

Birdsong, Speech, and Language: Exploring the Evolution of Mind and Brain

edited by Johan J. Bolhuis and Martin Everaert
foreword by Robert C. Berwick and
Noam Chomsky



The MIT Press

From The MIT Press



MITCogNet

© 2013 Massachusetts Institute of Technology

All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher.

MIT Press books may be purchased at special quantity discounts for business or sales promotional use. For information, please email special_sales@mitpress.mit.edu or write to Special Sales Department, The MIT Press, 55 Hayward Street, Cambridge, MA 02142.

This book was set in Syntax and Times Roman by Toppan Best-set Premedia Limited, Hong Kong. Printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Birdsong, speech, and language: exploring the evolution of mind and brain / edited by Johan J. Bolhuis and Martin Everaert; foreword by Robert C. Berwick and Noam Chomsky.
p. cm.

Includes bibliographical references and index.

ISBN 978-0-262-01860-9 (alk. paper)

1. Language acquisition. 2. Birds—Vocalization. 3. Birdsongs. 4. Speech acts (Linguistics) 5. Human evolution. 6. Cognitive neuroscience. 7. Neurolinguistics. I. Bolhuis, Johan J. II. Everaert, Martin.

P118.B56 2013

401'.93—dc23

2012028611

10 9 8 7 6 5 4 3 2 1

W. Tecumseh Fitch

Darwin's *Origin of Species* (1859) made little mention of human evolution. This initial avoidance of human evolution was no oversight, but rather a carefully calculated move: Darwin was well aware of the widespread resistance his theory would meet from scientists, clergymen, and the lay public, and mention of human evolution might have generated insuperable opposition. But Darwin's many opponents quickly seized on the human mind, and language in particular, as a potent weapon in the battle against his new way of thinking. Alfred Wallace, whose independent discovery of the principle of natural selection spurred Darwin to finally publish his long-developing "outline" of the theory in 1859, did not help by arguing that natural selection was unable to explain the origins of the human mind. Although Wallace had reservations about all evolutionary approaches to the mind, human language provided the most powerful argument, due to the respectable position of linguistics and philology in Victorian science.

Darwin's most formidable foe on the linguistic front was Friederich Max Müller, professor of linguistics at Oxford University, a very well-known and well-respected scholar (Stam, 1976). In his "Lectures on the Science of Language," delivered at the Royal Institution of Great Britain in 1861 and rapidly published thereafter (Müller, 1861), Müller launched a full frontal attack on Darwin and Darwinism, using his credentials in the "science of language" as a powerful bludgeon. Müller's position was uncomplicated: "Language is the Rubicon which divides man from beast, and no animal will ever cross it. . . . The science of language will yet enable us to withstand the extreme theories of the Darwinians, and to draw a hard and fast line between man and brute" (Quoted in Noiré, 1917, p. 73–74). For Müller, language was the key feature distinguishing humans from all animals. Müller's arguments were seen by many as convincing: his student Noiré dubbed him the "Darwin of the mind" and considered Müller "the only equal, not to say superior, antagonist, who has entered the arena against Darwin" (Noiré, 1917, p. 73). Müller's argument about the unbridgeable, qualitative difference between human language and all forms of animal communication, combined

with Wallace's opinions, provided arguments that Darwin by necessity took very seriously.

Thus, when Darwin finally broached the subject of human evolution in his second great book, *The Descent of Man and Selection in Relation to Sex* (1871), the need to provide a credible explanation for language evolution was a central concern. He rose to the challenge: his "musical protolanguage" model represents a powerful marriage of comparative data, evolutionary insight, and a biological perspective on language. Darwin's view of language was ahead of its time, and his model and arguments remain surprisingly relevant to contemporary debates. He clearly adopted a "multicomponent" view of language, one that recognized the necessity of several distinct mechanisms to produce the complex product that we now call language, rather than privileging any one factor as the single key to language in a monolithic sense. Among these several components, he presciently recognized the necessity for complex vocal learning, and recognized that this biological capacity, while unusual among mammals, is shared with many birds. The importance of vocal learning has often been forgotten, but also frequently reaffirmed by later scholars (Egnor & Hauser, 2004; Fitch, 2000; Janik & Slater, 1997; Marler, 1976; Nottebohm, 1976).

Darwin also adopted an empirical, data-driven approach to the problem at hand, exploiting what Botha (2009) has termed "windows" into language evolution. In particular, Darwin exploited a wide comparative database, drawing on not just his knowledge of nonhuman primate behavior, but also insights from many other vertebrates. Finally, and most characteristically, he resisted any special pleading about human evolution. He intended his model of human evolution to fit within, and remain consistent with, a broader theory of evolution that applies to beetles, flowers, and birds. Unlike Wallace (1905), who remained a human exceptionalist to his death, Darwin aimed to uncover general principles, like sexual selection and shifts of function, to provide explanations of unusual or unique human traits. While gradualistic, his model does not assume any simple continuity of function between nonhuman primate calls and language, and he clearly recognized the uniqueness of language in our species. In many ways, then, Darwin's model of language evolution finds a natural place in the landscape of the contemporary debate on language evolution, and it is surprising that his model has received relatively little detailed consideration in the modern literature (for exceptions see Donald, 1991; Fitch, 2006).

