Origins of Language
Converging Evidence in Language and Communication Research

Over the past decades, linguists have taken a broader view of language and are borrowing methods and findings from other disciplines such as cognition and computer sciences, neurology, biology, sociology, psychology, and anthropology. This development has enriched our knowledge of language and communication, but at the same time it has made it difficult for researchers in a particular field of language studies to be aware of how their findings might relate to those in other (sub-)disciplines.

CELCR seeks to address this problem by taking a cross-disciplinary approach to the study of language and communication. The books in the series focus on a specific linguistic topic and offer studies pertaining to this topic from different disciplinary angles, thus taking converging evidence in language and communication research as its basic methodology.

Editors

Marjolijn H. Verspoor
University of Groningen
m.h.verspoor@let.rug.nl

Wilbert Spooren
Vrije Universiteit Amsterdam
w.spooren@let.vu.nl

Advisory Board

Walter Daelemans
University of Antwerp

Leo Noordman
Tilburg University

Cliff Goddard
University of New England

Martin Pütz
University of Koblenz-Landau

Roeland van Hout
Tilburg University and University of Nijmegen

Elizabeth Closs Traugott
Stanford University

Volume 5

Origins of Language: Constraints on hypotheses
by Sverker Johansson

Library of Congress Cataloging-in-Publication Data
Sverker Johansson
Origins of Language: Constraints on hypotheses / Sverker Johansson.
p. cm. (Converging Evidence in Language and Communication Research, ISSN 1566–7774 ; v. 5)
Includes bibliographical references and index.

P116.O76 2005
401--dc22 2004066031
isbn 90 272 3891 X (Eur.) / 1 58811 629 8 (US) (Hb; alk. paper)

© 2005 – John Benjamins B.V.
No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands
John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA
Table of contents

List of Figures viii
List of Tables ix
Preface x

1. INTRODUCTION 1

2. WHAT IS LANGUAGE? 5

3. THE THEORY OF EVOLUTION 13
   3.1 Natural selection as an abstract process 14
   3.2 Variation, randomness, and mutation 15
   3.3 Limitations and misunderstandings of evolution 16
   3.4 Cultural evolution 20
      3.4.1 Evolutionary epistemology 22
      3.4.2 The three worlds of Karl Popper 24
      3.4.3 Memetics 25
      3.4.4 The interaction of cultural and biological evolution 29
   3.5 Evolutionary theory as applied to language 31
      3.5.1 Evolution of language per se 32
      3.5.2 Co-evolution of linguistic and biological organisms 35
   3.6 The time scale of evolution 39
   3.7 Summary 40

4. HUMAN ORIGINS AND EVOLUTION 41
   4.1 Mammals 42
   4.2 Primates 44
      4.2.1 Who is our closest living relative? 48
   4.3 Hominids 51
      4.3.1 Bipedal apes 52
      4.3.2 Growing brains 60
      4.3.3 Ourselves and other humans 66
   4.4 Summary 75

5. ANATOMICAL AND NEUROLOGICAL PREREQUISITES 77
   5.1 Sound production 77
      5.1.1 Vocal tract differences 78
5.1.2 The neural wiring of the vocal apparatus
5.1.3 Control of rapid sequencing
5.1.4 When did we start to speak?

5.2 Sound perception

5.3 Brain anatomy, modularity, and lateralization
   5.3.1 Why did large brains evolve?
   5.3.2 Lateralization
   5.3.3 Modularity
   5.3.4 Brain and language in other species, fossil and extant

5.4 Summary

6. ANIMAL COMMUNICATION IN THE WILD
   6.1 Do animal calls mean anything?
   6.2 Mental states of communicating animals?
   6.3 The evolution of animal communication
   6.4 Animal syntax?
   6.5 Summary

7. CAN NON-HUMANS BE TAUGHT LANGUAGE?
   7.1 Apes
   7.2 Dolphins
   7.3 Parrots
   7.4 Patterns of non-human language acquisition
   7.5 Summary

8. LANGUAGE, MIND, AND SELF
   8.1 What is Mind — the ‘hard problem’
   8.2 What is mind — the ‘easy problem’ — and do animals have it?
      8.2.1 Theory of mind
      8.2.2 Self-awareness
   8.3 Summary

9. HYPOTHESES OF LANGUAGE ORIGINS
   9.1 Historical background
   9.2 Dimensions of language evolution hypotheses
   9.3 Adaptation vs. spandrel
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.4 Early vs. late</td>
<td>167</td>
</tr>
<tr>
<td>9.4.1 Art and technology as proxies for language?</td>
<td>168</td>
</tr>
<tr>
<td>9.5 Gradual vs. sudden</td>
<td>170</td>
</tr>
<tr>
<td>9.6 Speech first vs. gestures first</td>
<td>173</td>
</tr>
<tr>
<td>9.6.1 Mimesis</td>
<td>175</td>
</tr>
<tr>
<td>9.6.2 Mirror neurons</td>
<td>176</td>
</tr>
<tr>
<td>9.6.3 Why switch to speech?</td>
<td>177</td>
</tr>
<tr>
<td>9.7 Innate vs. learned</td>
<td>178</td>
</tr>
<tr>
<td>9.8 Summary</td>
<td>190</td>
</tr>
<tr>
<td>10. WHY DID LANGUAGE EVOLVE?</td>
<td>193</td>
</tr>
<tr>
<td>10.1 Hunting</td>
<td>195</td>
</tr>
<tr>
<td>10.2 Tool making</td>
<td>198</td>
</tr>
<tr>
<td>10.3 Sexual selection</td>
<td>202</td>
</tr>
<tr>
<td>10.4 Child care and teaching</td>
<td>206</td>
</tr>
<tr>
<td>10.5 Social relations in groups and tribes</td>
<td>208</td>
</tr>
<tr>
<td>10.6 Miscellaneous ideas</td>
<td>214</td>
</tr>
<tr>
<td>10.7 Why us and not the other apes?</td>
<td>216</td>
</tr>
<tr>
<td>10.8 Summary</td>
<td>217</td>
</tr>
<tr>
<td>11. PROTO-LANGUAGE</td>
<td>219</td>
</tr>
<tr>
<td>11.1 Proto-speech</td>
<td>219</td>
</tr>
<tr>
<td>11.2 Proto-gestures</td>
<td>222</td>
</tr>
<tr>
<td>11.3 Proto-semantics</td>
<td>223</td>
</tr>
<tr>
<td>11.4 Proto-syntax</td>
<td>230</td>
</tr>
<tr>
<td>11.5 How can all the protos be combined?</td>
<td>240</td>
</tr>
<tr>
<td>11.6 Summary</td>
<td>242</td>
</tr>
<tr>
<td>12. CONCLUSIONS</td>
<td>243</td>
</tr>
<tr>
<td>References</td>
<td>248</td>
</tr>
<tr>
<td>Index</td>
<td>331</td>
</tr>
</tbody>
</table>
List of Figures

4.1 Geological periods. *Left:* The past 225 million years. *Right:* The past 65 million years. (m.y.a. = millions of years ago)

4.2 Family tree for the major groups of primates. Not all relationships are firmly established — the tree shows the best-supported hypothesis, but alternative possibilities exist.

4.3 Family tree for the apes

4.4 Skulls of extant African apes, intended for comparison with the various fossil skulls shown in the next section. Above: Chimpanzee (left) and Gorilla (right), Below: Human.

4.5 Overview of the approximate times at which the various hominid species existed.

4.6 One possible family tree for the hominids. Few of the relationships are firmly established — the tree shows what I consider the best-supported hypothesis, but many alternative possibilities exist.

4.7 *Australopithecus boisei,* ‘Zinj’, a robust australopithecine. Note the thick arches and ridges on the skull.


4.9 Fossils of intermediate forms between *Homo erectus* and *sapiens,* that may be attributed to *Homo heidelbergensis,* found at Steinheim (left) and Broken Hill (right).


5.1 The vocal tract of a human (left) and a chimpanzee (right). Note the lower position of the larynx and vocal chords, and the longer vertical part of the supralaryngeal tract, in the human.

5.2 Brain size plotted against body size for various species of vertebrates, with different symbols for different classes, as shown above.

5.3 The different lobes of the human brain. Broca’s and Wernicke’s areas are also indicated.

5.4 Various interior parts of the human brain.
5.5 Brain size of various fossil and extant hominids, plotted against the time when they lived. Data mainly from Aiello & Dean (1990), with post-1990 finds added. Note that absolute skull volume is plotted, not normalized to body size.

10.1 Examples of tools from two early cultures. Left: Oldowan, commonly associated with habilines. Right: Acheulean, commonly associated with Homo erectus.

### List of Tables

4.1a Overview of the various hominid species discussed in the text, including alternate names for the same fossils. The usage varies in the literature; the names designated as primary in the Table are used in the text. Species belonging to the genus Homo are listed in Table 4.1b. (m.y.a. = millions of years ago)

4.1b Overview of the various species of humans (genus Homo) discussed in the text, including alternate names for the same fossils. The usage varies in the literature; the names designated as primary in the Table are used in the text. (m.y.a. = millions of years ago)

5.1 Ontogenetic stages in linguistic and non-linguistic cognitive and social development.

11.1 A summary of the starting position and possible exaptations and early evolutionary steps, in the four areas of proto-language discussed in this section.

12.1 Evaluation of the viability of different combinations of locations along the five dimensions of language evolution hypotheses discussed in Chapter 9. The salience of the text indicates the confidence level of the assessment. The evidence and arguments that the table is based on are spread out throughout Chapter 9.
Preface

The origin of our capacity for language is a complex topic, requiring input from many disparate fields, from linguistics to paleoanthropology. Specialists in any one field are often unfamiliar with the evidence from other relevant disciplines.

I perceive a need for an integration of knowledge from all relevant fields, outside as well as inside linguistics, in a single work. My purpose in writing this book is to bring together the material needed for such an integration, and to take the first steps towards the synthesis needed for a thorough understanding of the evolution of language.

My principal aim here is not to ‘sell’ any new theory of my own. What is novel in this work lies more in the synthesis of and drawing conclusions from existing data, and the systematic evaluation of existing theories. Throughout the book, it is my intention that the line of argument be data-driven, not theory-driven.

A large part of the book has the character of a scholarly review, presenting in a coherent manner the relevant evidence and theories from all the disciplines involved, with ample references to primary sources. From there I proceed to review different hypotheses proposed for the origin of language, and evaluate the hypotheses in the light of the evidence reviewed earlier in the book. This leads to firm conclusions concerning which hypotheses remain tenable, and which do not.

As the book is mainly aimed at linguists, I have chosen to place less emphasis in my review on evidence from linguistics proper, and more on fields with which a linguist may be less familiar, notably evolutionary biology, primatology, and paleoanthropology. At the same time, I have tried to keep the book readable for both linguists and non-linguists interested in the field of language origins.

An evolutionary perspective permeates the book. But I wish to emphasize here that by no means is the origins of language solely a question of biological evolution — cultural and cognitive issues are likely to have been at least as important as biology for the origin of human language, even though less hard data are available.

During the gradual evolution of this text, I have received valuable feedback on earlier versions from Terrence Deacon, Jordan Zlatev, and Peter Kitzing.

My own background is in a totally different field — I was originally trained as a physicist, and got my Ph.D. in particle physics in 1990 at the University of Lund, Sweden (Johansson, 1990), and went on to a postdoc position in astroparticle physics. But my interests already then were broader than just physics, and I took a lot of extra courses on the side, actually ending up with second bachelor’s degree in economics. And then, in the early 1990s, I took an introductory
course in linguistics, taught by Professor Gisela Håkansson, still at Lund. And I was hooked... Thanks to Gisela for being such an inspiring teacher!

But unlike too many physicists, who come into a new field thinking they know it all, I didn’t start right away writing the book to reform the field of language origins. Instead I enrolled as a regular undergraduate student in linguistics, went through all the courses, and eventually defended my master’s thesis a couple of years ago (Johansson, 2002), while still making a living teaching physics — thanks to my advisor Jordan Zlatev, whose constructive criticism cleared away a lot of fuzzy thinking.

I am currently associate professor of physics and assistant dean at the School of Education and Communication in Jönköping, Sweden, where I mainly work with teacher education. As a professor with a broad background at a small college, I’m teaching a wide diversity of courses, not just physics proper but everything from introductory philosophy to human evolution. Next year I’ll be giving my first course in linguistics, with this book as the main text.
CHAPTER 1

INTRODUCTION

For centuries and millennia, we humans have regarded ourselves as the pinnacle of Creation, placing ourselves closer to our gods than to the other animals on earth. But as we have learnt more and more about other animals, we find more and more similarities with ourselves. The features that have been held up as uniquely human, as proof of the abyss separating us from the animal world, have one by one fallen by the wayside.

Human beings are indeed in many ways unusual animals, with some very peculiar adaptations. In most respects, however, the difference between us and other animals is a matter of degree only:

- We may be the most intelligent animal on this planet, but apes and dolphins aren’t totally devoid of intelligence either.
- Our species uses tools more than any other, but other species do use, and even make, tools (Ambrose, 2001).
- We have the most extensive body of social and cultural knowledge, but other species learn from each other as well (Nagell et al., 1993), and pass on cultural patterns (Vogel, 1998; Whitehead, 1998; Boesch & Boesch, 1990; Vogel, 1999a), leading to distinct cultures in different populations of chimps (de Waal, 1999; Whiten et al., 1999; Whiten & Boesch, 2001), orangutans (van Schaik et al., 2003a; Vogel, 2003), and whales (Rendell & Whitehead, 2001; Deecke et al., 2000, but see also e.g., Mann (2001) and Maestripieri & Whitham (2001)).
- We are the only species to run a global monetary economy, but trade and bartering are not unknown in the animal world (de Waal & Berger, 2000; Hyatt & Hopkins, 1998; de Waal, 1997; Westergaard et al., 2004).
- We are the only species to replace native ecosystems with our own crops on a scale visible from orbit — but ants (Mueller et al., 1998), termites (Aanen et al., 2002), and beetles (Farrell et al., 2001) have been practicing agriculture more than a thousand times longer than we have.
- We are the only species to engage in large-scale long-distance killing of our conspecifics — but murder and tribal warfare are not unknown among other
apies (Wrangham, 1999; Wilson & Wrangham, 2003; Wilson et al., 2004; Watts, 2004).

And so on, for practically every feature that has been proposed as uniquely human...

The more we learn about our relatives, the more similarities we find, and the more kinship we can observe — and feel! (Goodall, 1998; Fouts, 1997)

There remains only one important area in which it can still be argued that we are unique, and that is our habitual use of language. No other species has anything remotely approaching our language capability, and many linguists maintain that no other species has any language at all.

Nevertheless, we have overwhelming evidence that we evolved from an ape-like ancestor in just a few million years, and so our language capabilities must have evolved as well, presumably through some sequence of intermediate stages. We share a fairly recent common ancestor with chimpanzees, not more than ten million years ago, and slightly less recent common ancestors with various other primates, which provides a starting point for language evolution, but the continuation is not so obvious. Our relatives have their own communication systems, but it is by no means obvious whether there is any homology\(^1\) between them and human language.

Among modern humans, there is no evidence that any language spoken by any group of people is in any way intermediate between ‘full’ language and any evolutionary predecessors (Blake, 2001). The same applies to any historically attested language.\(^2\) All modern languages have an equally long history of diachronic change since the origin of language in our lineage. The diachronic processes of language change is a form of language evolution (cf. Section 3.5.1), but one which is not covered in this book. The history of human language will be followed only up to the point when our full language capacity is realized, not further through the diversification of the modern-day language families.

The lack of clear evolutionary transitional forms between non-language and language renders the elucidation of the process difficult, so difficult, in fact, that many scholars have abandoned the problem as intractable, to the point of deriding or even banning any attempt at solving it.

Among linguists, the infamous dictum of the Linguistic Society of Paris in 1866, forbidding speculation on the origin of language (Trabant, 2001; Pinker, 2000), was explicitly endorsed as late as 1991 (Lightfoot, 1991, cited in Newmeyer (2000)), and still appears to hold sway in some circles:

\(^{1}\)Homology is a technical term in biology, roughly meaning similarity due to shared ancestry (Webber & Ponting, 2004; Mindell & Meyer, 2001). It is occasionally used also in linguistic contexts (Pinker, 1998b).

\(^{2}\)The sole claim the contrary that I have found, Georgiev (1984, cited in an editorial comment in Jucquois (1991) ) does not appear to be widely accepted. See, however, Newmeyer (2003b) for a more nuanced picture of possible relationships between grammar and culture, and Bichakjian (2002) for an argument for diachronic evolutionary improvement of languages.
There is a long history of study of origin of language, asking how it arose from calls of apes and so forth. That investigation in my view is a complete waste of time. (Chomsky, 1988, p. 183)

But during the past few decades, starting with a conference in New York in 1975 (Harnad et al., 1976) and then taking off with the seminal paper by Pinker & Bloom (1990), there has been a growing trend towards increased interest in language origins, with dozens of books and scores of papers published by both linguists and others.

Still, even among those who do work on the issue of language origins, there is in general a lack of unity of purpose and methods. Linguists often express their disdain or concern when non-linguists tackle the problem:

We cannot leave the discussion of language origins to those researchers who have yet to understand the concerns of modern linguistics. (Wilkins & Wakefield, 1995, p. 27 (online version)).

But the issue of language origins is by no means a purely linguistic problem — there is just as much evolutionary biology, neurology, ethology, and social psychology involved. One can equally well reverse the quote:

We cannot leave the discussion of language evolution to those linguists who have yet to understand the concerns of modern evolutionary theory.

