eds.), The Evolution of Language: Proceedings of the 13th International Conference (EvoLang13). doi:10.17617/2.3190925

[16] **Martins, P. T.**, & Boeckx, C. (2020). Specifying the "Vocal" in Vocal Learning. In A. Ravignani, C. Barbieri, M. Martins, M. Flaherty, Y. Jadoul, E. Lattenkamp, H. Little, K. Mudd, & T. Verhoef (Eds.), The Evolution of Language: Proceedings of the 13th International Conference (EvoLang13). doi:10.17617/2.3190925

[17] **Martins, P. T.**, & Boeckx, C. (2018). Sound Production Learning and the Contiguum Hypothesis. In Proceedings of Evolang XII. doi:10.12775/3991-1.069