In this chapter, I aim to redress this neglect by considering Darwin's model of language evolution in detail. After discussing Darwin's main points and arguments, I briefly review additional data supporting Darwin's model that has appeared since his death. I also discuss the issue of meaning, about which Darwin had too little to say, but that can be resolved by the addition of a hypothesis due to Jespersen (1922). My conclusion is that, suitably modified in light of contemporary understanding, Darwin's model of language evolution, based on a "protolanguage" more musical

than linguistic, provides one of the most convincing frameworks available for understanding language evolution. The present book provides an appropriate place to discuss Darwin's model, given the heavy reliance of both on comparative data concerning birdsong. The timing of my writing, on the 150th anniversary of the *Origin* and the 200th of Darwin's birth, is also appropriate for a revival of interest in Darwin's compelling and well-supported hypothesis.

Language as an "Instinct to Learn"

Chapter 2 of the *Descent of Man* (Darwin, 1871), titled "Comparison of the Mental Powers of Man and the Lower Animals," is one of the most remarkable portions of the entire Darwinian corpus, noteworthy for its conciseness and its breadth of argument, in considering the evolution of the human mind. The first half of the chapter lays the groundwork for modern research in comparative cognition, arguing that animals have emotions, attention, and memory as well as many other mental traits in common with humans. However, Darwin's opponents, notably Müller, had already ceded the point that animals have memory, experience emotions, and so on. Language was the key issue, and one can imagine considerable anticipation of both pro- and anti-Darwinian readers as they turned to the section simply titled "Language."

In ten densely argued pages, Darwin considers some theoretical preliminaries, then lays out his theory of language evolution. The first stage involved a general increase in intelligence and complex mental abilities, and the second involved a sexually selected attainment of the specific capacity for complex vocal control: singing. The third stage was the addition of meaning to the "songs" of the second stage, which was both driven by, and in turn fueled, further increases in intelligence.

Theoretically, Darwin makes a number of important observations. First, he recognizes the crucial distinction between the language *faculty* (the biological capacity that enables humans to acquire language) and particular languages (like Latin or English). The former capacity, which Darwin refers to as "an instinctive tendency to acquire an art" (p. 56), is shared by all members of the human species. Darwin neatly bypasses the unproductive nature-nurture debate that has consumed so much scholarly energy by observing that language "is not a true instinct, as every language has to be learnt. It differs, however, from all ordinary arts, for man has an instinctive tendency to speak, as we see in the babble of our young children" (p. 55). As ethologist Peter Marler has put it, language is not an instinct, but an "instinct to learn" whose expression entails that both biological and environmental preconditions be fulfilled. It is this "instinct to learn" for which a biological, evolutionary explanation must be sought: a thoroughly modern perspective.

Second, although Darwin was well aware of the peculiarities of the human vocal tract, he argues that the human capacity for language must be sought in the brain, rather than the peripheral vocal tract. He acknowledges that “articulate speech” (by which he means vocalization augmented by controlled movement of the lips and tongue, p. 59) is “peculiar to man,” but he denies that this mere power of articulation suffices to distinguish human language, “for as every one knows, parrots can talk.” Darwin states that it is not speech, but humans’ “large power of connecting definite sounds with definite ideas” that is definitive of language, and that this capacity “obviously depends on the development of the mental faculties” (p. 54). By locating the language capacity in the human brain, Darwin’s viewpoint is again thoroughly modern.

Finally, Darwin recognized the relevance to language evolution of birdsong, which he considered the “nearest analogy to language.” Like humans, birds have fully instinctive calls, and an instinct to sing. But the songs themselves are learned. He recognized the parallel between infant babbling and songbird “subsong,” and recognized the key fact that *cultural* transmission ensures the formation of regional dialects in both birdsong and speech. Finally, he recognized that physiology is not enough for learned song: crows have a syrinx as complex as a nightingale’s but use it only in unmusical croaking. All of these parallels have been amply confirmed, and further explored, by modern researchers (Doupe & Kuhl, 1999; Marler, 1970; Nottebohm, 1972, 1975).

Darwin’s “Musical Protolanguage” Hypothesis

Darwin’s model of the phylogenesis of the language faculty, like most models today, posits that different aspects of language were acquired sequentially, in a particular order, and under the influence of distinguishable selection pressures. The hypothetical systems characterized by each addition can be termed, following Bickerton (1990) and Hewes (1973), “protolanguages.” Darwin’s first hypothetical stage in the progression from an apelike ancestor to modern humans was a greater development of protohuman cognition: “The mental powers in some early progenitor of man must have been more highly developed than in any existing ape, before even the most imperfect form of speech could have come into use” (p. 57). He elsewhere suggests that both social and technological factors may have driven this increase in cognitive power.