Neither aspect can be ignored — both biologists and linguists, and preferably people who understand all the relevant fields, are badly needed if progress is to be made. The same point is made several by prominent people in this field:

We argue that an understanding of the faculty of language requires substantial interdisciplinary cooperation. (Hauser & Chomsky & Fitch, 2002, p. 1569)

The ideal scholar in this field should combine a professional training in linguistics, paleoanthropology, evolutionary biology, neurology, psychology and primatology, at the very least. (Bickerton, 2001, p. 581)

Thus, language evolution research must necessarily be cross-disciplinary in order to provide sufficient constraints on theorizing to make it a legitimate scientific inquiry. (Christiansen & Kirby, 2003b, p. 300)

My principal purpose in this book is to provide and apply such constraints from different disciplines. The main questions to be addressed here are:

The theory of evolution, as applied to language origins — Chapter 3 reviews the general characteristics of evolutionary processes, and the possibilities and limitations that they offer. Which evolving systems are relevant for the origins of language? The biological evolution of human beings is certainly part of the story — but there is more to the evolution of language than just the bodily changes from apes to humans.

The evolution of humans — what does our general evolutionary history look like, and what implications does this have for the evolution of language? Our family tree, presented in Chapter 4, provides the backdrop against which the story of language origins unfolds, and will enable us to locate important events in time and space, and provide some hints of what the first speakers may have looked like.
The biological basis of language — what anatomical and neurological structures are necessary bases for the evolution of language, and when did our ancestors acquire them? Which structures were acquired specifically for linguistic purposes, and which ones had evolved for some other purpose and were later co-opted for language? These issues are addressed in Chapter 5.

The mental and cognitive basis of language — what relationship, if any, is there between mind and language? What aspects of mind do non-humans possess? What non-linguistic cognitive abilities are associated with the evolution of language? Mind and language is addressed in Chapter 8.

The cultural basis of language — what cultural and social structures are associated with the evolution of language, and when did our ancestors acquire them? Which structures were acquired specifically for linguistic purposes, and which ones had evolved for some other purpose and were later co-opted for language? Much less is known about this issue than about the biological side of the story; some ideas are discussed in Chapter 10.

Animal communication in the wild — what features characterize animal communication, and what qualitative differences, if any, are there between human language and animal communication systems? Animal calls with language-like features are reviewed in Chapter 6.

Teaching language to animals — do any non-humans acquire language, given appropriate training or rearing? Which aspects of language are most accessible to non-humans? Chapter 7 presents various attempts to teach language to apes, dolphins, and parrots.

Hypotheses of language evolution — among the multitude of hypotheses proposed to explain the origins of language, which classes of hypotheses are inconsistent with the available evidence, and which remain tenable? In chapter 9, hypotheses are classified along five different dimensions, and information from the preceding sections is used to constrain the array of possibilities and exclude some sections of this five-dimensional space.

What drove the evolution of language — what reproductive advantages did language confer on our ancestors, causing ‘language genes’ or ‘language memes’ to increase in frequency? And why did the same factors not give chimps language? This is the theme of Chapter 10 where a variety of ideas, from politics to sex, concerning why our ancestors found language so useful, are evaluated in the light of what we know about the habits of our ancestors and cousins.

Paths of language evolution — what transitional stages may language have gone through, on the path from ape-like to human-like linguistic abilities? Finally, in Chapter 11, we try to trace the actual steps towards language that our ancestors may have gone through, and identify a possible path from ape to human.
CHAPTER 2

WHAT IS LANGUAGE?

Language is a very complex concept, far from easy to define in any stringent manner that actually captures its richness. To begin with, there are several sub-concepts hiding under the common label ‘language’. Saussure (1916) distinguished famously between ‘langue’ and ‘parole’, where langue is the systematic structure underlying parole, the actual speech events. The distinction between ‘competence’ and ‘performance’ is along somewhat similar lines, but with more emphasis on competence as a property of the individual language-user, as opposed to the more social character of langue.

Lamb (2000) makes further distinctions in the same spirit with his four indexed ‘language$_{1,2,3,4}$’. Lamb’s language$_{1}$ corresponds to the Saussurean parole, and language$_{2}$ roughly to langue. Language$_{3}$ is a label for the cognitive processes that form the basis for language$_{1+2}$, and language$_{4}$, which Lamb himself considers an improper use of the word ‘language’, is used for innate language endowments of children, as in Pinker (1994) The Language Instinct.

It should, of course, be noted that parole or language$_{1}$ is the only aspect of language that is actually observable. We infer the existence of langue from the manifest existence of parole. While essentially all linguists would agree that langue in some sense exists, there is less agreement about its structure and content. A case can be made for regarding parole as the primary aspect of language (Weigand, 2002), even though much more theoretical effort by linguists has been spent on langue. But Itkonen (2003) argues instead for the logical and methodological priority of langue, with its social and normative aspects, over both parole and language$_{3}$.

Both langue and parole, or for that matter all four of Lamb’s languages, have their place in this book, as they cannot be decoupled in their origins. Language$_{1}$, the surface manifestation of language, is impossible to explain without language$_{3}$, the underlying cognitive processes, as well as the neural processes underlying them, and as soon as language$_{1}$ is anything beyond random noises, language$_{2}$ in some sense is needed as well. And in an account of the evolutionary origins of language, language$_{4}$ necessarily plays a central role, even though it remains
an open question (addressed in Section 9.7) to what extent language is actually language-specific in any strong sense.

A distinction that may need to be made explicit, is that between the different evolutionary processes that may be involved in the origin, development and evolution of language. The evolution of language or langue is a matter of cultural or memetic evolution, whereas the evolution of language is mainly a matter of normal biological evolution. Evolutionary processes as applied to different aspects of language are further discussed in Chapter 3.

It may be useful to distinguish also between those parts of the human language capacity — language, with some echoes of language — that are used exclusively for language, and those that have other uses as well. Hauser & Chomsky & Fitch (2002a) use the terms FLN and FLB (Faculty of Language Narrow/Broad sense). “FLN is the abstract linguistic computational system alone, independent of the other systems with which it interacts and interfaces.” (Hauser & Chomsky & Fitch, 2002, p 1571), whereas FLB also includes all those other systems that interface with the FLN, and handle the entire chain from incoming sound to the FLN, and from there to outgoing sound.1 The issue of how language-specific various components of our language capacity are will be a recurring theme throughout this book.

From a functional perspective, language is, among other things, a form of communication, and it is probable that it evolved for the purpose of communication (Pinker, 1998a; Li, 2002; Jackendoff, 2002) — or even for the related but simpler purposes of affecting the behavior (Catania, 2001) or mental state (Allott, 1994) of others. But other origins are not excluded; notably the possibility that language — syntax in particular — evolved as a tool for conceptual representation has attracted a fair amount of interest, see e.g. Bickerton (1995) and Newmeyer (2003a). We’ll return to the relation between language and thought in Chapter 8, but meanwhile communication will be treated as the main purpose of language — if it were not, why would we need a lot of cognitive machinery to transform between logical content and perceptible signals (Pinker, 2003)? Or as Harris (2001) puts it:

Why, and how on Earth, should private mental states (or the physical states of biological systems) have features which only have a role in communication — and what would it explain if they did? (p. 227).

But language is by no means the only form of communication used in either the animal or the human world, and language is certainly not synonymous with communication (Bickerton, 1995). Every social animal has some form or another of communication, forming a highly diverse assemblage of communication methods (Hauser, 1997), but few, if any, of these can be regarded as languages. And language also possesses additional capabilities, on top of its basic communicative purpose (Bickerton, 1995).

1 Or equivalent in non-speech modalities.
So what is it that is so unique about language, apart from the fact that it is our main method of communication? What sets language apart from all the grunts and tail-wagging and odors and whatnot that other species use? The nature of language has been contemplated by innumerable thinkers from Plato onwards; Everson (1994a) gives a historical overview of Western thought on this issue, with Coward (1990) adding some Eastern perspectives. The relationship between language, thought, and mind, occupies a central role in this debate, to which we shall return in Chapter 8, but this does not answer the question of what makes language unique. According to Győri (2001c), the uniqueness of human language stems from its dual function, both communication and cognition, something which makes it qualitatively different from the systems of other animals.

The discrete, particulate nature of language is a key feature, without which its infinite combinatorial richness would be impossible to achieve (Studdert-Kennedy & Goldstein, 2003). Particulate systems with subunits that can be recombined into new ‘utterances’ are not totally unknown among other animals — the songs of some birds are built this way — but most animal communication either doesn’t consist of discrete units, or the units cannot be recombined into new messages with emergent meaning. And in any case the two-level discreteness of human language, the ‘double articulation’ (Martinet, 1957), does appear unique — we recombine units into an infinity of combinations both at the level of phonology and at the level of syntax.

There is one other context in nature where it has been argued that language exists, and that is in our genetic code (Schrödinger, 1944). This code, as expressed with DNA on our chromosomes, has some language-like properties: it is particulate, combining discrete units, it has words, sort of (three-‘phoneme’ sequences coding for one amino acid), and it has a kind of grammar for the decoding of strings of such words, according to which some strings are grammatical and others not, and it does have the capability of ‘infinite use of finite means’, in its own way. But in other ways it is quite unlike human language, and its use does not remotely resemble our communication. For a review of this issue, see Searls (2002), with some different perspectives in Botstein & Cherry (1997), Tsonis et al. (1997), Bodnar et al. (1997), and Sebeok (1985). Collado-Vides (1992) proposes, furthermore, that the regulation of gene expression in the cell can be treated within a grammatical formalism.

Returning to human language, one important feature commonly regarded as unique is ‘displaced reference’ (Morford & Goldin-Meadow, 2001) or ‘detached representation’ (Gärdenfors, 1996), the ability to communicate about absent things and events. Animal communication, to the extent that it is referential at all (see Chapter 6) is generally a matter of ‘cued representation’ (Gärdenfors, 1996), re-

---

2But then, McArthur (1987) argues that it may be a mistake to focus on its uniqueness: “...en soulignant les notions du caractère unique du langage ..., on néglige des aspects important du langage.” (p. 157). Perhaps ‘what makes language unique’ is the wrong question to ask?
ferring only to things for which environmental cues are present — monkeys emit
alarm calls only when they actually perceive a leopard, they cannot use the calls to
discuss leopards in the absence of leopards.

Pinker (1998a; 1999) defines language as a system with two main components:
words and grammar, a finite (though extensible) set of symbols, and a likewise
finite set of rules for combining these symbols, giving us “the infinite use of finite
This quote expresses a major part of what makes human language appear unique.
Whether it actually is unique, and qualitatively different from animal communica-
tion systems, will be explored in Chapter 6.

But the division of language into words and rules is certainly not original with
Pinker — it has a history stretching back at least to Aristotle with his onoma/rhema
pair. Launey (2004) identifies two separates paths here, one from Aristotle to
the modern functionalists, focused on predication, and the other, from Dionysius
Thrax to Chomsky and colleagues, focused on syntax, with the main difference
being whether syntax emerges from semantic and cognitive considerations, or if it
is an autonomous system of relations between abstract parts of speech.

McArthur (1987) focuses on the ‘word’ component, calling language “une tech-
nologie sémiotique” (p. 159), a theme echoed in Sebeok’s (1985) argument for
semiosis as the root of language. In the same spirit, Deacon (1997) emphasizes
the symbolic aspects of language. The distinction between symbols and other sig-
nals is important here. A signal can be just about any information that an organism
receives, with or without communicative intent or awareness or reference. Shared
reference is, in contrast, a key property of symbols:

Conventional symbol systems are grounded in an intersubjective meaning-field in
which speakers represent, through symbolic action, some segment or aspect of re-
ality for hearers. (Sinha, 2001, p. 4, emphasis in original).

Implicit in this quote is also that symbols, unlike signals, are strictly and inten-
tionally communicative, and presuppose a shared awareness of the universe of
discourse. But a system characterized only by joint reference is regarded by Sinha
(2001) as merely proto-symbolic — the emergence of full symbolization requires
construal, entailing the elaboration and conceptualization of the joint references of
proto-symbolic systems. The origin of symbols and semantics is further discussed
in Section 11.3.

Symbolic language requires at least third-level intentionality sensu Grice (1957)
— the speaker intends the listener to understand what the speaker intends. This in
turn requires both speaker and listener to represent each other as beings capable
of understanding and intent, i.e., both must have a ‘theory of mind’ (Gärdenfors,
1996). Issues surrounding our theory of mind (Premack & Woodruff, 1978) are
discussed in Section 8.2.1.

But the main thrust of much of modern linguistics has been aimed at the ‘rule’
component, the grammar, following the second of the tracks of Launey (2004) dis-
cussed above. The syntactical work of Noam Chomsky (1957; 1965; 1982; 1986;
What is language?

1995, among others) is of course seminal here. This has led to an emphasis on syntax as the core of language, and the principal defining feature of language. The formal structures of Lamb’s language are regarded as central, the principal object under study is competence, with performance largely disregarded. A grammar in the Chomskian sense is an autonomous abstract computational structure, used for generating and parsing grammatical strings. Much work has been done on the syntax of different languages within this paradigm, leading to subtle, complex and highly abstract rules for generating grammatical sentences.

Language acquisition becomes a matter of identifying the correct target grammar, a difficult task for which an innate Universal Grammar is needed. The Chomskian view of syntax as a monolithic innate structure does not easily lend itself to evolutionary explanations, but Section 9.7 and 11.4 aim to resolve this problem.

Syntax, in the Chomskian generative paradigm, is an autonomous system totally decoupled from questions of meaning and function. But even though the sleep of colorless green ideas is as grammatical as it is furious, not a few linguists feel that the Chomskian ‘syntactocentrism’ (Jackendoff, 2002) may be a mistake. Other aspects of language cannot be neglected, and something central in language is missing when syntax is divorced from meaning. The growth of cognitive linguistics and cognitive semantics (Lakoff, 1987; Langacker, 1987; Ungerer & Schmid, 1996; Allwood & Gärdensfors, 1999), can be seen as an attempt to restore the connections between language and general cognition, between syntax, function, and meaning (Fauconnier, 1999). Jackendoff (2002) can be seen as an attempt to move away from syntactocentrism while remaining within the spirit of the Chomskian paradigm.

Cognitive linguistics returns to the first of Launey’s (2004) two tracks, the one stretching from Aristotle to functionalism and beyond, with language use based on and emerging from semantics and cognition. Performance, the actual use of language is in focus, rather than competence. Language is regarded as a dynamic process, not an abstract and static grammar (Langacker, 2001; Taylor, 2002a).

Key here is our experience of the world, our perceptions and how we conceptualize them and build cognitive structures (Ungerer & Schmid, 1996). Logical rules and objective definitions do not form a prominent part of our cognitive and conceptual structures, so it makes little sense to use them as a base for our syntactic and semantic theories. Central to cognitive linguistics is how linguistic form is strongly affected by meaning. But there is no direct link from the world ‘out there’ to our linguistic descriptions of it — everything is filtered through our cognition, and colored by how we conceptualize the situation.

To begin with, individual concepts are not built through formal definitions — prototypes and family resemblance are regarded as more important (Lakoff, 1987); cf. Wittgenstein (1953). Concepts are organized in hierarchical categories, but these are built neither top-down nor bottom-up. Instead it is argued that we start
with basic concepts somewhere in the middle of the hierarchy\(^3\) and build from there both upwards towards generalization and downwards towards specification — ‘tree’ is a more natural concept for us than either ‘oak’ or ‘plant’.

Metaphor and metonymy are regarded not just as literary and rhetorical devices, but as important conceptual and linguistic tools, heavily used for extending conceptual structures into new areas, and forming new concepts by analogy. Mapping of conceptual structures between domains is the central process here (Lakoff & Johnson, 1980). A simple metaphor can be seen as just a projection of structure from a source domain to a target domain, but more complex cases are ubiquitous and routinely used and created in language. Fauconnier & Turner (1998) present a more general framework for handling cross-domain mapping, with their concept of ‘mental spaces’, and the operation of conceptual integration, ‘blending’ between them, going from input spaces to a generic space, and from there to the output space, the blend. The relation between metaphor and blending is further discussed in Grady et al. (1999).

Syntactical structures are in cognitive linguistics analyzed principally through conceptual and attentional structures, such as image-schemas, figure/ground, trajector/landmark, frames, and scripts. It is evident in these analyses how both our experience of the world and our attention colors our syntax (Ungerer & Schmid, 1996). Talmy (2000) elaborates on an analysis of the semantics of the grammatical subsystem and closed-class words, within a cognitive framework.

The generative and cognitive perspectives on language have rather different implications for language origins. In the cognitive perspective, language would emerge through the evolution of human cognition, and explanations for its origins can hardly be divorced from explanations for our general cognitive evolution. In the generative perspective, on the other hand, language is an autonomous faculty, disconnected from cognitive and functional considerations, rendering evolutionary explanations somewhat problematic. As the generative perspective poses more obstacles for the evolution of language than does the cognitive perspective, and generativists more often oppose evolutionary explanations, the generative perspective will receive more explicit attention in this book, with analyses of the empirical force of their objections to language evolution.

\(^3\)Mandler (2004b), however, presents evidence that young children may start their conceptualization of the world higher up in the hierarchy.
Further reading

CHAPTER 3

THE THEORY OF EVOLUTION

The purpose of this chapter is to summarize some biological basics that are needed in order to discuss the evolution of language. First the theory of evolution by natural selection is presented as a general concept, together with some extensions and limitations of evolutionary processes. After that, some ideas concerning the application of evolutionary theory to language origins are discussed.

Evolutionary ideas have a long history. Some aspects were discussed already by Aristotle and Heraclitus in classical times (Wilkins, 2003), as well as by numerous Enlightenment writers. That all species may have had a common ancestor was first proposed by Pierre Maupertuis (1745), and both Immanuel Kant (1790) and Erasmus Darwin (1795), grandfather of the famous Charles, entertained similar ideas.

But evolution was not really taken seriously until Charles Darwin (1859) presented both a plausible mechanism — natural selection1 — and a mass of empirical evidence. *The Origin of Species* (Darwin, 1859) is still a good introduction to the theory of evolution, though more recent works are of course also available, a few examples of which can be found in the Further Reading section at the end of this chapter.

After Darwin’s seminal work, evolution was rapidly accepted among biologists. Today, it is totally unthreatened as the ruling paradigm in biology2 — as Theodosius Dobzhansky (1973) puts it: “*Nothing in biology makes sense except in the light of evolution.*” (p. 125). Nevertheless, among non-biologists evolutionary ideas still encounter resistance. As discussed on page 161, Chomsky and other leading linguists have had their reservations, as do, unsurprisingly, various

---

1Natural selection was not strictly original with Darwin — as is well known, it was independently discovered by Alfred Russell Wallace, whose work was presented together with Darwin’s at a joint session of the Linnaean Society (Darwin & Wallace, 1859). It is less well known that the core concept of natural selection was published already in 1818 in an obscure work by William Charles Wells, and again in 1831 by Patrick Matthew, though neither attracted any attention at the time (Wilkins, 2003).

2It is sometimes claimed that e.g., punctuated equilibrium (Eldredge & Gould, 1972) threatens the paradigm. But the debate around punctuated equilibrium is firmly within the evolutionary paradigm — see e.g., Gould (2002), and Section 9.5 below — and is mainly a matter of emphasis and perspective, however acrimonious.
religious fundamentalists, e.g. Morris (2001) and Tassot (2004). But evolutionary concepts are troublesome not only to linguists and religious fanatics. Ferrari & Chi (1998) argue that these difficulties are due largely to the ontological structure of evolutionary explanations, with its emphasis on process, something which differs markedly from the event ontology prominent elsewhere in scientific explanations, and also from the essentialist thinking common among humans; cf. page 161. Related to this problem, but on a more fundamental level, is the Sapir-Whorf-like argument of Blinkhorn (2001) — it is very difficult to discuss adaptive structures with their pseudo-design, without using words and grammatical constructions with teleological implications. Language is good for discussing intent and purpose, but not at all good for discussing undirected processes that nevertheless have apparently purposeful results. Blinkhorn (2001) connects this ‘mentalistic bias’ in language with theories relating the origin of language with the emergence of a theory of mind (see Chapter 8), and Weiss (2002) provides further discussion of what he calls “the irrepressible illusion of teleology” (p. 4) in evolution.

3.1 Natural selection as an abstract process

The most important concept in the theory of evolution is natural selection. The process of natural selection is applicable to a very large class of systems, far beyond the confines of biological evolution for which it was originally intended. The prerequisites for natural selection to work are:

1. A class of objects exists, the objects in which are similar but not identical; some variation between objects exists in some features.
2. The objects have different rates of reproduction and/or mortality.
3. The differences in reproduction/mortality rates have some correlation with the variable features.
4. If the objects reproduce, then the features of the copies are correlated with those of the original (parent) object(s).

Given the above conditions, the objects will evolve. Objects with features correlated with higher reproductive rates, or lower mortality rates, will increase in frequency, at the expense of their less successful brethren.

This description of natural selection is intentionally phrased in highly abstract terms. The objects are by no means limited to being living creatures, but can be anything from universes (Harrison, 1995) to religious dogma (Cullen, 1999; Gottsch, 2001); results from evolutionary theory at this level of abstraction can

---

3 This is echoed in Steven Pinker’s comment on Chomsky’s anti-evolutionary stance (cf. page 161):

I don’t think he [Chomsky] likes the style of explanation that comes from Darwin’s theory, of a lot of statistical variation and incremental advantage being repeated iteratively over hundreds of thousands of generations. His style is more Platonic, where there must be a way of deducing specifics from general principles. (Pinker, quoted in Shermer (2001)).
be applied to both. For slightly different, but equally abstract, accounts of natural selection, see Hull et al. (2001), Calvin (1997) or Gabora (1997).

In the abstract description above, there is nothing corresponding to the biological distinction between genotype and phenotype. In principle, such a distinction is not needed for natural selection to do its job — but conceptual clarity may be gained by separating the two roles of what was above called ‘objects’:

1. Storing the information that needs to be transferred to the next generation, in order to maintain the correlation of features between parent and copy.
2. Interacting with the world, and either succeeding or not in producing a copy carrying the above-mentioned information.

Dawkins (1976) generalized the geno-phenotype distinction with his concepts of replicator and vehicle, where the replicator is the information carrier and the vehicle is whatever equipment the replicator uses to interact with the world in its quest for replication. In biology, the replicators are normally genes, and the vehicles are mainly the bodies of organisms. The ‘vehicle’ concept was later generalized and renamed interactor by Hull (1980, cited in Wilkins (1998a)). It should, however, be noted that the interactor need not be separate from the replicator, and there is no need for ‘the’ interactor to be a single coherent entity like the biological organism — Dawkins’ The extended phenotype (1982) is an extended argument for the possibility of more distributed interactors, a point he also makes in much briefer form in Dawkins (1994). See Szathmáry (2000), Nanay (2002), and Godfrey-Smith (2000) for more on replicators.

### 3.2 Variation, randomness, and mutation

Natural selection is not the only component of evolution. The variability that natural selection acts upon must come from somewhere, usually from random mutations. ‘Random’ is used here in the special sense of random with respect to reproductive success (fitness), with no correlation between the probability of any specific mutation and its effect on fitness. There is, however, no requirement that the mutations are random — natural selection works perfectly fine regardless of the source of variation, as long as the rate of introduction of new mutations is not larger than the rate at which natural selection can weed out failed variants. Darwin’s major insight was that evolution can work, and generate the appearance of design, even if mutations are random and undirected.

---

4The actions of replicators are described here with mentalistic language, giving the misleading impression that replicators are conscious agents with intentions. This is common in biology, and should not be interpreted literally, but instead as shorthand for the underlying evolutionary processes that produce the appearance of design and intentions, despite being mindless and undirected. As Blinkhorn (2001) noted, human languages are more suited to describing intentionality than evolutionary processes.

5Note, however, that there are many more specific mutations that decrease fitness or leave it unchanged, than there are that increase fitness.
Jean Baptiste Pierre Antoine de Monet, Chevalier de Lamarck (1809), formulated his own theory of evolution a generation before Darwin, based on the inheritance of acquired characters. This acquisition of characters can be regarded as non-random mutations. Lamarckian evolution would have worked in biology, even better than Darwinian, if it hadn’t been for the empirically observed facts that the pattern of biological mutations is indeed effectively random with respect to fitness, and the interactor’s acquired characters aren’t transmitted back to the replicator, and are thus not replicated. See Wilkins (2001) for a review of Lamarckism, as well as a critical perspective on the role of Lamarckism in cultural evolution.

There are considerable amounts of randomness elsewhere in the evolutionary process as well, making the outcome non-deterministic, and occasionally causing changes in entities that do not correlate with higher reproductive potential, a process known in biology as ‘genetic drift’. But the random aspects of evolution can never be responsible for the exquisite and complex adaptations that are ubiquitous in living creatures, not least in ourselves. Natural selection is the only reasonable scientific explanation here. But the process is nevertheless subtler than what is usually presented at school.

3.3 Limitations and misunderstandings of evolution

Evolution is a remarkably powerful process, displaying the appearance of ingenuity and creativity and design despite its partially random basis. But there are a few limitations of evolutionary processes, as well as a couple of common misunderstandings, that may be worth pointing out:

*Evolution does not plan ahead.*

A feature is selected (or not) on the basis of its *current* utility, nothing else. An ape will not acquire a structure because it in some distant future will be vital for an as yet unattained language capability; it will be selected for only if it’s of some use to the ape right then and there (though possibly for some different purpose). This has an important corollary: if a feature evolved because it’s useful for language, then language was in place before that feature evolved.

*Evolution will find local optima only.*

The process is limited by the features already present, and will only take routes that are accessible by genetically small steps from the current situation, with every step being advantageous (or at least not too harmful). This means that it’s a lot easier and more likely to have an existing structure adapted for a new purpose, than to build a new structure from scratch (Wimsatt, 1999).
This can lead very far, but not everywhere, and not in a straight line. The process can easily get stuck in functional but suboptimal solutions, and every organism carries a lot of evolutionary baggage, accumulated along its historical path.

*Evolution does not take into account ‘the good of the species’ or any other higher-level goals.*

In the case of biological evolution, a feature is selected (or not) on the sole basis of whether it promotes the spread of copies of the genes of its individual possessors. A mutation giving an ape some language capacity, will not spread and become common just because it benefits the tribe of that ape — only if it actually benefits that individual ape, and its close relatives carrying copies of the same gene with the same mutation.

In theory, group-level features, and evolution at the group level, are possible (Price, 1972; Johnson & Boerlijst, 2002), but no clear examples are known from real life that cannot be explained by inclusive-fitness effects at a lower level. For reasonable values of the relevant parameters, individual selection will overwhelm any group-level effects, unless the group is so tightly integrated that it effectively functions as a single organism, e.g. an ant colony. Csányi (2001) invokes group selection between human tribes in connection with the origin of language, but it remains to be shown that humans are sufficiently ant-like in their social organization.

The same applies in principle to other evolutionary processes such as cultural evolution — a feature will spread and increase in frequency if and only if it increases the fitness of whatever entity carries the feature (though some care is required in identifying the relevant entities).

*Evolution does not repeat itself.*

The evolution of a complex feature is a rare occurrence, so it is very unusual for a complex feature to evolve twice in different organisms. The corollary of this is that if we observe the same complex feature in two related organisms, we can safely assume that it evolved only once, and that their common ancestor possessed it already (Byrne, 2000). In the case of language, this means that any language-related features displayed by e.g., chimpanzees today, were present already in the common ancestor of us and chimpanzees, and did not evolve for human-level linguistic purposes.

If evolution does appear to repeat itself, with the what appears to be the same complex feature turning up in unrelated animals, this implies either that the similarities are merely superficial, as with the wings of birds and bats, or that the feature isn’t as complex as it appears, and is not a difficult evolutionary step.

---

6This applies even in the rare cases of apparent group selection, e.g. Wilson (1997), though the details can be subtle.
In principle, this non-repeatability applies also to non-biological evolution, such as cultural evolution (see Section 3.4 below). However, the ‘family tree’ of cultural features need not coincide with the biological family tree, and the presence of the same cultural element in two organisms need not indicate that their common biological ancestor had that element (or had culture at all, for that matter). But the corresponding cultural conclusion can be drawn, that the presence of the same complex cultural feature in two separate cultures indicates the existence of a common cultural ancestor.

A corollary of non-repeatability is that evolution, beyond very minor changes (Teotonio & Rose, 2001), is effectively irreversible (Darwin, 1859; Dollo, 1893).

*The original and current functions of a feature need not be the same.*

Questions like ‘What good is half a wing?’ are sometimes rhetorically asked in the context of attempts to prove the impossibility of gradual evolution of complex features. But already Darwin (1859) provided an adequate response, mainly in terms of half a wing being useful for some other purpose than powered flight. Darwin’s argument has been elaborated and extended, and nowadays it is generally accepted that wings and feathers had other uses before birds started flying with them.

Exaptation, a word coined by Gould & Vrba (1982), refers to this type of feature that evolved for some other purpose, or as a simple side effect of the evolution of unrelated features (in which case Gould & Lewontin (1979) call it a spandrel), but turned out to be, quite accidentally, useful for a new purpose, different from that for which it evolved. This process of co-opting, with new selection pressures superseding old ones and ‘hijacking’ a feature for a new purpose, is further analyzed in Shelley (1999). The role of exaptations and spandrels in the evolution of the human language capacity is further discussed in Section 9.3.

Previously, exaptations were called preadaptations, a usage still occasionally seen, e.g., in Shelley (1999), Corballis (1999), Christiansen & Kirby (2003b), Hurford (2003b) and Botha (2001a). The word ‘preadaptation’, however, can easily but erroneously give the impression that the future use of a ‘preadaptation’ was planned in advance — see the first point in this section.

Clark (2000) commits a related fallacy in the context of language evolution, arguing that the first speaker cannot have had anybody to speak to, so how could language get started? This ancient argument — it goes back at least to Lucretius in the first century B.C. (Everson, 1994b) — would be troublesome only if the origin of language were a sudden leap in a single step from no language at all to full modern grammar. But, as discussed in Section 9.5, this is extremely unlikely. If language evolved gradually, in many small steps, there will never be a lonely ‘first speaker’ with nobody to talk to — the differences in language ability

---

7 Nowadays, this is mainly the province of religious evolution-deniers. For a recent and superficially more sophisticated example of the ‘half-a-wing’ fallacy, see Behe (1996)
within the same population at any given time will never be large. And comprehen-
sion is generally easier than production, developing earlier in ontogeny and likely
in phylogeny as well — both humans and other animals consistently understand
more than they can say (Burling, 2000). Comprehension of the signals of others
is possible even without there being any communicative intent on the part of the
sender — very likely our pre-linguistic ancestors were already trying to figure out
each other’s intentions (Bickerton, 2003). The ‘first speaker’ need only evolve the
ability to exploit this comprehension exaptation, a point made already by Condillac
(1746) (see p. 158). From there on, the signals could evolve into wordlike symbols
through a process of conventionalisation or ritualisation (Condillac, 1746; Burling,
2000).

In general, the evolution of systems that are tightly coupled, either between or
within organisms, is not a problem. The system can have evolved gradually from
a less tightly coupled state. Complex, tightly coupled, superficially monolithic
systems are not rare in biology. But the fact that a system is inextricably monolithic
today does not necessarily imply that it has always been so. There are several
possible types of evolutionary paths into tightly coupled complexity (Thornhill &
Ussery, 2000). Lenski et al. (2003) demonstrate the evolution of complex features
in computer simulations of evolving entities.

Evolution is not a matter of ascending a ladder towards an ultimate goal.
There are actually three errors in this ladder metaphor:

1. There is no well-defined notion of general ‘progress’ in biological evolution,
no clear up or down on the ladder, no ‘higher’ or ‘lower’ species (Mogie, 2000).
Every evolutionary step is in the direction of increased fitness right then and
there, but what happens to be more fit here and now need not be ‘better’ in any
absolute sense, and might be totally unfit in a different environment.

2. The pattern of evolutionary change is rarely a single sequence of ancestors and
descendants — the typical family tree looks more like a bush than a ladder,
with lots of branches diverging in different directions. Our own history is no
exception — two million years ago, and again just a few tens of millennia ago
(Gibbons, 2004), there were several different lineages of proto-humans alive at
the same time (see Figure 4.5). This also means that an ancestral species need
not disappear just because a new species evolves — if the speciation event
involves only a subpopulation of the ancestral species, the rest of the ancestral
species is unaffected. This is occasionally used by creationists as an argument
against human evolution in general: ‘If we’re descended from apes, how come
there are still apes around?’, an argument that is firmly rooted in this erroneous
view of evolution as a ladder.

For discussions of some of the complexities surrounding the concept of ‘progress’ in this context, see
Radick (2000b) and van der Meer (2000).
3. Evolution is not directed towards any goal. As already noted, it takes whatever direction happens to increase fitness here and now, and does not take into account the distant future. Even in the special cases of goal-directed evolution like artificial breeding of dogs or intentional language change, changes can only be effected indirectly, by increasing the here-and-now fitness of the desirable dogs and words. In the case of dogs, this is done simply by providing the desirable ones with mates and ample food — directing language change is more tricky. As for evolution ‘in the wild’, its course has meandered and bifurcated so much over the gigayears that it much more resembles a random walk from a simple starting point, than a goal-directed process with a purpose in mind.

That the ladder is a fallacy is clearly demonstrated by *Homo floresiensis* (Brown et al., 2004), discussed on page 65. *H. floresiensis*, with its tiny brain and late date, shows that (1) in an unusual environment, smaller brains may be ‘better’, enhancing fitness, (2) there were several species of humans living in parallel in the recent past, forming a bush, not a ladder, and (3) human evolution may in the ‘right’ circumstances change direction totally, not moving towards any *sapiens*-like goal.

*Organisms do not evolve in a vacuum.*

They adapt to their environment — but they also shape their environment, and other evolving systems in the environment adapt in response. The result is a complex weave of interacting systems, the results of which can be quite different from the naïve picture of an organism adapting to a static environment (Nowak & Sigmund, 2004; Wuketits, in press).

### 3.4 Cultural evolution

Culture is a concept that is not entirely trivial to define, and there is little consensus among students of culture about what the proper definition should be (Boesch, 2003; Byrne et al., 2004). But without worrying about the finer details, we can apply ‘culture’ as a label for the complex patterns of social behavior, beliefs, institutions, and artifacts that are characteristic of all human societies. The cultural patterns of humans are strongly variable between groups, but are often reasonably homogeneous within groups, and preserved from generation to generation with some degree of fidelity. This cultural continuity within groups can be explained in three different ways (Aunger, 2001):

1. Genetic inheritance (biological evolution).

2. Individual learning (different individuals learning from the same material context arriving at the same result).
3. Transmission (social learning, memes).