Next, Darwin outlines the crucial second step: what I have dubbed “musical protolanguage” (Fitch, 2006). Having noted multiple similarities with birdsong, he argues that the evolution of a key aspect of spoken language, vocal imitation, was driven by sexual selection, and used largely “in producing true musical cadences, that is in singing” (p. 56). He suggests that this musical protolanguage would have

been used in both courtship and territoriality (as a “challenge to rivals”), as well as in the expression of emotions like love, jealousy, and triumph. Darwin concludes “from a widely-spread analogy” (amply documented with comparative data later in the book) that sexual selection played a crucial role in driving this stage of language evolution, in particular suggesting that the capacity to imitate vocally evolved analogously in humans and songbirds.

The crucial remaining question is how emotionally expressive musical protolanguage made the transition to true meaningful language—how, in Humboldt’s words, humans became “a singing creature, only associating thoughts with the tones” (Von Humboldt, 1836, p. 76). This leap, from nonpropositional song to propositionally meaningful speech, remains the greatest explanatory challenge for all musical protolanguage theories (see Mithen, 2005). Darwin (1871, p. 56), citing the previous writings of Müller and Farrar (1870), suggests that articulate language “owes its origins to the imitation and modification, aided by signs and gestures, of various natural sounds, the voices of other animals, and man’s own instinctive cries.” Darwin thus embraces all three of the major leading theories of word origins of his contemporaries (see Fitch, 2010). Once protohumans had the capacity to imitate vocally, and to combine such signals with meanings, virtually any source of word forms and meanings would suffice, including onomatopoeia (an imitated roar for *lion*, or “whoosh” for *wind*) and controlled imitation of human emotional vocalizations (mock laughter for *play* or *happiness*). The attachment of specific and flexible meanings to vocalizations required only that “some unusually wise ape-like animal should have thought of imitating the growl of a beast of prey. . . . And this would have been a first step in the formation of a language” (p. 57).

Darwin does not suggest that the evolutionary process would stop with the initial acquisition of meaning. For “as the voice was used more and more, the vocal organs would have been strengthened and perfected” (p. 57). Additionally, language would have “reacted on the mind by enabling and encouraging it to carry on long trains of thought,” which “can no more be carried on without the aid of words, whether spoken or silent, than a long calculation without the use of figures or algebra” (p. 57). Thus began the interactive evolutionary spiral that led to modern human language, and human intelligence, today.

Signaling Modality: Vocalization or Gesture?

Darwin also explicitly acknowledges the role of gesture in conveying meaning, echoing Condillac’s earlier arguments (Condillac, 1747/1971) and presaging contemporary discussions (Arbib, 2005; Corballis, 2003; Hewes, 1973; Stokoe, 1974; Tomasello & Call, 2007). Darwin is aware of the power of signed language: he reminds us that using his fingers “a person with practice can report to a deaf man every word

of a speech rapidly delivered at a public meeting” (p. 58). He also acknowledges the value of gesture in conveying meaning, and allows that vocal communication would have been “aided by signs and gestures” (p. 56). Nevertheless, he argues against gestural theorists, because the preexistence in all mammals of “vocal organs, constructed on the same general plan as ours” would lead any further development of communication to target the vocal organs rather than the fingers.

Darwin clearly believes that the power of speech is neural, not peripheral, citing the early aphasia literature as a demonstration of “the intimate connection between the brain, as it is now developed in us, and the faculty of speech” (p. 58). Comparing the vocal organs and brain, he concludes that “the development of the brain has no doubt been far more important” (p. 57). And although he uses a continuity argument to support the early and sustained role of speech, he firmly acknowledges the abrupt modern *discontinuity* in the linguistic system that has evolved. Thus, like many other insightful commentators (e.g., Donald, 1991; Hockett & Ascher, 1964), Darwin recognizes that posing phylogenetic continuity and modern discontinuity as in any way opposed is to create a false dichotomy. The treelike nature of phylogeny guarantees that both are core parts of the evolutionary process.

Darwin Redux: Modern Comparative Data

Summarizing, Darwin suggests that the first step on the road to human language was a general increase in intelligence in the hominid lineage. In a typically pluralistic fashion, he recognizes both “social intelligence” (“Machiavellian intelligence” in the modern trope (Byrne & Whiten, 1988)) and technological/ecological intelligence (e.g., for tool use) as playing important selective roles. Given our modern understanding of hominid evolution, this first stage might be provisionally linked to the genus *Australopithecus* or perhaps early *Homo* (e.g., *Homo habilis*).