Human cultural universals may have a genetic basis, but the vast complexity of culture makes it vulnerable to the ‘poverty of the genes’ argument (see page 186), and the wild variability of culture in both space and time is simply too fast to be explained by the slow and cumbersome process of biological evolution (Boyd & Richerson, 2001).

Individual (non-social) learning, including the ‘ecological inheritance’ of Laland & Odling-Smee (2001), appears unlikely as an explanation for human cultural continuity, both because it requires a degree of identity between the thought processes of different individuals that is hardly consistent with our observations of learning in non-cultural contexts, and because of the quintessentially social nature of human culture. Furthermore, processes that are straightforwardly interpreted as the social transmission of culture are readily observed in the interactions of human adults and children.

The situation may differ in other animal species, where individual learning and imitation may play a significant role in those rudiments of culture that have been observed — but it may be argued that a common behavioral pattern does not qualify as culture, unless it is based on a common understanding, transmitted in a process of social learning (Fragaszy, 2003). In *Homo sapiens* transmission, cultural learning, clearly emerges as the by far most plausible explanation for cultural continuity within groups and across generations. Tomasello (1999a) argues for cultural learning as the main adaptive breakthrough of humans, leading to our cumulative growth of culture and knowledge, and playing a major role in the evolution of human cognition.9

The transmission of culture entails the transfer of some kind of information describing the culture to a new member, and the construction of a replica of that information in the mind of the new member. Cultural evolution is the process of evolution applied not to the reproduction of biological organisms, but to this transfer of cultural patterns and components of social systems (Campbell, 1965; Pagel & Mace, 2004). The application of evolutionary ideas to the study of cultures has a long history, but has remained peripheral in the field. The reasons for this may be both practical — the application is not easy — and cultural — students of culture commonly regard concepts from the natural sciences as foreign, if not with outright hostility (Alvard, 2003) — and historical — past attempts at applying evolutionary concepts in cultural studies have a mixed and partially unsavory record. There is a regrettable tendency also for the debate to become polarized, with proponents arguing that evolution works everywhere, and opponents arguing that since culture and biology aren’t identical, evolution cannot work in culture (Gabora, 2001).

---

9This does not mean that learning and social transmission are unimportant in other animals — see e.g., Mameli (2002).
The core logic of cultural evolution remains the same abstract process of natural selection described on page 14 above, but beyond this any analogies with biological evolution should be handled with care, since there are several important differences between biology and culture (Wimsatt, 1999; Gabora, 2001). To begin with, the biological concepts relevant to evolution, such as genes, reproduction, mutations, organisms, and species, are reasonably well understood and can be reliably operationalized. The corresponding concepts in cultural evolution, however, remain elusive and difficult to define and operationalize, and it is not at all clear whether all the biological concepts mentioned above actually have relevant analogs in culture (Wimsatt, 1999). Attempts to shoehorn culture into an exact biological template are doomed to fail. But the mere fact that there are differences between biology and culture is not in itself sufficient to invalidate an evolutionary approach to culture, as long as the distinguishing features of culture are taken into account, and the conceptual machinery of Darwin is appropriately adapted (Gabora, 2001).

One of the aspects making cultural evolution more difficult to study than biological evolution, is that the relevant biological concepts generally correspond to concrete measurable physical objects and events, whereas cultural concepts are of a more abstract nature. Even though aspects of culture may be instantiated in artifacts and minds, it is not really appropriate to say that a culture is its artifacts. Artifacts may be manifestations of a certain cultural concept, but the concepts themselves are independent of the artifacts. The ontological status of a cultural concept is in fact not at all obvious, and has been the source of some confusion.

3.4.1 Evolutionary epistemology

Evolutionary epistemology, baptized by Donald Campbell (1974) and reviewed in Callebaut & Pinxten (1987) and Gontier et al. (in press), is closely related to the field of cultural evolution. It concerns the application of evolutionary reasoning to the origin of cognition and knowledge. There are major subdivisions within evolutionary epistemology, sufficiently distinct that it may be unfortunate that the same name covers all of them (Vollmer, 1987; Gillièron, 1987):

The biological evolution of cognition

There are many components of the human mind that appear to be innate, with a long list of proposed cognitive universals that goes back to the geometrical axioms of Euclid, and Kant’s category of synthetic a priori knowledge (Kant, 1781). It

\[10\] Even though biology is more complex than is often appreciated outside the field (Speel, 1997). For example, the species concept as normally defined is applicable only to sexually reproducing organisms, and, as noted by Dawkins (1982), the biological gene is very difficult to define stringently.

\[11\] Though Campbell (1997) regrets this choice of name, preferring to call it ‘Selection theory’, after Simmel (1895) and Baldwin (1909).
is argued in this context by Herbert Spencer (1883), William James (1890), and most influentially Konrad Lorenz (1941), as well as more recently by e.g., Vollmer (1987), Shepard (2001) and Schwartz (2001), that knowledge that for us is perceived as *a priori* self-evident, is actually evolutionary adaptations of our cognition to the world around us, so that Kant’s synthetic *a priori* knowledge is really evolutionarily *a posteriori* (Campbell, 1997; Red’ko, 2000). Such ‘knowledge’ is not even expected to be true in any strict universal sense, as adaptation is a purely local process, at best producing knowledge that works well enough in the local environment (Wozniak, in press). One of the traditional *a priori* truths, Euclid’s parallellity axiom, has indeed been found to be empirically false (Einstein, 1915; Dyson et al., 1920), though it remains an excellent approximation in our everyday world.

This type of evolutionary epistemology may have implications for the innateness debate in linguistics — see Section 9.7 — but does not really concern cultural evolution. Chomsky, oddly enough considering his stance on linguistic innateness, argues against this form of cognitive innateness, stating that: “We cannot appeal to this *deus ex machina* to explain the convergence of our ideas and the truth about the world. Rather, it is largely a lucky accident...” (Chomsky, 1988, p. 158).

**The ontogenetic development of cognition**

Jean Piaget (1972) has done the seminal work here, with his ‘Genetic Epistemology’, describing the adaptive mechanisms behind the development of cognitive abilities in children. It is debatable to what extent Piaget’s model is evolutionary in any strict sense — Gili`eron (1987) would exclude it from evolutionary epistemology, whereas Apostel (1987) has a broader view of evolution in general and evolutionary epistemology in particular. In a linguistic context, Piaget’s approach is best known from his debate with Noam Chomsky about language acquisition in 1975 (Piattelli-Palmarini, 1994).

**The cultural evolution of knowledge**

This has mainly been studied in the context of scientific knowledge. This type of evolutionary epistemology is foreshadowed already in the 19th century by Huxley (1893, cited in Pennock (2000)) and Simmel (1895, cited in Coleman (2002)), as well as in Karl Popper’s influential work *The Logic of Scientific Discovery* (1959). It has later been more fully developed by Campbell (1974), Toulmin (1972), as well as Popper in his *Objective knowledge — an evolutionary approach* (1979). This type of evolutionary epistemology is reviewed and further analyzed by Wilkins (1995; 1998a), who proposes an epistemology even more Darwinian than Campbell’s and Popper’s, in contrast with Hussey (1999) who characterizes evolutionary epistemology as fundamentally Lamarckian. In what follows, Popper’s framework for cultural evolution will be used.
3.4.2 The three worlds of Karl Popper

Karl Popper (1979) proposed a framework for handling the ontology of both cultural, mental and material entities in a consistent manner. Popper’s three-world system distinguishes what might be called three different planes of existence:

**World 1:** The normal physical universe, inhabited by everything from galaxies to gnats.

**World 2:** Mental space, inside our heads, inhabited by our thoughts and emotions and memories.

**World 3:** The world of disembodied abstract ideas, notably cultural concepts.

The three worlds are not entirely independent of each other. World 1, the mindless physical universe, is not dependent on the other worlds. World-2 objects do appear to depend on the existence of world-1 brains. World-3 objects, however, are less straightforward. Popper regards a world-3 object as existing regardless of whether it is currently instantiated in either world 1 or world 2 — e.g., a melody keeps on existing as a world-3 object even if all recordings of it (world-1 objects) are destroyed and all memories of it (world-2 objects) are forgotten. In this respect, Popper’s world 3 has much in common with Plato’s world of Ideas or Forms, despite their totally different metaphysical status:

... Plato was the discoverer of the third world. ... Plato’s third world was divine; it was unchanging and, of course true. Thus there is a big gap between his and my third world: my world is man-made and changing. (Popper, 1979, p. 122).

Furthermore, for Plato, only idea-world objects really exist, objects in the other worlds being mere shadows of them. Popper, on the other hand, does not appear to grant primacy to any of the worlds, but regards all three worlds as in some sense real.

A world-3 object will apparently survive indefinitely, once it is born. But unlike the Ideas of Plato, it is indeed born, created when it is first instantiated in world 2 — the first time an idea emerges in a mind, as a world-2 object, the corresponding abstract world-3 object is born. In this sense, the contents of world 3 is a human creation, as stated by Popper in the quote above.

World-3 objects lead a passive existence within their world 3, and interact only through their world-1 and world-2 instantiations. But when a world-3 idea is instantiated in a mind, as a world-2 object, that mind can communicate the idea to others, planting new instantiations in new minds, either through direct communication or through intermediate world-1 objects (books and other information-storing artifacts). In a new host mind, the idea will interact with other ideas in the same...

---

12 Alternatively, if intentionality, consciousness and meaning can actually emerge in a computer or a robot, e.g., as envisioned by Zlatev (2001a), world-2 objects would instead be hosted by world-1 silicon. But some world-1 substrate is required, unless one postulates the existence of disembodied souls, which would live in world 2 alone, but which would bring us too far into theological territory.

13 Or even in world 1, at least in special cases — Popper invokes the example of a computer-generated table of logarithms (Popper, 1979, p. 115).
mind, and likely be modified. The modified version is in a sense a new world-3 object, a descendant of the old one, similar to the parent but not identical. This is something quite close to Darwin’s concept of descent with modification. But not all world-3 objects reproduce equally often — some ideas are more likely than others to be communicated, and adopted by new minds. Thus, some world-3 ideas will be fruitful and spread into many minds, generating many descendants, whereas others will be forgotten by all minds and stranded uninstantiated in world 3, for all practical purposes extinct.

In the case of scientific hypotheses, Popper makes the parallel with Darwinism explicit:

[T]he growth of our knowledge is the result of a process closely resembling what Darwin calls ‘natural selection’; that is, the natural selection of hypotheses: our knowledge consists, at every moment, of those hypotheses which have shown their (comparative) fitness by surviving so far in their struggle for existence:... (1979, p. 261, emphasis in original).

This evolutionary perspective also resolves the slight tension between the individual and the social in Popper’s discussion of world 3. World-3 objects are born out of individual human minds, and are thus not intrinsically social — but their evolution is fundamentally social, requiring an interacting population of world-2 instantiations, just as biological evolution is normally based on an interacting population of organisms, and it is only through this social process that the objective knowledge that is the goal (and title) of Popper (1979) can emerge in any meaningful sense.

3.4.3 Memetics

Popper mainly limited his studies of world-3 evolution to the growth of scientific knowledge. Memetics, introduced by Dawkins (1976), is an attempt to extend Popper’s Darwinian analogy to include not only scientific hypotheses but all inhabitants of Popper’s world 3, cultural, linguistic, or whatever, into a reductionist framework usable for micro-level studies of their evolution. The core concept here is the meme, intended to be the cultural equivalent of the biological gene, and closely corresponding to Popper’s concept of a world-3 object. Other precursors of the meme concept include the ‘symbolate’ of White (1959, cited in Gatherer (1997)), and there are hints of similar reasoning also in the writings of Peirce (1931–35) on the evolution of symbols (Gatherer, 1997). For a popular review of memetics, see Blackmore (2000a).

The literature on memetics is not entirely clear on the proper meaning of the meme concept, something which has severely hampered progress in the field. Some memeticists call just about anything remotely mental or cultural a meme, diluting the concept into meaninglessness, whereas others insist on strictly operationalizable definitions of memes as something physically tangible, gaining stringency but possibly losing in the process the ‘soul’ of the concept that makes it
fruitful. This difficulty of defining a core concept in a way that is both stringent and fruitful is, however, not unique to memetics — in practice, concepts acquire their meaning through usage and joint reference, only rarely by formal definition. What hampers memetics is not so much the lack of definitions, as the lack of consistent usage and joint reference. I would agree here with Wittgenstein, who in a different context argues for the acceptance of fuzzier concepts:

> [A]ll concepts need not be unitary and crisp, but rather that different phenomena falling under the concept may be connected by ‘family-resemblances’. (Zlatev, 2001a, p. 163, citing Wittgenstein (1953)).

Similarly Hull (2001) argues for an emphasis more on finding applications where memes are fruitful and less on definitional battles.

There is no total agreement either on which of Popper’s worlds is the appropriate home of memes — see e.g. the exchange between Gatherer (1998), Wilkins (1999) and Speel (1999), where the contest appears to be between world 1 (material artifacts) and world 2 (mental objects), though Wimsatt (1999) adds the level of institutional structures, clearly relevant to cultural evolution but not easy to fit in either world 1 or 2.

In my opinion, placing memes in world 3, following Popper’s (1979) analysis of the evolution of scientific hypotheses described above, is more appropriate, and lends itself to more fruitful analysis, despite the fact that world 3 is not directly observable — the study of biological evolution was an eminently successful enterprise for a full century before biological genes were directly observed in any reasonable sense. The effects of genes, before the discovery of DNA, were observable through the resulting phenotypes, and similarly the effects of world-3 memes are observable through their instantiations in world-2 thoughts (as manifested by world-1 behavior) and world-1 artifacts. Thus, knowledge of the biochemical substrate of genes (DNA) was not needed for the development of Mendelian genetics as a science — likewise, the concept of memetics is not rendered incoherent by our inability to identify the substrate of memes (Deumert, 2003). As emphasized by Aunger (1999), Gabora (1997) and Dennett (1995), the evolutionarily relevant aspect of both genes and memes is their information content, not how they are physically encoded:

> What is preserved and transmitted in cultural evolution is information — in a media-neutral, language-neutral sense. Thus the meme is primarily a semantic classification...!

Possibly a parallel with quantum physics may be illustrative as well. The wave function $\psi$ plays a role in quantum physics that is as vital as that of a gene in biological evolution or a meme in memetics. The history of quantum physics is a century of unparalleled success, both as an eminently fruitful research paradigm and as a source of practical applications that include all modern computer equipment, with a rock-solid empirical foundation — but also a century of abject failure in the quest for the identity and nature of the wave function. Some of the brightest minds of the 20th century, including Albert Einstein, spent decades trying to define and operationalize and make sense of the wave function. The end result is that we have still no idea what a wave function really is — but we can state with some confidence that it is not an ordinary world-1 object in any reasonable sense, not anything tangible or directly accessible. But the fact that quantum physics is worse off than memetics regarding the ontological status of a central concept, does not prevent quantum physics from being an enormously productive science.
tion, not a *syntactic* classification that might be directly observed in ‘brain language’ or natural language. (Dennett, 1995, p 353-354, emphasis in original).

Still, the lack of observable meme substrate means that the molecular-level reductionist revolution in genetics that has been so successful in recent decades is unlikely to have a memetic counterpart in the foreseeable future.

Just like with cultural evolution in general, the analogy between genes and memes should not be pushed too far. Not all biological concepts have clear memetic counterparts, and like I said about culture above, attempts to squeeze memetics into a biological straightjacket are doomed to failure. For an example of this fallacy, see the attempt by Holdcroft & Lewis (2000) to criticize memetics, which they do precisely by stretching the gene-meme analogy beyond the breaking point, and then complaining that it broke down. With my placement of memes in world 3, the world-3 entity would be the replicator *sensu* Dawkins (1976) and its world-1&2 instantiations would be interactors (or vehicles). Other authors make different and sometimes contradictory choices here, with e.g., Gabora (1997) arguing for world-2 replicators and world-1 interactors, whereas Gatherer (1998) proposes the opposite.

Whatever the location of memes, it is quite clear that memes need hosts in order to reproduce. This is sometimes interpreted as memes being akin to viruses, who also co-opt host reproductive machinery in order to replicate themselves — but given the very different nature of a meme and its host, it might be better to use the less loaded image of a meme needing fertile minds in the same sense as a plant needing fertile soil for its reproduction.

Still, even if we disregard the definitional debates, there are a number of weaknesses in the meme-gene analogy that even proponents of memes recognize (Dennett, 1995):

- Low-fidelity replication, with an effective mutation rate that may be high enough to wash out any selection effects.
- Mutations may be purposely directed by humans, opening the door to Lamarck.
- Blending inheritance, instead of Mendelian particulate genes, washing out innovations.

There are also other oddities in the criticism of Holdcroft & Lewis (2000), like their complaint:

Memetics does not explain the emergence of the very first memes. (p. 180)

At the same time, they dismiss the same issue for genes:

But though it is interesting to speculate about the origins of DNA and RNA, evolutionary or otherwise, no one supposes that an inability to answer such questions invalidates explanations [...] in terms of variations in the genotype. (p. 170).

Liane Gabora has later (Gabora, 2004) reconsidered the issue of cultural replicators, and no longer accepts memes as replicators at all, since memes (or ideas) do “not consist of coded self-assembly instructions.” (p. 127). Here she has a point, in that memes indeed do not code in the sense that genes do. The replicator in Gabora (2004) is instead the *worldview*, a coherent autocatalytically closed network of ideas.
Independent invention of the ‘same’ meme may be common. The first objection, if true, is fatal to the idea of memetic evolution. But we have no reason to believe it to be generally true — it is not difficult to find examples of memes that have persisted with little change over many generations. Linguistic memes, for one thing (see Section 3.5.1 below), but also many traditions and religious ideas.