The second stage is the least intuitive: that before vocalizations were used meaningfully they were used, so to speak, aesthetically, to fulfill many of the same functions for which modern humans use music today (courtship, bonding, territorial advertisement and defense, competitive displays, etc.). This idea that complex vocalizations (and thus some aspects of phonology and syntax) might have preceded the ability of speech to convey propositions and distinct meanings is the most challenging aspect of Darwin’s model. But Darwin uses the comparative database, and particularly detailed analogy between learned birdsong and human song and speech, to show that this step is not just plausible but well documented: it has occurred in many other species. Indeed, modern data shows that vocal learning, without propositional meaning, has evolved independently in at least three other clades of mammals (cetaceans, pinnipeds, and bats) and three clades of birds (parrots, hummingbirds, and oscine songbirds) (Janik & Slater, 1997; Jarvis, 2004). Such conver-

gent evolution, or repeated independent evolutionary development of a comparable ability, provides our strongest empirical basis for estimating the likelihood of a particular type of evolutionary event (Harvey & Pagel, 1991). Many of the chapters in this book affirm, and extend, the observations of parallels between language learning and birdsong that Darwin offered in 1871. Thus, whether intuitive or not, Darwin's focus on, and hypothesis for, the evolution of vocal learning is consistent with a wealth of evolutionary and comparative data.

Difficulties with Darwin's Model: Evolving Phrasal Semantics

How did man become, as Humboldt somewhere defined him, "a singing creature, only associating thoughts with the tones"?

—Otto Jespersen (1922, p. 437)

Despite its many virtues, some important problems remain with Darwin's model that have impeded its acceptance today. The first and most important is his explanation of the addition of meaning. Darwin's explanation, typical for his day, was concerned only with *word meanings* (what today would be termed "lexical semantics"). But from the viewpoint of modern linguistics, his model seems wholly inadequate to deal with large swaths of semantics, particularly those aspects tied in with the interpretation of whole phrases and sentences ("phrasal semantics"). Modern formal semantics has developed rigorous models of this aspect of linguistic meaning (Dowty, Wall, & Peters, 1981; Guttenplan, 1986; Montague, 1974; Portner, 2005), and it is far more complex and difficult to explain than lexical semantics. Although one can hardly blame Darwin for not foreseeing these relatively recent developments in linguistics, they nonetheless raise substantial difficulties for his model. For much of the syntactic "glue" that binds sentences together into large, meaningful wholes (function words, inflection, bound morphemes, word order, and a host of other elements) cannot be understood as resulting from onomatopoeia or imitation of emotional expressions. Nor can they be readily understood as "inventions" of some uniquely intelligent individual: all evidence suggests that these indispensable linguistic tools develop reliably in individuals of normal intelligence (Bickerton, 1981; Kegl, 2002; Mufwene, 2001; Mühlhäusler, 1997; Senghas, Kita, & Özyürek, 2005). This key aspect of language thus seems to have a biological basis. Darwin does recognize the phenomenon today called "grammaticalization": he states that "conjugations, declensions &c., originally existed as distinct words, since joined together" (p. 61). But he offers no model for the origin of these distinct words, and it is hard to see how onomatopoeia or similar processes could have generated this original syntactic and semantic "glue." Thus, complex phrasal semantics remains unexplained by Darwin's model.

However, this oversight was remedied long ago by the linguist Otto Jespersen (1922). Jespersen's basic insight involves recognizing the link, in humans, between musical and linguistic phrases, and working conceptually backward from there. Jespersen suggested a form of protolanguage in which, initially, whole propositional meanings attached to entire sung phrases, but where there was no consistent link between the individual *conceptual* components of the meaning, and component parts of the musical phrases (syllables and notes). Thus, there were no "words" as we now understand them. From this "holistic" starting point, Jespersen argued that a cognitive process of analysis began, which slowly isolated individual chunks of the musical phrase (syllables, or multisyllabic "phraselets" – what today we call "words") and associated them with individual components of the meaning (e.g., nouns, verbs, and adjectives, whose precursors were already present in the conceptual systems of our prelinguistic ancestors).

Jespersen's hypothesis of a "holistic protolanguage" has recently been rediscovered and championed by linguist Alison Wray (1998, 2000) and neuroscientist Michael Arbib (2005). Both cite considerable additional evidence supporting this "analytic" model, including data from modern adult language, child language acquisition, and cognitive neuroscience. Supporters of the more intuitive "synthetic" model of protolanguage, in which words evolved first followed by syntactic operations for combining them (e.g., Bickerton, 1990), have subjected holistic models to extensive criticisms (Bickerton, 2007; Tallerman, 2007, 2008). However, I argue that most of these critiques miss their mark if the notion of a musical protolanguage is accepted as a starting point (see Fitch, 2010). Jespersen/Wray's model of holistic protolanguage thus dovetails nicely with the musical protolanguage hypothesis, in ways that I believe resolve many, if not all, of these criticisms (Fitch, 2006; Mithen, 2005).