The second objection simply means that memetic evolution need not be purely Darwinian in a strict sense, which is not a serious problem — culture is different from biology — though Wilkins (2001) argues that cultural evolution really is Darwinian. Hull (2002) doesn’t find any Lamarckian inheritance in culture either, in any coherent sense. In any case, an evolutionary perspective on culture “does not ... necessitate a denial of human agency” (Deumert, 2003, p. 25).

The third objection, blending inheritance, does not invalidate memetic evolution per se, but would, if true, hamper the kind of cumulative growth of complexity and adaptedness that is a major strength of Darwinian evolution (Dennett, 1995) and characteristic of human cultural evolution (Tomasello, 1999a), rendering memetics uninteresting as a possible explanatory framework for cultural complexity. Wimsatt (1999) argues, however, that cultural evolution is not blending in this sense. The fourth objection, independent invention, does not invalidate memetic evolution, but makes the reconstruction of memetic phylogenies very difficult.

Dennett’s objections are all empirical claims about the nature of memes. It is likely that his claims are true for at least some classes of cultural phenomena, rendering a memetic analysis of those phenomena unlikely to be fruitful. However, this need not be the case in all areas of culture — the fact that memes are useless in some areas does not prove that they are useless in all (Aunger, 2001). Wilkins (1998a) addresses the concerns of Dennett (1995) in the context of the evolution of scientific theories, Heinrich & Boyd (2002, cited in Heinrich & McElreath (2003)) does the same in a more general cultural context, and these concerns as applied to language will be addressed in Section 3.5.1 below.

Within limited domains, some serious and careful empirical work has been done within a memetic paradigm, The study of birdsong, e.g., Baker (1996) or Majoros (2002), is one area where memetics has been fruitful. Dirlam (2003) presents a memetic analysis of the rise and fall of data analysis methods in developmental

---

17 Rose (1998) suggests that the enormous difference in cultural complexity between humans and other animals, discussed in Tomasello (1999a), is due to humans crossing a fidelity threshold in social learning, with fidelity getting high enough for cultural evolution to ‘take off’.

18 Though Deumert (2003) seems to argue for the opposite position in the paragraph preceding the one containing the quote above. There she says that “[a] result of human agency, cultural information (...) cannot be conceptualised as an autonomous replicator which evolves (in analogy with genes) independently of its vehicle.” (p. 24). But here she is simply wrong, misinterpreting the ‘selfish gene’ stance. The fate of a gene is certainly affected by the fate and actions of its vehicle, but it can still evolve. Both memes and genes are subject to the whims of their human vehicle, who may choose or not to transmit both memes and genes. If anything, memes have a less intimate relation with human vehicles than genes do — from the meme’s perspective, human agency is simply part of their environment.
research, as well as a more general framework for memetic analysis. Butts & Hilgeman (2003) attempt to deduce the memetic structure potentially underlying religious beliefs of Americans. Suicide is studied from a memetic perspective by Marsden (1998). There is also active research on what idea-intrinsic factors determine the transmissibility and memorability of novel ideas, e.g., the studies reviewed in Barrett (2000), as well as Jacobson (1998), concerning religious concepts. Research like this, even though not all of it is done within an explicitly memetic paradigm, is clearly important for placing memetics on a sound empirical footing.

Unfortunately, these modest successes are often overshadowed by the approach seen in the writings of some of the more vocal proponents of memetics, e.g., Blackmore (2000b), an approach characterized more by enthusiasm than by stringency. This over-enthusiasm has caused a counter-reaction of skepticism against everything memetic, a polarization that does not facilitate progress in the study of cultural evolution. But the skeptics do have a point both in that much work in memetics is wildly speculative and conceptually confused, and in that some proponents of memetics are naïve and ignorant about the cultural sciences (Kuper, 2001; Bloch, 2001).

Nevertheless, despite all the weaknesses and problems with memetics in its current state, it contains a valuable core idea, that can contribute to our understanding of the evolution of culture and ideas, and language. If the application of evolutionary ideas to culture is ever to become more than hand-waving, something along the lines of memetics will be needed. But before that will happen, memeticists need to get their act together, and demonstrate that the memetic research program can actually produce results that are both novel and empirically supported. Outside limited domains, the empirical basis of memetics remains modest.

3.4.4 The interaction of cultural and biological evolution

Cultural and biological evolution are by no means two completely independent processes. Cultural evolution wasn’t even possible until biological evolution had provided us with the requisite cognitive skills, and cultural evolution even today is very likely biased by our biological predispositions.

But there are links in the other direction as well, from culture to biology. For one thing, modern human material culture will keep alive many people who would have died young in a hunter-gatherer tribe, relaxing the selection pressures otherwise weeding out e.g., people like me who need glasses to see that there’s a lion hiding in the grass.

Incidentally, the suicide studies may lead to a clear example of a meme enhancing its own fitness at the expense of the reproductive success of its host, something which would help to establish the reality of memes as independent replicators — cf. the ‘selfish gene’ concept of Dawkins (1976).
This kind of gene/culture coevolution is far from a new idea — it has roots in the early 19th century, well before Darwin, in the writings of Lamarck (1809; 1820, cited in Delisle (2000)), who proposed a model of human origins heavily based on the interaction and feedback between biological and cultural processes.

One interesting case of gene/culture coevolution is the retention of lactase (a milk-digesting enzyme) in adults in certain cultures. Most humans lose the ability to produce lactase as soon as they are weaned, and cannot drink significant amounts of milk as adults. But in some human populations with both a long cultural tradition of herding and a dietary need for nutrients most easily obtained through the consumption of milk, a biological evolutionary change has taken place, allowing them to drink milk as adults (Fuller, 2000; Kiple & Ornelas, 2000).

Our mate preferences are strongly affected by cultural factors, which means that cultural evolution directly modifies the biological reproductive success of different individuals, changing the course of biological evolution — in a culture where e.g. people with freckles are culturally regarded as particularly attractive, people with freckles will have more opportunities to reproduce, and genes for freckles will spread in the population. Durham (1991, cited in Rose (1998)) proposes that similar meme/gene coevolution may be a significant process in humans.

A subtler effect of culture on biology is the Baldwin effect (Baldwin, 1896; Weber & Depew, 2003; Turney et al., 1996; Steels, 1997; Lachapelle et al., in press), known from various other fields of behavioral evolution. The Baldwin effect is a pseudo-Lamarckian process in which a learned behavior can become genetically encoded over evolutionary time. The logic of the process is the following:

If a certain behavior is highly advantageous for the members of a species, the behavior may, in a cultural species, first turn up as a cultural innovation, with no biological basis. The first generation of children after the innovation may start out by learning the behavior without innate help other than general learning abilities. But in each new generation, those individuals who acquire the behavior quickly and effectively are favored by natural selection. This is a clear recipe for the evolution of an innate predisposition to acquire the behavior, or even a dedicated acquisition device. In the long run, the behavior may under certain circumstances become entirely genetically hardwired, bypassing the acquisition step.

However, as shown by Deacon (2004b), the circumstances under which the Baldwin effect will work are rather narrow, making Baldwinian evolution unlikely in practice. In many cases, the presence of a learning mechanism may just as well relax the selection pressure on innate mechanisms, leading to their degradation, or ‘devolution’.

The ‘genetic assimilation’ of Waddington (1942; 1957) is another pseudo-Lamarckian idea, based on developmental canalization, invoked by Briscoe (2003) and Dor & Jablonka (2001) in their scenarios of language evolution. It is sometimes mixed up with the Baldwin effect, e.g., by Dor & Jablonka (2001), though they are quite distinct in principle (Deacon, 2004b).
3.5 Evolutionary theory as applied to language

The evolution of language is usually discussed exclusively in terms of the biological evolution of the human language faculty. This is a limited and problematic view, since several other evolving systems are involved, evolving at different rates, and with complex mutual interactions (Mameli, 2001; Christiansen & Kirby, 2003b; Catania, 2001). The most important evolving system here is language itself, considered as a meme-complex. Memetic evolution of language will be discussed in Section 3.5.1 below. But there are also other evolving systems that are relevant for language:

- **Evolution of other human bodily and mental faculties**, that are now used for linguistic purposes, but may have arisen as adaptations for some other purpose, as exaptations. This is a matter of normal biological evolution, but their evolution need not be originally connected with the emergence of human language, and may well have taken place in our remote proto-ape ancestors, which largely removes the timing constraints discussed below.

- **Evolution-like processes in ontogeny**
  - The ontogenetic development of neural connections in the brain has some parallels with natural selection (Changeux, 1985; Edelman, 1987; Czikó, 1995; Deacon, 1997; Sireteanu, 1999; Chenn & Walsh, 1999; Buller & Hardcastle, 2000). In those cases that have been studied in detail, e.g., motor neurons and vision (Deacon, 2000), the neural connections are found not to be genetically preprogrammed. Instead, neurons at first form connections in large numbers, rapidly and to all appearances randomly. As neural input arrives, those connections which are heavily used multiply and are strengthened and those which rarely carry any signals wither and are eventually pruned. The result is a pattern of connections that is adapted to handle the type of input received during this sensitive period. It is not entirely obvious how an innate language acquisition device with a genetically determined universal grammar (Chomsky, 1965) could arise through such a process of ‘neural Darwinism’; cf. the discussion of ontogenetic plasticity on page 111.

  - The process of language acquisition is not purely deductive. Variation and selective retention appears to play a role as well, as in the genetic epistemology of Piaget (1972) (see p 23), in a way reminiscent of hypothetico-deductive science — or evolution.

- **Evolution of human social systems, both biological and cultural evolution** (Hill, 1997; Laland et al., 2000; Caporaal, 1996). Humans, like most other mammals, certainly have a set of biological adaptations for social life, which have evolved in the usual Darwinian fashion. But the wild diversity of social systems in different human cultures demonstrates that our social system — unlike that of most other animals — is not genetically biologically determined in any
strong sense. Instead, the most important evolutionary process in this context is cultural evolution, with our genetic adaptations for sociality as a more-or-less deeply buried substrate.

3.5.1 Evolution of language per se

Language, as an abstract structure, is not some pre-existing essence, a goal for a teleological evolutionary process to strive towards, contra Bever (1982), but instead an evolving entity of its own (Darwin, 1871; Wilkins, 1998b; Keller, 1989; Diller, 1997), a huge and successful system that can be regarded as an entity (or population of entities) in Popper-world 3 (see Section 3.4.2 above, as well as Popper (1979)).

The discussion of evolution-like processes in the history of language actually predates Darwin, beginning with the conclusion of Sir William Jones in 1786, that Sanskrit, Latin, and Greek all were descended from a common ancestor through a process of gradual modification. These studies significantly influenced Darwin (1859) as he transferred the concept of descent with modification from linguistics to biology (Gatherer, 1997). In *Descent of Man* (1871), Darwin further developed the parallels between language evolution and biological evolution (Radick, 2002).

Within linguistics, the study of the phylogeny of languages has remained a vital field of research ever since, though usually purely descriptive, without explicitly invoking evolutionary parallels except at a very superficial level. In the 19th century, there was some contact between historical linguists and evolutionary biologists, but it was fairly short-lived (Hull, 2002; Croft, 2002). For quite some time in the 20th century, biological metaphors were shunned by linguists (Deumert, 2003), but there has been some recent revival of interest (Hull, 2002; Rexová et al., 2003).

Still, the processes of language change may well be open to evolutionary interpretations. Language can be modeled in terms that closely parallel Popper’s (1979) evolutionary epistemology, with scientific hypotheses as evolving world-3 entities. ‘Language’ in the abstract sense would similarly be a world-3 entity, and its implementations in human minds would be world-2 objects, and of course the communicative use of language encodes it into sound patterns and other world-1 objects. This view of language has considerable similarities with memetics as well, and may be analyzed within a memetic paradigm, particularly since at least some of the objections against memetics do not apply to language. Returning to Dennett’s (1995) list from page 27 above:

- **Low-fidelity replication**, does not apply to language, as language is clearly transmitted from generation to generation with only minor changes.
- Mutations may be purposely directed by humans, does apply to language, but does not invalidate an evolutionary approach (Wilkins, 2002).
- Blending inheritance may apply at the phonetic level, but not at the level of the lexicon or grammar — if two parents use different words for the same concept,
their children will learn one or the other (or likely both) but will not blend the words together.

- Independent invention of the ‘same’ meme is unlikely to be a problem, given the conventional and arbitrary nature of language.

Within this Popper-inspired quasi-memetic framework, language evolution can be considered on several different levels. Popper (1979) appears to regard ‘language’ not as just a world-3 entity, but as a major part of the framework for world 3, with a role much like the spacetime of world 1 (Einstein, 1915; Johansson, 1997b):

Main thesis: our conscious subjective knowledge (world 2 knowledge) depends on world 3, that is to say on (at least virtually) linguistically formulated theories.

(Popper, 1979, p. 74, emphasis in original).

But ‘language-as-spacetime’ is not a tremendously useful perspective for language evolution. On less abstract levels, there are several linguistic entities in different worlds that are better candidates for evolutionary considerations.

**Meme level.**

Reductionistically, one may adopt the equivalent of the ‘gene’s eyes view’ of Dawkins (1976), which would be the smallest linguistic units that can be coherently replicated, the memes of language. Memes are notoriously difficult to pinpoint in many cultural contexts, but language, being both more coherent and more thoroughly analyzed, may offer some hope of success. A non-exhaustive list of plausible linguistic meme candidates might include:

- Individual sounds.
- Individual words or morphemes (Worden, 2000).
- Individual grammatical rules (or parameter settings in a Universal-Grammar framework).
- Utterances, either generic (Zlatev, 1997) or actually occurring ones (Croft, 2000).\(^{21}\)

This would appear to be a useful level of analysis for the study of contemporary language change, on a rather short time scale. The spread of a newly-coined word, for example, is about as clear an example as one can get of memetic evolution.

**‘Organism’ level.**

The organism level in biology may be regarded, following Dawkins (1976), as a coherent set of genes working together as a team, forming a common vehicle, though as shown by Santelices (1999) the organism concept in biology is not trivial. Individual genes of a human being do not do anything useful on their own —

\(^{20}\)Or ‘lingueme’, as Croft (2002) calls them.

\(^{21}\)Croft (2000) builds his theory of language evolution around utterances as the basic units, and natural selection among utterances as the central process.
they are meaningful evolutionary units only in the context of all the other genes of our genome. Similarly, individual language memes like the words or rules or utterances mentioned above are not meaningful in isolation, only as parts of a coherent system with, at a minimum, a set of generic utterances and a set of words to fill the slots in them, and more normally consisting of a large battery of grammatical rules and tens of thousands of words as well as all the various other bits and pieces that make up a language. The lowest-level entity in which all these meme-level pieces are gathered together in a coherent whole would be the idiolect of an individual language-user (Mufwene, 2002).

In biology, the lowest-level entity in which all human genes are working together as a coherent whole is an individual human being, as a biological organism. An idiolect can be regarded as a ‘linguistic organism’ in the same sense. It is quite accidental that the organism levels of biology and linguistics very nearly coincide, and perhaps unfortunate, as it may invite confusion as well as over-extension of biological analogies; this will have to be kept in mind.

This organism level may be the most fruitful for the study of the origin of the human language capacity. Very little interaction can be expected between individual language memes and individual human genes, so the meme/gene level is less likely to yield interesting insights into this issue. Direct interaction, and possible co-evolution, may instead be expected at the system level, between the human being as a system, with emergent properties beyond the sum of the genes, and the idiolect as a system of, but similarly beyond the sum of, language memes. This interaction will be discussed more at length in Section 3.5.2 below.

One might have considered languages, rather than idiolects, as linguistic individuals. If all speakers of a language had actually spoken exactly the same way, with no differences between idiolects, then languages would have been tenable individuals. Different idiolects in the same community are indeed similar — but not identical. We all know that no two individuals speak exactly the same way, even if they are members of the same speech community (Davis, 2001). A fact that is less well known (and is often brushed aside in theoretical linguistics) is that there are considerable individual differences also in grammaticality judgements (Johnstone, 2000). Given that there is no total unity among the speakers of the ‘same’ language, it is not clear that a language exists as a coherent entity in any sense other than as a set of idiolects (Davis, 2001; Mufwene, 2001; Mufwene, 2002). In this way, languages more resemble biological species than biological individuals, and are relegated to the next level below.

‘Species’ level.

In biology, a species can be regarded as a set of organisms that are mutually reproductively compatible. Alternatively, from a gene perspective, a species is a gene pool within which genes can flow freely. A linguistic analogy of the species concept would then be either a population of mutually compatible idiolects, or a
‘meme pool’ within which language memes can flow freely. With either perspective, a linguistic species concept is indicated that is quite close to our everyday notion of a language;\textsuperscript{22} cf. Croft (2000).