Sexual Selection

A second problem with Darwin's model remains unresolved at present: his focus on sexual selection as the force driving the evolution of musical protolanguage. Appearing as it did as a few pages in a lengthy tome introducing and then extensively documenting the very idea of sexual selection, this aspect of Darwin's theory has the virtue of explaining a core aspect of human evolution using a broad principle abundantly demonstrated in the evolution of other species. As throughout his work, Darwin eschewed "special pleading" for our own species. The central difficulty for this beautiful hypothesis is posed by two ugly facts about modern human language: it is equally developed in males and females, and it is expressed very early in ontogeny, essentially at birth (Fitch, 2005a). These aspects of language differentiate it sharply from most sexually selected traits, which are strongly biased to develop in the more competitive sex (typically males), and only at sexual maturity. If anything,

human females have superior language skills to those of males (Henton, 1992; Kimura, 1983; Maccoby & Jacklin, 1974), and language is remarkable in its very early development, with at least some early tuning to phonology already occurring in utero before birth (DeCasper & Fifer, 1980; Mehler et al., 1988; Spence & Freeman, 1996).

There are several potential responses to the difficulty that these facts pose: one is to argue that during the musical protolanguage stage, sexual selection was the driving force, and song was (as in most bird species) expressed mainly in males at sexual maturity. Then, at a later stage (presumably during the evolution of meaningful language) some other selective force kicked in, so that language became equally (or better) expressed in females, and was pushed to develop early. A candidate selective force is kin communication: that selection for information transmission between parents and their offspring, or more generally between adults and their younger kin. I have suggested that kin selection drove this second stage of the evolution of propositional semantic content (Fitch, 2004, 2007; for an exploration and critique of this idea, see Zawidzki, 2006). This kin-selection scenario neatly explains the early ontogenetic appearance of language in infants (the earlier offspring begin absorbing their elders' knowledge, the better), and its bias toward females (who are the primary caregivers in all hominoids). The continued presence of meaningful speech in males is easily explained by the dual facts that immature males must also learn, and that, unusually in humans, adult males play an important role in childrearing (whether the father, or male siblings of the mother, is irrelevant to this fact). Finally, this kin-selection model has the virtue of explaining why language evolved in humans and *not* in other "musical" lineages. Humans combine an extended childhood, with ample time to acquire knowledge, with very small reproductive output. The fact that ape babies are born singly, and rarely, makes the survival of each individual hominid infant a crucial component of reproductive success in the great ape lineage (Fitch, 2007; Hrdy, 1999, 2004).

An alternative possibility is that sexual selection was, and remains, an important driving force in human cognitive evolution, including language (Miller, 2001), but that human pair bonding has "changed the rules" in significant ways, so that both sexes are choosy, and both compete for high-quality mates. Some comparative data can be cited in support of this second option. Recent data shows that female birdsong is not so uncommon as thought by Darwin, who considered female song to be a simple aberration (Langmore, 2000; Riebel, 2003; Ritchison, 1986). There is some evidence suggesting that sexual selection can indeed drive female birdsong, though it seems clear that female song is a secondary derivation of male song in most lineages (Langmore, 1996). While these observations provide some support for the idea that the dual-sex expression of human language could result from sexual selection, it is important to recognize that female song still appears to be numerically speaking

exceptional and that *any* model based on sexual selection will have difficulty explaining the extremely early development, and productive use, of language in human infants.

A final possibility is that sexual selection *never* played a role in the evolution of music or of language. The popular notion that music evolved for courtship (Miller, 2000, 2001) stands on a surprisingly weak empirical footing compared to a less obvious but better documented function of music: mother-infant communication (Trainor, 1996; Trehub, 2003a, 2003b). Mothers sing to their infants all over the world, even those who claim to be unable to sing (Street, Young, Tafuri, & Ilari, 2003), and infants both prefer song to speech, and respond to song in manifestly adaptive ways (e.g., engaging with and getting excited by play songs, and being lulled to sleep by lullabies (Trehub & Trainor, 1998)). These observations suggest that music originally functioned in a childcare context, as it continues to do today. By this model, the use of music in bonding among adults is simply a side effect of this central function, and its occasional use in courtship is a red herring (Dissanayake, 2000; Falk, 2004; Trehub & Trainor, 1998). This final possibility is clearly compatible with the kin-selection arguments advanced above, but here there would be no intervening stage of language evolution in which sexual selection ever played a dominating role. Even Darwin was occasionally wrong.

Terminological Niceties: Musical or Prosodic Protolanguage?

A final, less crucial difficulty with Darwin's model is terminological. Darwin himself seemed to conceive of his presemantic protolanguage in terms directly comparable to modern-day music (or at least he provides no indication that this is *not* the case). He concludes that "musical notes and rhythm" were present in this protolanguage, and that they were deployed "in producing true musical cadences, that is in singing" (1871, p. 56). This is why I term his model "musical protolanguage." However, modern human music consists not just of song, but also instrumental music, so this appellation might immediately have connotations of drumming, whistling, or flutes that are not, strictly speaking, relevant to language evolution. More pertinently, if we take the musical protolanguage model seriously, we must acknowledge that modern music may not necessarily preserve the state of this protolanguage precisely, and that both music and language have changed in the interim (Brown, 2000). That is, Darwin's hypothetical communication system was protomusic, not music per se. Adopting the logic of comparative reconstruction, we can then ask which aspects of modern speech and song are shared, and thereby reconstruct this earlier system (Fitch, 2005b). The central shared aspects are prosodic and phonological: the use of a set of primitives (syllables) to produce larger, hierarchically structured units (phrases) that are discretely distinctive. But two key "musical" aspects are not

shared between speech and song: namely discrete-pitched notes, and temporal isochrony (a steady beat). I have used this comparison of modern speech and song to argue for a subtly different model from that of Darwin, which I termed “prosodic” rather than “musical” protolanguage, in which protolanguage consisted of sung syllables, but *not* of notes that could be arranged in a scale, nor produced with a steady rhythm (Fitch, 2006). This prosodic protolanguage model thus includes the “sung cadence” aspect of Darwin’s model, while rejecting both his “notes” and “rhythm” (at least as normally construed). Both of these aspects of (most) modern song are, by hypothesis, more recent developments in music, not present in protolanguage. I see this as an adjustment of Darwin’s hypothesis, fully in keeping with its spirit. Furthermore, it is unclear from his writings whether Darwin would have disagreed with this adjustment.

A different reconstruction of the common ancestor of music and language, involving both discrete pitches and isochronic rhythm (as well as tone-based meaning) is given by Brown (2000). Brown also argues that his hypothetical protolanguage, which he dubs “musilanguage,” could not have evolved by normal neo-Darwinian selection and thus demands a group selection explanation. This remains its clearest, and most dubious, distinction from what is otherwise just a rediscovery of Darwin’s basic hypothesis (for critiques see Botha, 2008; Fitch, 2010).

Conclusions

I have argued that Darwin’s model for language evolution, “musical protolanguage,” suitably updated, provides a compelling fit to both the phenomenology of modern music and language, and to a wealth of comparative data. By placing vocal control at the center of his model, Darwin availed himself of the rich comparative database of other species who have independently evolved complex vocal imitation, and he thus explains two of the features of human language that set it off most sharply from nonhuman primate communication systems: vocal learning and cultural transmission. The biggest missing piece in Darwin’s model, as I see it, is a reasonable explanation of phrasal semantics (and the aspects of syntax that go with it), but this gap was filled by Jespersen by 1922. Together, these hypotheses provide one of the leading models of language evolution available today (for an enthusiastic book-length exploration see Mithen, 2005), and one that has been repeatedly rediscovered by later scholars (e.g., Brown, 2000; Livingstone, 1973; Richman, 1993). While many aspects of what has now become a family of models remain to be explored empirically (the issues surrounding sexual, kin, and group selection remain particularly unclear), this is a model worthy of detailed consideration and elaboration today. Most importantly, Darwin’s model makes numerous testable empirical predictions (for example, about the partially overlapping nature of the brain mechanisms

underlying music and spoken language, and their genetic basis) that can be addressed in the coming decades. The fact that it was born of, and supported by, the similarities between birdsong and human speech and song makes it particularly relevant to the current book, and the 200th anniversary of Darwin's birth seems an opportune time for his model of language evolution to regain the prominence it deserves.

References

- Arbib, M. A. (2005). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences*, *28*, 105–167.
- Bickerton, D. (1981). *Roots of language*. Ann Arbor, MI: Karoma Press.
- Bickerton, D. (1990). *Language and species*. Chicago: University of Chicago Press.
- Bickerton, D. (2007). Language evolution: A brief guide for linguists. *Lingua*, *117*, 510–526.
- Botha, R. (2008). On musilanguage/“HmMMM” as an evolutionary precursor to language. *Language & Communication*, *29*, 61–76.
- Botha, R. (2009). Theoretical underpinnings of inferences about language evolution. In R. Botha & C. Knight (Eds.), *The cradle of language* (pp. 93–111). Oxford: Oxford University Press.
- Brown, S. (2000). The “Musilanguage” model of music evolution. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 271–300). Cambridge, MA: MIT Press.
- Byrne, R. W., & Whiten, A. (1988). *Machiavellian intelligence: Social expertise and the evolution of intellect in monkeys, apes and humans*. Oxford: Clarendon Press.
- Condillac, É. B. de. (1747/1971). *Essai sur l'origine des connaissances humaines* (Nugent, T., Trans.). Gainesville, FL: Scholar's Facsimiles and Reprints.
- Corballis, M. C. (2003). From mouth to hand: Gesture, speech and the evolution of right-handedness. *Behavioral and Brain Sciences*, *26*, 199–260.
- Darwin, C. (1859). *On the origin of species*. London: John Murray.
- Darwin, C. (1871). *The descent of man and selection in relation to sex*. London: John Murray.
- DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science*, *208*, 1174–1176.
- Dissanayake, E. (2000). Antecedents of the temporal arts in early mother-infant interaction. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 389–410). Cambridge, MA: MIT Press.
- Donald, M. (1991). *Origins of the modern mind*. Cambridge, MA: Harvard University Press.
- Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and human speech: Common themes and mechanisms. *Annual Review of Neuroscience*, *22*, 567–631.
- Dowty, D. R., Wall, R. E., & Peters, S. (1981). *Introduction to Montague semantics*. Dordrecht: Reidel.
- Egnor, S. E. R., & Hauser, M. D. (2004). A paradox in the evolution of primate vocal learning. *Trends in Neurosciences*, *27*, 649–654.