One aspect that has been invoked as a problem for the language-species analogy, and the culture-biology analogy in general, is the frequency of hybridization and horizontal transmission between languages and cultures, unlike traditional Darwinian evolutionary theory where normally it is assumed that only your direct descendants inherit your genes (Gould, 1987). But the formation of new species through hybridization (Hull, 2002; Wirtz, 1999; DeMarais et al., 1992; Goodfriend & Gould, 1996)\textsuperscript{23} and horizontal transmission between species (Syvanen, 1994; Poelarends et al., 2000; Brown & Doolittle, 1999; Olendzenski et al., 2001), does occur in biology as well. And at least hybridization mergers are rare between languages (Hull, 2002). Cultural evolution does, however, permit wholesale horizontal transmission of information (Takahasi, 1999) on a scale rarely encountered in biology, at least among us eukaryotes. Loanwords is a linguistic example of horizontal transmission. Wholesale hybridization between languages is also far from unknown (Croft, 2000; Ansaldo, 2001)

The species level of analysis is relevant for the study of the historical development of languages and language families, but hardly for the ultimate origins of language. And, as already noted, the diachronic development of human languages will not be covered in this book. For work in that direction, see e.g. Rexová et al. (2003), Searls (2003), Gray & Atkinson (2003), and Ben Hamed & Wang (2004), who are all successful examples of the use in historical linguistics of tools developed by biologists for the study of the relationships of biological species. Mufwene (2004) discusses speciation and extinction of languages in an evolutionary framework. Bichakjian (2002) also deserves mention in this context, with his attempt to treat the evolution of Indo-European languages within a thoroughly Darwinian framework.\textsuperscript{24}

Wilkins (2002) reviews some further issues in this kind of linguistic evolution.

3.5.2 Co-evolution of linguistic and biological organisms

Each human has an idiolect, his or her own version of language. The idiolect is for this purpose regarded as one individual organism; the inner structure of this

\textsuperscript{22}In biology, our intuitive ‘folk biology’ notion of what a species is, across cultures, corresponds remarkably closely with the biological species concept, much more closely than the correspondance between folk and scientific concepts in other areas and on other levels (Atran, 1998). Our intuitive ‘folk linguistics’ might also be similarly reliable at the species level, despite the old joke of defining a language as a dialect with an army.

\textsuperscript{23}It’s the same Gould in Gould (1987) and Goodfriend & Gould (1996), despite the apparently opposite conclusions concerning the role of hybridization in biology. Gould’s magnum opus on biological evolution (Gould, 2002) does not clarify the issue.

\textsuperscript{24}I have some caveats about the details of his functional hypotheses, though, and particularly about his treatment of directionality and progress in evolution; cf. the ‘ladder’ metaphor on page 19 above.
‘organism’ is presumably built from individual language memes, but need not concern us here. The habitat of an idiolect is an individual human mind, where the idiolect resides as a symbiont — or parasite (Mufwene, 2004). Idiolects reproduce whenever somebody acquires a language — language acquisition does not take place in a vacuum, pre-existing ‘parent’ idiolects are always present as sources of language input. The resulting idiolect of the language acquirer is the descendant of the idiolects that provided input, in the Darwinian sense of descent with modification required for an evolutionary process — the ‘child’ idiolect is normally very similar to, but not identical with, the ‘parent’ idiolects.

Language evolution, in this sense — cf. Christer Johansson (2001) — can be seen as a process of natural selection between our individual idiolects, with the most fit idiolects contributing the most to the idiolects of the next generation of people. The features of an idiolect that provides input to many language acquirers become more common in the next generation, whereas the features of an idiolect from which nobody acquires language disappear.

The features mentioned above correspond to the meme-level components of an idiolect. These memes, entities in Popper-world 3, are instantiated in world-2 idiolects, and the evolution of idiolects can reductionistically be regarded as a sequence of changes in the frequency of instantiation of different memes. New linguistic memes are created in individual human minds, same as for other world-3 objects, in the context of the world 2 of that individual mind, including its resident idiolect, but may then spread (or not) in a social process establishing their conventionalized meaning in world 3.

The evolution of idiolects (and individual language memes) is in several respects as much Lamarckian as Darwinian:

- Idiolects can acquire new features during their lifetime, and pass on those features to their offspring. New words are being acquired throughout the entire lifespan, and changes in grammar and phonology are not rare either. A complication here is that there is no clean demarcation between features acquired ‘at birth’ à la Darwin and features acquired later à la Lamarck, but it is nevertheless quite clear that the latter process plays a more prominent role here than in biological evolution.

- ‘Mutations’ of an idiolect are not necessarily random, but can be directed towards increased fitness. This includes both failure to acquire features that are difficult to learn, and the deliberate addition of features, e.g., new words that enhance communication, or speech patterns that affect social status.

The fitness of an idiolect is to some extent determined by the biological fitness of its host, i.e. the host’s reproductive success, since your biological descendants

---

25 The same mind can be the host for several idiolects, in the case of bilinguals or multilinguals.
26 Note that ‘parent’ here does not imply that the hosts of the parent idiolects are necessarily the biological parents of the language acquirers, though in practice that is commonly the case, nor does it imply that the number of parent idiolects is necessarily equal to two. See Wimsatt (1999) for more on the complexities of memetic parenthood.
commonly have your idiolect as a major source of language input. High biological reproductive success means having many surviving children, which indirectly also means that your idiolect will be a primary source of input data for language acquisition for as many new idiolects as you have children.

Apart from the reproductive success of the host, a large number of social and cultural factors will also likely influence the reproductive success of an idiolect (Nettle, 1999; Croft, 2000).

More interesting in a linguistic context are factors that are intrinsic to language, and that may affect the reproductive success of an idiolect (beyond the reproductive success of its host). Some selection for communicative functionality can be expected (Bates & MacWhinney, 1982; Nettle, 1999; Bichakjian, 2002; Briscoe, 2003), largely because an idiolect that’s more efficient for communication enhances the biological fitness of its host, but also because others may be more likely to adopt memes from a communicatively superior idiolect, and language learners may be more likely to learn from (and get a larger fraction of their idiolect from) a better communicator. Communicative functionality may be a matter of intrinsic logic, having a grammar that respects semiotic constraints and avoids ambiguity, or a matter of respecting human processing constraints (Deacon, 2003a; Givón, 1995).

Ease of processing more generally may also affect the fitness of an idiolect. This applies both at the neuromuscular level — some sounds are easier to produce than others, as shown by their earlier acquisition in ontogeny — and at the cognitive level — some possible grammars may require more short-term memory than others.

Campbell (1997) discusses the process whereby words evolve to match what he calls ‘striking discriminanda’ (p. 23, online version), giving language a conceptual basis ‘cutting Nature at its joints’. But since mutations of idiolects, including the coining of new words, may well be in the direction of increased functionality as well, due to human design, it is difficult to distinguish selection effects from design effects.

In the transmission process itself there will be severe selection for learnability (Deacon, 1997; Briscoe, 2003) — an idiolect that is difficult for children to acquire will not be passed on effectively; either the children will acquire some other idiolect (from the other parent or from someone else) or they will acquire a modified idiolect that’s not identical to the parent’s, presumably modified in the direction of improved learnability. Creole formation may be an extreme example of such modification, with the pidgin of the parents being modified into the creole of the children.

This selection for learnability applies regardless of whether the children have an innate language acquisition device or not — the evolution of idiolects will be driven towards learnability with whatever cognitive equipment children happen to
have (Newport, 1982). Johansson (1997a)\textsuperscript{27} presents a computer simulation of such evolution towards learnability, and Briscoe (2002a) simulates the coevolution of language and a language acquisition device.

At the same time as idiolects are subject to learnability selection, children are subject to selection for the ability to acquire an idiolect, and parents for the ability to ensure that their children learn. Language is so vital in human society that failure to acquire an idiolect that is communicatively usable will severely reduce biological reproductive success, both directly, and through other learning being hampered by the lack of communication.

Within modern human societies, with well established languages, selection relating to language, both biological and memetic, may be expected to be largely stabilizing (Deacon, 2003b), since our idiolects are already quite well adapted to their ecological niches, and humans are likewise well adapted to language acquisition.\textsuperscript{28} This means that the main effect of selection is to gradually weed out ‘failures’ — such as SLI children on the biological side, or functionally detrimental inventions on the idiolect side. Ordinary diachronic language change over the few thousand years for which we have solid data, basically resembles a random walk around roughly the same level of functionality — there is no support either for claiming that the fitness of Italian idiolects is higher or lower than that of Latin ones, or for claiming that modern Italians are better or worse at language acquisition than Romans.

Significant directional evolution today is observed mainly in the formation of new creole languages from pidgins. But in the distant past, before either language or the human language capacity had attained their modern level of refinement, significant idiolect evolution can be expected to have occurred. As soon as our species-unique capabilities for cultural transmission (in which idiolect transmission is included) had reached the threshold needed for cumulative cultural evolution (Tomasello, 1999a), this cultural evolution of language would pick up speed as well, growing from whatever primitive communication systems were in use at the time (which must nevertheless have been non-trivial, or cultural transmission would be ineffective) towards modern human language.

The general ‘cultural explosion’ would have provided the cognitive tools for rapid idiolect evolution, as well as considerable selection pressure towards improved communication concerning more and more complex issues. Rapid growth of the size and complexity of idiolects would result, straining the abilities of learners and speakers, and placing selective pressure on our biological language equipment as well.

\textsuperscript{27}No relation to the author.

\textsuperscript{28}Note that nothing here is implied with respect to innate language acquisition devices à la Chomsky, an issue to which we’ll return in Section 9.7. That humans are adapted to language acquisition means only that human beings have (and have acquired through evolution) what it takes to acquire a language, whatever that may be.
3.6 The time scale of evolution

Biological, classical Darwinian, evolution of human beings is by far the slowest of the various levels of evolution discussed in this chapter, with many generations required for notable changes to occur. Cultural evolution, including memetic language evolution, is orders of magnitude faster (Whitmeyer, 1998), not least because it is basically Lamarckian — variations can be directed, and acquired characters can be transmitted and inherited — but also for reasons having to do with cultural structures being less rigidly interlocked than biological structures (Wimsatt, 1999). The evolution-like processes during ontogeny are, of course, the fastest, as they need to be completed on a timescale of months or years.

The issue of time scales is highly relevant, as it has been argued by various people, from Chomsky (1990) to religious fundamentalists, that the time elapsed during human evolution is insufficient for the Darwinian gradual evolution of such a complex system as human language. Worden (1995) makes the same argument, in more quantitative terms, and with more real biology behind it than either Chomsky or the fundamentalists can muster. Worden’s argument is not completely watertight, but he does have a point in that the emergence of language would be remarkably rapid if it were a pure process of biological evolution gradually, step by step, producing a highly complex innate, genetically determined, language acquisition device. From this it can be concluded that such arguments “depend on inaccurate assumptions about biology or language or both.” (Pinker & Bloom, 1990, p 707) and that “the claim that linguistic principles are innate or genetically programmed is an oversimplification with little precise meaning.” (Mueller, 1996, p. 25, online edition).

The latest version of Chomskian grammar (Chomsky, 1995; Berwick, 1997; Lasnik, 2002) might appear to offer a way out, in that his ‘Minimalist program’ (MP) actually proposes a very simple core grammar, with basically just a single rule — as argued by Lorenzo & Longa (2003): “A biological interpretation of the MP points to the conclusion that a minimum of specifically linguistic genetic codification can suffice...” (p. 646). However, the implementation of the single rule of MP appears quite complex, and it is not at all obvious that it can be genetically encoded in any simple way as argued by Lorenzo & Longa (2003).

In any case, this complex interplay of evolution at different levels is not a unique feature of the evolution of language, but is a general, if often neglected, feature of the evolution of complex traits, particularly those with a behavioral component (Vancassel et al., 1996; Lipp, 1996; Laland et al., 2000). And even the biological part of the evolutionary process need not be straightforward, but may take a variety of routes (Thornhill & Ussery, 2000).

---

29 But see Wilkins (2001).
30 Chomsky’s rhetoric here is frequently quoted with some glee by these fundamentalists, e.g. Morris (2001), in supposed support of their own anti-evolutionary and anti-scientific agenda.
3.7 Summary

Evolution is a very general process, applicable far outside the purely biological realm for which the theory evolved. Natural selection is the core of evolution, a process in which those who are good at reproducing themselves increase in numbers at the expense of those who are not.

Key points to keep in mind:

- Any population of entities will evolve, given:
  - Variation
  - Differential replication
  - Heredity
- Common pitfalls:
  - Evolution does not plan ahead
  - Evolution will find local optima only
  - Evolution does not take into account ‘the good of the species’
  - Evolution does not repeat itself
  - The original and current functions of a feature need not be the same
  - Evolution is not a matter of ascending a ladder towards an ultimate goal
  - Evolution does not take place in a vacuum

The evolution of language is not just a matter of the biological evolution of us language-users — other evolving systems have to be taken into account:

- Biological evolution of other human faculties
- Evolution-like processes in ontogeny
- Evolution, cultural and biological, of human social systems
- Evolution of language per se

Coevolution of these different systems, particularly between human cognition and language itself, may well play a prominent role in the emergence of modern human language.

Further reading

CHAPTER 4

HUMAN ORIGINS AND EVOLUTION

An overview of our family tree is given here, as there will be frequent reason further on to refer back to various branching points in the phylogeny. This chapter will dwell mainly on our own line of descent through the fossil record, with scant attention paid to other branches — but the reader should keep in mind that our history is not a simple progression up a straight ladder (cf. page 19), ours is just a single twig in a thriving bush.

The origin of mammals is chosen as a starting point, because several relevant organs (notably brains and ears) underwent significant restructuring at that time. The history of life before that point is treated at length in any number of textbooks, e.g. Cowen (1995), and popular works, e.g. Fortey (1997), to which the reader is referred. Likewise, there is no shortage of books about human evolution, from Darwin (1871) onwards; see the Further Reading section at the end of this chapter. Most of them contain detailed accounts of recent fossils of human or near-human beings, together with large or small doses of speculation about our origin and evolution, a tradition that goes back at least as far as Huxley (1863) and Haeckel (1897).

There are a number of standard scenarios — or perhaps narratives is a better word (Landau, 1991) — that are repeated in discussions about human origins. The four main features distinguishing us from our relatives are terrestriality, bipedalism, encephalization, and a technological culture, and scenarios of human evolution attempt to account for the origin of these four in various different orders. Most possible permutations of these four have been proposed at one time or another, though the one of Darwin (1871) has probably been the most popular. In Darwin’s scenario, the order is terrestriality – bipedalism – encephalization – culture, which is also close enough to the current consensus in paleoanthropology (though reality is not as neatly divided into stages as the narratives are). Other narratives with different permutations can be found in e.g., Keith (1915), Wood Jones (1916), Osborn (1916), Elliot Smith (1924), and Gregory (1934). But all these scenarios have an underlying common theme, resembling heroic folktales in their structure (Landau, 1991).
The study of human evolution, touching as it does upon our own reasons for existing, is a highly emotional subject for many people. Tempers run high also among professional paleoanthropologists (Lewin, 1987), and a sizeable fraction of the general population even today refuse to recognize our kinship with the other primates, as reviewed in Scott (1997). The issues in human evolution thus appear more contentious in the literature than is really motivated by the magnitude of the actual scientific disagreements, which are generally minor.

One source of unnecessary confusion comes from the naming of fossil species. Living species are moderately well-defined, but the partitioning of a set of fossils into different species involves a degree of intrinsic arbitrariness. Particularly in the case of so-called chronospecies, where a single lineage changes through time, the divisions are purely a matter of taste. Different paleoanthropologists use different names for the same fossils, and the same names for different fossils; do not mistake naming conflicts for genuine disagreements about the course of human evolution. For an example of a pure naming conflict, see e.g., the exchange between Kennedy (1999) and Wood (1999), and the comments of Groves (1999). I shall endeavor to be consistent in my own use, and to note where confusion may arise.

The fossil record of all life forms is necessarily incomplete. In order for us to know of an ancient species, a dead creature first has to be buried before decomposing, then remaining intact in the sediments over millions of years of geological processes, then being re-exposed on the surface, and finally found by a paleontologist before being destroyed by erosion. The odds against this happening are very long indeed, and so we are most likely aware of only a tiny fraction of all the species that have ever lived. There may have been thousands of primate species alone that we have seen no trace of (Fleagle, 2002). So gaps in the family tree will be the rule rather than the exception. Nevertheless, we do have enough fossils to discern some patterns, notably in the case of recent human ancestors, on whom a disproportionate amount of paleontological effort is spent.

4.1 Mammals

Mammals arose during the Triassic,\(^1\) some 200-odd\(^2\) million years ago, from a group of reptiles, in parallel with the evolution of dinosaurs from another group of early reptiles. The transition from reptiles to mammals is a smooth sequence, amply documented in the fossil record, with numerous intermediate forms, so closely similar anatomically that it is difficult to select one of them and say ‘Here is the first mammal.’ One of the important skeletal differences between reptiles and mammals

---

\(^1\)The various geological periods that are mentioned in the text are shown in Figure 4.1.
\(^2\)The exact age is undetermined, and largely a matter of definition at that (Rowe, 1988). Jenkins et al. (1997) argue for a somewhat older (Mid-Triassic) age than most authors, based on recent fossil finds indicating considerable diversity already 200 million years ago.
Figure 4.1. Geological periods. Left: The past 225 million years. Right: The past 65 million years. (m.y.a. = millions of years ago)

lies in the middle ear, where reptiles have just a single bone, and mammals have three, significantly improving the frequency range and sensitivity of their ears; see Section 5.2.

A detailed review of the reptile-mammal transition, with references to the primary paleontological literature, can be found in Rowe (1988) (but see also Wible (1991) for some critical perspectives on Rowe (1988)). More accessible, and possibly more complete, is the fossil review of Hunt (1997).