- Falk, D. (2004). Prelinguistic evolution in early hominins: Whence motherese? *Behavioral and Brain Sciences*, 27, 491–503.
- Farrar, F. W. (1870). Philology & Darwinism. *Nature*, 1, 527–529.
- Fitch, W. T. (2000). The evolution of speech: A comparative review. *Trends in Cognitive Sciences*, 4, 258–267.
- Fitch, W. T. (2004). Kin selection and “Mother Tongues”: A neglected component in language evolution. In D. K. Oller & U. Griebel (Eds.), *Evolution of communication systems: A comparative approach* (pp. 275–296). Cambridge, MA: MIT Press.
- Fitch, W. T. (2005a). The evolution of language: A comparative review. *Biology and Philosophy*, 20, 193–230.
- Fitch, W. T. (2005b). The evolution of music in comparative perspective. In G. Avanzini, L. Lopez, S. Koelsch, & M. Majno (Eds.), *The neurosciences and music II: From perception to performance* (Vol. 1060, pp. 29–49). New York: New York Academy of Sciences.
- Fitch, W. T. (2006). The biology and evolution of music: A comparative perspective. *Cognition*, 100, 173–215.
- Fitch, W. T. (2007). Evolving meaning: The roles of kin selection, allomothering and paternal care in language evolution. In C. Lyon, C. Nehaniv, & A. Cangelosi (Eds.), *Emergence of communication and language* (pp. 29–51). New York: Springer.
- Fitch, W. T. (2010). *The evolution of language*. Cambridge: Cambridge University Press.
- Guttenplan, S. (1986). *The languages of logic*. Oxford: Blackwell.
- Harvey, P. H., & Pagel, M. D. (1991). *The comparative method in evolutionary biology*. Oxford: Oxford University Press.
- Henton, C. (1992). The abnormality of male speech. In G. Wolf (Ed.), *New departures in linguistics* (pp. 27–59). New York: Garland.
- Hewes, G. W. (1973). Primate communication and the gestural origin of language. *Current Anthropology*, 14, 5–24.
- Hockett, C. F., & Ascher, R. (1964). The human revolution. *Current Anthropology*, 5, 135–147.
- Hrdy, S. B. (1999). *Mother Nature*. New York: Pantheon Books.
- Hrdy, S. B. (2004). Comes the child before man: How cooperative breeding and prolonged postweaning dependence shaped human potentials. In B. Hewlett & M. Lamb (Eds.), *Hunter gatherer childhoods* (pp. 65–91). Hawthorne, NY: Aldine/de Gruyter.
- von Humboldt, W. (1836). *Über die Kawi-Sprache auf der Insel Java*. Berlin: Druckerei der Königlichen Akademie der Wissenschaften.
- Janik, V. M., & Slater, P. B. (1997). Vocal learning in mammals. *Advances in the Study of Behavior*, 26, 59–99.
- Jarvis, E. D. (2004). Learned birdsong and the neurobiology of human language. *Annals of the New York Academy of Sciences*, 1016, 749–777.
- Jespersen, O. (1922). *Language: Its nature, development and origin*. New York: Norton.
- Kegl, J. (2002). Language emergence in a language-ready brain: Acquisition issues. In G. Morgan & B. Woll (Eds.), *Language acquisition in signed languages* (pp. 207–254). Cambridge: Cambridge University Press.