The early Mesozoic mammals were all small, and most likely nocturnal, superficially resembling modern-day shrews but belonging to groups that are now extinct. The earliest placentals\(^3\) are 125 million years old (Stokstad, 2002), but

\(^3\)The placentals (or more formally Eutheria) are mammals equipped with a placenta, and thus capable of nourishing their young inside the body for an extended period, unlike the marsupials and the egg-laying monotremes.
their diversification into the familiar orders of today’s mammals, from perissodactyls to primates, did not take place until around the time of the mass extinction at the Cretaceous-Tertiary (K/T) boundary \(^4\) 65 million years ago. The exact timing of events is unresolved, and somewhat contentious (see e.g., Benton (1999) versus Easteal (1999)), but shortly after the dinosaurs were gone, the mammals (as well as the birds) had undergone a major adaptive radiation. All modern orders, including our own, can trace their origin to this period. The rapidity of the diversification is unfortunate, since it makes the relationships between major groups of mammals difficult to resolve, though some progress has been made recently (Murphy et al., 2001).

### 4.2 Primates

Our own order of mammals, the primates, have been around for at least 60 million years, and possibly even longer than that. According to some authorities, primates are members of a larger grouping, Archonta, which also includes Chiroptera (bats), Dermoptera, \(^5\) and Scandentia (tree shrews), whereas others leave out the bats, placing primates, dermopterans, and tree shrews together in Euarchonta (Helgen, 2003; Scally et al., 2001), possibly grouped with rodents and lagomorphs in Euarchontoglires (de Jong et al., 2003), but the evidence remains inconclusive (Allard et al., 1996; Miyamoto, 1996; Heesy, 2001; Springer & de Jong, 2001; Liu et al., 2001; Sargis, 2002a; Novacek, 2001; Silcox, 2002). There is some paleobiogeographical evidence indicating a Northern Hemisphere origin for primates (Helgen, 2003).

Unambiguous primate ancestors are recognizable from the early Paleocene onwards (Bloch & Boyer, 2002; Sargis, 2002b; Kirk et al., 2003), with a sometimes extremely detailed fossil record (Gingerich, 1980; Gingerich, 1983), but molecular studies (Easteal, 1999) as well as indirect paleontological estimates (Tavaré et al., 2002) indicate a somewhat earlier Cretaceous origin. The earliest primate fossils were most likely nocturnal (Kay & Kirk, 2000), like most modern lemurs, and were generally lemur-like. Some recent finds from this earliest period are reviewed in Ringe (2002), and a nice overview of both fossil and extant primate groups can be found in Shoshani et al. (1996).

---

\(^4\) Mostly famous for the demise of the dinosaurs, even though they formed only a small minority of all the animals that went extinct at the same time. For more on the K/T extinction, see e.g., Burke et al. (1994). Benton (1995) reviews the general history of mass extinctions.

\(^5\) The order Dermoptera lacks a generally accepted English name. They are sometimes called ‘flying lemurs’, but that appellation is reasonable only under the Archonta hypothesis. They look like large flying squirrels, a lemur-like body with fore- and hindlimbs connected by extended skin flaps, enabling them to glide from tree to tree. Apart from the flight-skin, they resemble the lemurs enough that there is ongoing debate about whether certain early fossils are primitive dermopterans or primitive primates (Bloch & Silcox, 2001).
The extant primates can be divided into two major groups: the *Strepsirhini* (lemurs and their allies; see Roos et al. (2004)) and *Anthropoidea* (monkeys, apes, and us). There is also a third, minor, group, the tarsiers, which are small nocturnal animals from Southeast Asia.Traditionally, they have been placed with the lemurs, but most recent authors classify them with the *Anthropoidea* in *Haplorhini* (Goodman et al., 1998; Kay et al., 1997; Ross, 2000; Poux & Douzyery, 2004), a classification that has received some recent fossil support (Ducrocq, 2001; Ni et al., 2004). The general shape of the primate family tree is sketched in Figure 4.2.

The split between the strepsirhines and anthropoids appears to have occurred quite early on, around 50–55 million years ago, or possibly even earlier (Goodman et al., 1998). At that time, there were two prominent fossil groups, the adapids and the omomyids. The adapids (Godinot, 1998) have lemur affinities, and the omomyids (Rasmussen, 1996) are likely ancestral to tarsiers, and may be ancestral to us as well (Ross et al., 1998; Ni et al., 2004; Martin, 2004), though some authors, like Shoshani et al. (1996), are not convinced. For more on strepsirhine origins and evolution, see e.g., Rasmussen & Nekaris (1998).

Our own suborder, the anthropoids, which includes monkeys as well as us and the other apes, is characterized by a larger brain than most other mammals of similar size (see Figure 5.2 on page 90), by adaptations to a diurnal, largely arboreal lifestyle, as evidenced by binocular color vision with a fovea6 (Kay & Kirk, 2000; Ross, 2000; Barton, 2004), and by grasping feet, obviously adapted for climbing rather than for the leaping common among strepsirhines and tarsiers. The origin and evolution of anthropoids is reviewed in Ross (2000).

The earliest anthropoids are found in Eocene deposits, around 40 million years old. *Eosimias* may possibly be the earliest monkey fossil (Kay et al., 1997; Ducrocq, 2001), but the evidence is not conclusive (Gunnell & Miller, 2001); the early primates are not all that different from each other, and most of the fossils are fragmentary, making it hard to tell who is the ancestor of whom. Furthermore, there is a general shortage of mammalian fossils from the relevant time frame (Hunt, 1997), with gaps in many lineages, aggravating the problem of tracing the early lines of anthropoid descent. The split between the *Catarrhini* (Old World monkeys and apes) and *Platyrrhini* (New World monkeys) also occurred in this time frame. The platyrrhines do not further concern us here; for more on their origin and phylogeny, see e.g., Houle (1999) or Bauer & Schreiber (1997). For an overview of catarrhine systematics and anatomy, see http://mac-huwis.lut.ac.uk/~wis/lectures/primate-adaptation/07OwMsAndHominids.pdf

After the split, the anthropoids do have a fairly clear fossil record. *Aegyptopithecus*, living in Egypt some 31 million years ago, is a good candidate for the most recent common ancestor of all extant catarrhines (Goodman et al., 1998; Kay

---

6Our color vision is very likely an adaptation to frugivory (Wolf, 2002). The fovea is a spot in the eye with a denser concentration of photoreceptors, improving visual acuity in daylight, at the expense of some night vision.
et al., 1997; Shoshani et al., 1996; Benefit, 1999), though some authors propose a different, more gibbon-like (so far hypothetical) ancestor (Köhler & Moya-Sola, 1997b). An Old World monkey, clearly distinct from apes, first turns up with *Victoriapithecus*, 15 million years ago (Benefit & McCrossin, 1997; Köhler & Moya-Sola, 1997b; Benefit, 1999), though Pilbeam & Young (2004) cite monkey finds from 19 million years ago.

The apes, or hominoids, more formally known as superfamily *Homoidea*, are first represented by a diverse group of Miocene fossils, of which *Proconsul* is the best known representative (Rae, 1999). *Morotopithecus* has been proposed as the first hominoid, just above 20 million years old (Gebo et al., 1997; Young & MacLatchy, 2004; MacLatchy, 2004), but the fossils assigned to it are, according to Gommery et al. (2002), a mixture of bones from two species, *Afropithecus* and *Ugandapithecus*. Apes appear to have been significantly more widespread and common during the Miocene than they are today with somewhere around thirty known species then (Gibbons & Culotta, 1997) against less than ten surviving species today (but see Jablonski & Kelley (1997) for a dissenting view). Overviews of the fossils can be found in e.g. Pilbeam (1996) or Shoshani et al. (1996), with a historical perspective in Mann & Weiss (1996), and the relationships between the extant apes, together with a few key fossils, are shown in Figure 4.3.
An interesting early ape, which appears to have gone partway along the human route towards bipedality and hand use, is Oreopithecus. The morphology of its feet and pelvis is remarkably similar to that of early hominids (Köhler & Moya-Sola, 1997a; Rook et al., 1999), and its hands may possibly have been capable of the kind of precision grip (‘opposable thumb’) that is often regarded as typically human (Moya-Sola et al., 1999, but see also Susman (2004)). Oreopithecus lived on a Mediterranean island in what is now Italy, and may have evolved bipedality as an adaptation to insular conditions, with an absence of large predators (cf. the flightless birds formerly common on oceanic islands). It appears unlikely that we are descended from Oreopithecus, but it may nevertheless cast light on the evolution of our own anatomical adaptations.

Otherwise, the Miocene scene is complex, and several different candidates have been proposed as ancestors of the living apes and humans (Gibbons & Culotta, 1997; Pilbeam, 1996; Ward et al., 1999a; Gibbons, 1998c). Gibbons split off at an early stage, with Laccopithecus a possible ancestor (Fuentes, 2000), and there is fairly solid evidence of a separate lineage leading to orangutans dating back to at least 15 million years ago (Pilbeam, 1996; Chaimanee et al., 2003; Chaimanee et al., 2004). The orangutan lineage most likely includes the genera Sivapithecus, Ramapithecus, Lufengpithecus, Khoratpithecus, Gigantopithecus, and possibly also Ankarapithecus and Griphopithecus. Sivapithecus and Ramapithecus are sometimes conflated into a single genus. Ramapithecus was once, in the 1930s–1950s, erroneously interpreted as a human ancestor, on the basis of fragmentary
fossils. When a complete face was subsequently found, the orangutan affinities became clear. *Gigantopithecus* survived into fairly recent times, co-existing with early humans (Ciochon et al., 1996), and makes a plausible ‘yeti’ or ‘bigfoot’, in the unlikely case that such creatures aren’t completely fictional.

### 4.2.1 Who is our closest living relative?

With the orangutan on a separate branch in the family tree, we are left with ourselves, the gorilla, and the chimpanzee, shown together in Figure 4.4. According to the vast majority of recent work on the subject (Lewin, 1993; Shoshani et al., 1996; Mann & Weiss, 1996; Pilbeam, 1996; Goodman et al., 1998, among many others), the three of us form a monophyletic group of African apes, with a common ancestor somewhere around ten million years ago. It is unclear whether this branch had its origin in Africa or in Eurasia (Cote, 2004). *Ouranopithecus* (Richmond, 1999), *Kenyapithecus* (Gibbons & Culotta, 1997), *Otavipithecus* (Senut & Pickford, 2004), and *Dryopithecus* (Köhler & Moya-Sola, 1998; Kordos & Begun, 2002) are candidate fossils from the lineage leading to that common ancestor.

The closest relatives within this African group are the two chimpanzee species *Pan troglodytes* (common chimp) and *Pan paniscus* (pygmy chimp, or bonobo). Beyond that, the relationship between chimps, gorillas, and humans is very nearly an unresolved trichotomy; the family tree split into three branches at the same time, or nearly so (Deinard & Kidd, 1999). Considerable efforts have gone into determining whether the chimpanzee or the gorilla is our closest relative.

Traditionally, the similar gross morphology and habitat of the chimp and gorilla, notably their unique shared way of walking, supporting their weight on the knuckles of their hands (Dainton & Macho, 1999), have been considered sufficient grounds for grouping them together, to the exclusion of humans (Pilbeam, 1996; Mann & Weiss, 1996; Lewin, 1993). More recent morphological analyses, however, taking into account a larger number of more informative characters, may reach different conclusions; Shoshani et al. (1996) and Lockwood et al. (2004), for example, present morphological support, for a *Pan-Homo* clade. With soft-tissue characters included as well, Gibbs et al. (2000) find that they can firmly exclude a *Gorilla-Pan* connection. Also, the oldest fossils from the human line do show some similarities with chimpanzees (White et al., 1994; Wood, 1994). Furthermore, the hands of the australopithecines (early human ancestors; see Sec-

---

7See, however, Milinkovitch et al. (2004) for a genetic analysis of purported yeti remains, the results of which are not consistent with orangutan affinities (but note also the publication date of their article).

8Schwartz (1984; 1987) is, as far as I can determine, quite alone in considering the orangutan our closest relative; see also Shoshani et al. (1996, table 1).

9In the following, there will be frequent references to chimps of both species. When referring to identified individuals, ‘chimpanzee’ (or chimp) means *Pan troglodytes*, but ‘chimpanzee’ will also be used as a collective label for both species, when referring to their branch of the family tree in general — in a phrase like ‘the language abilities of chimps’, bonobos are implicitly included.

10A clade is a group consisting of a common ancestor, and all that ancestor’s descendants.
tion 4.3.1 below) display traces of having evolved from knuckle walking ancestors (Richmond & Strait, 2000; Richmond & Strait, 2001; Corruccini & McHenry, 2001; Richmond et al., 2001), making knuckle walking ineligible as evidence for a chimp-gorilla relationship. Hints of knuckle-walking also turn up early in the fossil record, with *Kenyapithecus* (Benefit, 1999), well before the likely time frame of the human-chimp-gorilla split. Pilbeam (1996) gives a review of the situation.

Evidence from shared parasites also points towards a human-chimpanzee relationship — both our intestinal worms (Brooks & McLennan, 2003) and our lice (Kittler et al., 2003; Retana Salazar, 1996) are each other’s closest relatives. The only known exception is the genital lice, ‘crabs’, that infest the pubic hair of humans — their closest relatives inhabit gorilla fur; it would be improper to speculate on the origins of that relation. I am not aware of any parasites uniting chimps and gorillas to the exclusion of humans.

From the 1960s onwards, molecular data have been increasingly used for resolving phylogenetic issues (Pauling & Zuckerkandl, 1963; Zuckerkandl & Pauling, 1965, cited in Hey (2001)). A wide variety of molecular tests have been used: immunological, protein sequence, chromosome pattern (Williams, 1999), gene expression (Uddin et al., 2004), and direct DNA (Gibbons, 1998d; Kaessmann et al., 1999) comparisons. Molecular comparisons among the hominoids have been published by the dozen, with a strong majority supporting the grouping of chimps with humans, to the exclusion of gorillas (Pilbeam, 1996; Arnason et al., 1996b; Pilbeam & Young, 2004). Shared DNA-element insertions are particularly compelling evidence joining humans and chimpanzees (Salem et al., 2003). There are, however, dissenters such as Deinard et al. (1998) who still prefer a *Pan-Gorilla* grouping, and there is some molecular support for that position (Barbulescu et al., 2001), but the bulk of molecular studies still points in the opposite direction. The current situation is summarized by Pilbeam & Young (2004).

We can now conclude that the preponderance of the evidence points towards the *chimpanzee* being our closest living relative:

...demonstrated once again the close relationship of humans and chimpanzees, with gorillas more distant.

Inasmuch as any issue in such a contentious area as paleoprimatology can be considered settled, this is settled. (Pilbeam & Young, 2004, p. 308)

The timing of the split between us and the chimpanzees is not strongly constrained by either molecular or paleontological data. The current best estimate is around five million years (Pääbo, 2003), though both older and younger dates have their supporters. Arnason and associates (1996a; 1996b; 1998), argue for a somewhat older split, around 10 to 15 million years ago, whereas Takahata & Satta (1997) and Easteal & Herbert (1997) prefer less than five million years. The latter, however, is becoming difficult to reconcile with some of the fossils discussed in the next section.
Figure 4.4. Skulls of extant African apes, intended for comparison with the various fossil skulls shown in the next section. Above: Chimpanzee (left) and Gorilla (right). Below: Human.
4.3 Hominids

Our own exclusive branch of the family tree, traditionally known as the family Hominidae, separated from the branch leading to the chimpanzees a bit more than five million years ago. During those five million years, we have evolved a number of distinct traits. Cognitive traits will be discussed in later chapters, but here is a list of the main bodily changes (adapted from Carroll (2003), with some changes):

- Size of brain
- Shape of skull, face flattened and tucked in beneath the braincase
- Skull balanced on top of vertical spine
- Shape of torso and thorax
- Longer legs and shorter arms
- Small canine teeth
- Shorter jaws and weaker jaw muscles
- Shorter fingers, but longer thumb
- Pelvis twisted to fit bipedal posture
- Modified hair cover — less on the body, more on the head, sexually dimorphic
- Chin
- S-shaped spine

Not all of these changes appeared at once — instead, they appear one by one, spread out across the entire history of us hominids. And in some cases, like the sexually dimorphic hair cover, we simply have no idea when it evolved, since hair is rarely preserved on fossils.\(^{12}\)

Unfortunately, relevant fossils have been sparse from the period around the branching point between us and chimps, and we still do not have any fossils that

\(^{11}\)Yet another example of a rather uninteresting naming conflict is the perennial disagreement about which Linnaean rank we should be accorded. Do we have just our own genus Homo, or our own subtribe Hominina, or our own tribe Hominini, or our own subfamily Homininae, or are we a full family? Traditionally, we have been placed in our own family (and the great apes in a family of their own as well, the Pongidae), but as evidence of the close relationship between us and chimps has accumulated, it has become more and more clear that it’s not really appropriate to place us in separate families. If we attempt to look at the issue without anthropocentrism, it may well be biologically appropriate to place us and the chimps in the same genus (Wildman et al., 2003; Gross, 2003), so that we would be Homo sapiens, Homo paniscus and Homo troglodytes.

I have nevertheless chosen to follow tradition in this work, and use the word ‘hominid’ to refer to the human branch of the family tree only, after we and the chimps went our separate ways — cf. Wolpoff (2003, footnote 1, p 666). Readers should be aware that some authors use ‘hominin’ in the same sense as I use ‘hominid’. For those authors, ‘hominid’ may include other apes.

Linnaeus (1758) actually included orangutans in the genus Homo, as Homo nocturnus (Cela-Conde & Ayala, 2003), but that is more likely due to his near-total lack of knowledge about apes, than to any 18th-century understanding of how close we are.