- Kimura, D. (1983). Sex differences in cerebral organization for speech and praxic functions. *Canadian Journal of Psychology*, *37*, 19–35.
- Langmore, N. E. (1996). Female song attracts males in the alpine accentor *Prunella collaris*. *Proceedings of the Royal Society B: Biological Sciences*, *263*, 141–146.
- Langmore, N. E. (2000). Why female birds sing. In Y. Espmark, T. Amundsen, & G. Rosenqvist (Eds.), *Signalling and signal design in animal communication* (pp. 317–327). Trondheim, Norway: Tapir Academic Press.
- Livingstone, F. B. (1973). Did the Australopithecines sing? *Current Anthropology*, *14*, 25–29.
- Maccoby, E. E., & Jacklin, C. N. (1974). *The psychology of sex differences* (Vol. 1). Stanford, CA: Stanford University Press.
- Marler, P. (1970). Birdsong and speech development: could there be parallels? *American Scientist*, *58*, 669–673.
- Marler, P. (1976). An ethological theory of the origin of vocal learning. *Annals of the New York Academy of Sciences*, *280*, 386–395.
- Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, *29*, 143–178.
- Miller, G. F. (2000). Evolution of music through sexual selection. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 329–360). Cambridge, MA: MIT Press.
- Miller, G. F. (2001). *The mating mind: How sexual choice shaped the evolution of human nature*. New York: Doubleday.
- Mithen, S. (2005). *The singing Neanderthals: The origins of music, language, mind, and body*. London: Weidenfeld & Nicolson.
- Montague, R. (1974). Universal Grammar. In R. H. Thomason (Ed.), *Formal philosophy: Selected papers of Richard Montague* (pp. 222–246). New Haven, CT: Yale University Press.
- Mufwene, S. S. (2001). *The ecology of language evolution*. New York: Cambridge University Press.
- Mühlhäusler, P. (1997). *Pidgin and creole linguistics* (Rev. ed.). London: University of Westminster Press.
- Müller, F. M. (1861). The theoretical stage, and the origin of language. In F. M. Müller (Ed.), *Lectures on the science of language*. London: Longman, Green, Longman, and Roberts.
- Noiré, L. (1917). *The origin and philosophy of language*. Chicago: Open Court.
- Nottebohm, F. (1972). The origins of vocal learning. *American Naturalist*, *106*, 116–140.
- Nottebohm, F. (1975). A zoologist's view of some language phenomena, with particular emphasis on vocal learning. In E. H. Lenneberg & E. Lenneberg (Eds.), *Foundations of language development* (pp. 61–103). New York: Academic Press.
- Nottebohm, F. (1976). Vocal tract and brain: A search for evolutionary bottlenecks. *Annals of the New York Academy of Sciences*, *280*, 643–649.
- Portner, P. H. (2005). *What is meaning: Fundamentals of formal semantics*. Oxford: Blackwell.
- Richman, B. (1993). On the evolution of speech: Singing as the middle term. *Current Anthropology*, *34*, 721–722.

- Riebel, K. (2003). The “mute” sex revisited: Vocal production and perception learning in female songbirds. *Advances in the Study of Behavior*, *33*, 49–86.
- Ritchison, G. (1986). The singing behavior of female Northern Cardinals. *Condor*, *88*, 156–159.
- Senghas, A., Kita, S., & Özyürek, A. (2005). Children creating core properties of language: Evidence from an emerging sign language in Nicaragua. *Science*, *305*, 1779–1782.
- Spence, M. J., & Freeman, M. (1996). Newborn infants prefer the maternal low-pass filtered voice, but not the maternal whispered voice. *Infant Behavior and Development*, *19*, 199–212.
- Stam, J. H. (1976). *Inquiries into the origin of language*. New York: Harper & Row.
- Stokoe, W. C. (1974). Motor signs as the first form of language. In R. W. Wescott (Ed.), *Language origins* (pp. 35–49). Silver Spring, MD: Linstock Press.
- Street, A., Young, S., Tafuri, J., & Ilari, B. (2003). Mothers’ attitudes towards singing to their infants. *Proceedings of the 5th Triennial ESCOM Conference*, *5*, 628–631.
- Tallerman, M. (2007). Did our ancestors speak a holistic protolanguage? *Lingua*, *117*, 579–604.
- Tallerman, M. (2008). Holophrastic protolanguage: Planning, processing, storage, and retrieval. *Interaction Studies: Social Behaviour and Communication in Biological and Artificial Systems*, *9* (1), 84–99.
- Tomasello, M., & Call, J. (2007). Ape gestures and the origins of language. In J. Call & M. Tomasello (Eds.), *The gestural communication of apes and monkeys* (pp. 221–239). London: Erlbaum.
- Trainor, L. J. (1996). Infant preferences for infant-directed versus noninfant-directed play-songs and lullabies. *Infant Behavior and Development*, *19*, 83–92.
- Trehub, S. E. (2003a). The developmental origins of musicality. *Nature Neuroscience*, *6*, 669–673.
- Trehub, S. E. (2003b). Musical predispositions in infancy: An update. In I. Peretz & R. J. Zatorre (Eds.), *The cognitive neuroscience of music* (pp. 3–20). Oxford: Oxford University Press.
- Trehub, S. E., & Trainor, L. J. (1998). Singing to infants: Lullabies and play songs. *Advances in Infant Research*, *12*, 43–77.
- Wallace, A. R. (1905). *Darwinism: An exposition of the theory of natural selection with some of its applications*. New York: Macmillan.
- Wray, A. (1998). Protolanguage as a holistic system for social interaction. *Language & Communication*, *18* (1), 47–67.
- Wray, A. (2000). Holistic utterances in protolanguage: The link from primates to humans. In C. Knight, M. Studdert-Kennedy, & J. R. Hurford (Eds.), *The evolutionary emergence of language: Social function and the origins of linguistic form* (pp. 285–302). Cambridge: Cambridge University Press.
- Zawidzki, T. W. (2006). Sexual selection for syntax and kin selection for semantics: Problems and prospects. *Biology and Philosophy*, *21*, 453–470.