For a critical review of the Linnaean classification system as applied to primates, see Groves (2004).

\(^{12}\)Our general loss of body hair is dated by Rogers et al. (2004) to 0.56–1.2 million years ago, based on the evolution of a skin-color gene, but alternative interpretations of their data cannot be excluded. Neufeld & Conroy (2004) discuss the origins of human head hair, but without reaching any clear resolution.
clearly belong to our last common ancestor, though some possible candidates have been forthcoming recently (Gibbons, 2002d). One is *Sahelanthropus tchadensis*, a skull found recently in Chad (Brunet et al., 2002; Wood, 2002; Whitfield, 2002), about 6-7 million years old (Vignaud et al., 2002), with a puzzling mixture of features making it difficult to classify. The discoverers, Brunet et al (2002; 2004), emphasize its human-like features, as do e.g., Pilbeam & Young (2004), whereas others see more similarities with gorillas (Chalmers, 2002; Gibbons, 2002b; Wolpoff et al., 2002).

Another possible common ancestor is the set of teeth (Ungar et al., 1994) and bones (Senut et al., 2001; Pickford & Senut, 2001; Holden, 2000; Balter, 2001c), just below 6 million years old (Sawada et al., 2002), found at Lukeino, Kenya, and assigned to *Orrorin tugenensis*. There are some indications that *Orrorin* was bipedal (Pickford et al., 2002; Galik et al., 2004), which would actually place it on the human side of the branching point, but this remains controversial (Gibbons, 2004). Other small fossil fragments that may belong near the branching point in the family tree have been found in several places in Kenya, but they are too fragmentary for firm conclusions (Schick & Toth, 1993). *Ouranopithecus*, who lived in Europe around 9-10 million years ago, has also been proposed as a human ancestor (Koufos & de Bonis, 2004; de Bonis & Koufos, 2004), but belongs more likely on the trunk of the family tree before the last common ancestor of us and the other African apes (Richmond, 1999) — the morphology is unclear, and the age is too old, in conflict with most molecular results.

But even without a clearcut ancestor fossil, if we look at what traits, both physical and behavioral, we and chimps and gorillas and earlier apes have in common, we can make some plausible inferences about the probable characteristics of this ancestor. Basically, we can expect it to be rather similar to a chimpanzee, an inference that is borne out by the earliest human fossils (Lewin, 1993; Tappen & Wrangham, 2000; Richmond et al., 2001; Pilbeam & Young, 2004).

On the branch leading towards humanity, after that common ancestor, we do have a fair number of fossils, thousands of individual finds, classified into a large number of species. Many different names will be introduced in the following sections; for ease of reference, they are listed in Tables 4.1a and 4.1b on pp. 61–62. Figure 4.5 shows when they lived, and a rough sketch of a possible family tree can be seen in Figure 4.6.

4.3.1 Bipedal apes

The very early history of the human lineage is an exclusively African affair, which is not very surprising (it was foreseen already by Darwin), considering that our nearest relatives still live there.

13But note that the authors of Wolpoff et al. (2002) include the discoverers of *Orrorin*, discussed in the next paragraph.
The oldest known fossils that with some confidence can be assigned to the human line are those in the genus *Ardipithecus*, from 4 - 6 million years ago, found in Ethiopia. The 4-million-year-old ones are assigned to the species *Ardipithecus ramidus* \(^{14}\) (White et al., 1994; Wood, 1994), and the earlier ones to *A kadabba* \(^{15}\) (Haile-Selassie et al., 2004).

The discoverers of *Ardipithecus* and of the above-mentioned *Sahelanthropus* and *Orrorin* all consider their own fossil to be a human ancestor, and the others to be side branches (Cela-Conde & Ayala, 2003), though Haile-Selassie et al. (2004) is more moderate than previous publications. From the discoverers of *Orrorin*:

*Orrorin* is a 6-million-year-old biped, *Ardipithecus* could be related to chimpanzees and *Sahelanthropus* is possibly related to gorillas. (Senut & Pickford, 2004, p. 268)

The case for *Ardipithecus* appears to be somewhat more compelling. Regardless of which, if any, of these claims is correct, these finds make it clear that our earliest history is complex, with many branches on the family tree (Begun, 2004).

Some aspects of *Ardipithecus’* anatomy, particularly the teeth, are very nearly halfway between chimps and humans, but some details in the arms and skull base indicate that it was probably bipedal. This is a general theme in early human evolution; ape-like features in the head and teeth (including an ape-sized brain; see Section 5.3), but a human-like postcranial (below the neck) skeleton, adapted for bipedal walking. Obviously, the legs evolved first, probably right after we split from the apes, and the quintessential human quality that we call intelligence didn’t evolve until a few million years later (McHenry, 1994).

But the transition to pure bipedalism wasn’t instantaneous; apelike body proportions, with long arms and short legs, persist in some hominid fossils for another two million years, as do various climbing adaptations (Shreeve, 1996; McHenry & Berger, 1998; Oliwenstein, 1995; Clarke & Tobias, 1995), and apelike balance organs (Spoor et al., 1994). The early hominids, of the genera *Ardipithecus* and *Australopithecus*, did walk on two legs, and their gait (Lincoln, 1998) and footprints \(^{16}\) look very much like ours. Likewise, the load-bearing patterns in their hipbones (Macchiarelli et al., 1999; MacLatchy & Bossert, 1996; MacLatchy, 1996) and spines (Sanders, 1998) are consistent with bipedalism, if not identical to ours. Early australopithecines have pelves with a mixture of human and ape features, but later ones approach the human morphology (Haeusler, 2002).

---

\(^{14}\)Formerly assigned to *Australopithecus*.

\(^{15}\)The *kadabba* fossils were originally granted only subspecies status as *A ramidus kadabba* (Haile-Selassie, 2001), but were later elevated to a species of their own (Haile-Selassie et al., 2004), based on new distinctive fossil material.

\(^{16}\)A fossil trail, 3.7 million years old, with very humanlike footprints from at least two individuals, was found at Laetoli, Tanzania (Leakey & Hay, 1979). The footprints themselves do not reveal who made them, but the only fossils from the same time of beings that might possibly have walked bipedally are australopithecine.

Footprints from hominids are quite rare as fossils. The next set is ten times younger, from Italy some 300,000–400,000 years ago (Mietto et al., 2003), at which time our ancestors had bodies almost exactly like ours.
The evidence hints at an early australopithecine lifestyle mixing bipedal walking on the ground, with some modest amount of tree climbing. By three million years ago, the evidence indicates a diet obtained from more open terrain, implying less climbing. The diet may possibly have included meat from grass-eaters (Sponheimer & Lee-Thorp, 1999; Vogel, 1999b), but Ungar (2004) shows that australopithecine teeth were better adapted for hard, brittle food — teeth good at chewing tough elastic food like meat came along with early *Homo* (see Section 4.3.2 below).

The evolutionary reasons for the transition to bipedalism are not entirely clear. Many possible scenarios have been proposed, as reviewed in Richmond et al. (2001). Locomotor efficiency is commonly cited as a major factor — chimps move around almost as much in the trees as on the ground, so chimpanzee anatomy is a compromise between climbing ability and ground-walking ability, and not well optimized for either purpose (Pontzer & Wrangham, 2004). Our early ancestors were likely similar, which means that they already spent a fair amount of time in an upright position in the trees, and that adaptations for quadrupedal walking were not deeply integrated in their anatomy, facilitating the move to bipedalism (McHenry, 2004). Human bipedal walking efficiency (in terms of calories per kilometer) is comparable to an average quadruped of the same size (Steudel-Numbers, 2001), but chimpanzee ground-walking efficiency is significantly poorer (Leonard & Robertson, 2001; Lewin, 1993, but see also Steudel-Numbers (2003)). If early australopithecines adopted a lifestyle requiring longer walking distances, as a response to a drier climate with sparser resources, or possibly as a response to larger group size (Isbell & Young, 1996, and Section 10.5 below), there would be considerable selection pressure towards improving ground-walking efficiency, which may explain the origins of bipedalism. If combined with selection for hand use, this may have lead to bipedalism rather than to improved quadrupedal knuckle-walking. The climate record of East Africa over the past 4 million years does show aridification episodes, but not quite at the right time — the climate got drier around 3.5 million years ago, but our ancestors had most likely already been bipedal for a while then (Wynn, 2004; Cane & Molnar, 2001).

Chimps (both common chimps and bonobos) can walk bipedally if they want to (Videan & McGrew, 2001), though with poor energy efficiency (Wang et al., 2003; D’Août et al., 2002), and with some apparent discomfort. They often do so when they have something to carry, or have injured hands, or when picking fruit in trees (Stanford, 2002a), so the step towards habitual bipedality is not large, and may have needed only a small selective push. In controlled laboratory settings, incentives to carry food, and opportunities to pick elevated food, were both found to increase the incidence of bipedal walking in chimps (Videan & McGrew, 2002), so an environmental change providing similar incentives to our ancestors may have been enough to catalyse the change into habitual bipedalism. But chimpanzee bipedal walking is different from the human gait in several respects, partly due to
Figure 4.5.  
Overview of the approximate times at which the various hominid species existed.
Figure 4.6. One possible family tree for the hominids. Few of the relationships are firmly established — the tree shows what I consider the best-supported hypothesis, but many alternative possibilities exist.

morphological constraints in the chimpanzee (Günther et al., 1998; D’Août et al., 2002) (though monkeys trained from an early age to walk bipedally can achieve a more human-like gait (Hirasaki et al., 2004)), so some redesign is likely to have taken place early in the hominid lineage.
Just beyond 4 million years of age is the next fossil species, *Australopithecus anamensis* (Leakey & Walker, 1997; Leakey et al., 1998; Ward et al., 1999b; Ward et al., 2001). It is in many ways similar to *Ardipithecus ramidus* from which it may be descended (or maybe just a close cousin), but slightly more humanlike, particularly in the teeth. The changes in the teeth, with thicker enamel, larger molars but smaller anterior teeth, and different wear patterns, can be interpreted as adaptations for eating harder and more abrasive food than the soft fruits eaten by chimpanzees. *A anamensis* was definitely bipedal, with sufficient finds from the legs to demonstrate directly their mode of locomotion, unlike the indirect indications found in *Ardipithecus* and *Orrorin*.

After *anamensis* comes what may be most famous individual fossil of them all: ‘Lucy’ (Johanson & Taieb, 1976; Johanson & Edey, 1981), more formally known as *Australopithecus* *afarensis*. Lucy was found as one fairly complete female skeleton; many other fossils from the same species have subsequently been found. It is closely related to, but more humanlike than, *A anamensis*. The oldest fossils of *A afarensis* are just a few thousand years younger than *A anamensis* (Kappelman et al., 1996). Anatomically, she follows the same pattern as the two previous species, with a lower body adapted for bipedality, but with an apelike skull. Her brain is just marginally larger than that of a chimpanzee.

Lucy’s species appears to have been fairly widespread in East Africa from four to three million years ago. The other hominid species mentioned so far also come from the same region, so it would appear that East Africa is the cradle of humanity. There may, however, be other explanations for the shortage of fossil finds elsewhere:

- Early success in East Africa prompted paleoanthropologists to concentrate their efforts there.
- Fossil-bearing strata of the right age may not exist, or may not be accessible elsewhere in Africa.
- The environments elsewhere may not have been as conducive to fossilization as those of East Africa.

17 A few people (Senut, 1996; Strait et al., 1997; Wood & Collard, 1999) prefer to call her *Praeanthropus africanus*. For three reasons, I do not think such a change is well motivated (see also Groves (1999)):

- *Australopithecus afarensis* is in common and popular use
- Naming her anything with *africanus* invites confusion with *Australopithecus africanus*
- The difference between Lucy and the other australopithecines is nowhere near large enough to motivate a genus-level distinction.

Cela-Conde & Ayala (2003) also prefer *Praeanthropus*, but propose to extend the genus to encompass also a number of other australopithecines, as well as *Orrorin*.

18 Her gender has been questioned (Häusler & Schmid, 1997), but female remains the most likely interpretation (Wood & Quinney, 1996; Bouhallier et al., 2004).

19 How and under what circumstances fossils become fossils is an entire science of its own, named taphonomy (Gastaldo et al., 1996). One of its conclusions is that the likelihood of an individual be-
An expedition a few years ago to Chad, 2500 km west of where Lucy was found, did nevertheless surprise everybody by finding an australopithecine fossil (Brunet et al., 1995). It was named *Australopithecus bahrelghazali* (Brunet et al., 1996), but the finds made so far are too fragmentary to draw any firm conclusions about its evolutionary significance (Leakey & Walker, 1997). It does, however, show that the history may be more complicated than just a straight-line evolution from *ramidus* through *anamensis* to *afarensis*, and onwards, a complexity confirmed by another contemporary fossil species, *Kenyanthropus platyops*, made public recently (Leakey et al., 2001; Lieberman, 2001a; Balter, 2001b). There may well be many more branches on the family tree than previously thought, also at this early date.

The situation after Lucy, around 2–3 million years ago, is certainly very complex, with several coexisting species (Tattersall, 2000). First on the scene is *Australopithecus africanus*, a gracile hominid basically similar to Lucy, and a plausible descendant of hers, with a few more humanlike traits. This was the first australopithecine to be found, in the 1920s in South Africa (Dart, 1925), a find first greeted with considerable skepticism (Tattersall, 1995), as it fit poorly with then-current theories of human evolution. As the human brain is our most important adaptation, it was considered most logical that the large brain would have evolved first, with other human traits, like bipedality and reduced teeth and jaws, coming later. The australopithecines demonstrate the opposite path, with their basically humanlike bodies and ape-sized brains (Lewin, 1993).

But as more bits and pieces of australopithecines were collected, particularly after the Piltdown forgery was debunked, the debate was settled. South Africa remained the main home of *Africanus* — the first semi-complete skeleton of which was found there recently (Clarke, 1998) — as well as of a different, more robust species, *Australopithecus robustus*.

Today, three different species of robust australopithecines are known: *A robustus*, *A boisei* and *A aethiopicus*; see Figure 4.7. The relationships between the three are unclear; they are similar enough that many researchers regard them as a clade ( Strait et al., 1997), and some, e.g. Susman (1998), even place them in their own genus, *Paranthropus*. Others, however, argue that the similarities are just due to parallel evolution of adaptations to a similar lifestyle that involves heavy chewing (Skelton & McHenry, 1998; Morell, 1999; McCollum, 1999). A coming a fossil is strongly dependent on the environment. In some environments, like rain forests or mountains, fossilization is exceedingly rare. Any jungle or mountain dwelling hominids are thus highly unlikely to be found. This may explain the dearth of chimp and gorilla fossils.

---

20This fossil has recently been re-dated to about 4 million years old (Partridge et al., 2003). There are, however, considerable doubts about the reliability of the dating method used (Gibbons, 2003b), and other dating methods give younger dates (Berger et al., 2002). Such an old date is also difficult to reconcile with other fossil finds that are more securely dated.

21It can be a bit confusing when the names *Paranthropus* and *Australopithecus* are mixed, like in Delsont’s (1997) review of Suwa et al. (1997). Even worse, *A boisei* was initially placed in yet another separate genus, *Zinjanthropus* (Leakey, 1959).
Figure 4.7. Australopithecus boisei, ‘Zinj’, a robust australopithecine. Note the thick arches and ridges on the skull.

recently discovered skull from Ethiopia looks like a plausible transitional form between aethiopicus and boisei, being intermediate in both time and morphology, and found in the right geographical region (Alemseged et al., 2002). But robustus remains disconnected from the other two.

Whatever their relationships, the robusts were quite successful, spreading over most of Africa and persisting for 1,500,000 years (Kullmer et al., 1999). They thus coexisted for quite some time with early Homo, before going extinct, possibly out-competed by their brainier cousins, or even hunted by them. It is sometimes argued that the robusts were too narrowly specialized in their diet, and unable to cope with ecological changes, but Wood & Strait (2004) find no support for this position.

There are also some disagreements concerning the relationships between the robusts and other hominids (Wood & Collard, 1999). Humans, in the stricter sense, have traditionally been regarded as descendants of gracile australopithecines, with the robusts as just a side branch. There is, however, some recent support for associating some or all of the robusts with the main human lineage, to the exclusion of A africanus (Strait et al., 1997; Skelton & McHenry, 1998; Braga, 1998). Apart from the details in the skull and teeth considered by the preceding authors, Susman (1994; 1998) found, interestingly enough, that A robustus possesses anatomical structures in the hand that are regarded as human adaptations for tool use (Aiello, Allen (1988) speculates on the possibility of their not going extinct, in his novel Orphan of Creation.

22 The graciles include A afarensis, africanus and garhi, which lack the heavy-duty jaws and teeth of the robusts.
1994), and bone tools have been found in association with robustus remains, tools that were apparently used for digging into termite nests (Backwell & d’Errico, 2001; Shipman, 2001). Among the gracile australopithecines, a africanus is a possible ancestor of later humans, but not all the details fit. A direct link from a afarensis to Homo, bypassing a africanus, has been a fairly popular view, but there is a significant gap in both time and morphology to be bridged in that case. A new fossil species, australopithecus garhi, discovered fairly recently (Asfaw et al., 1999), has attracted a lot of attention, since it fits quite neatly into that gap, in both time and space and morphology. But the evidence is still insufficient for any firm conclusions — the ancestor of Homo is almost certainly found among australopithecines, but we cannot tell which one without more fossil evidence.

In conclusion, the australopithecines were the main proto-human group from around 5 to 2 million years ago, with a rich diversity of species. The australopithecines had roughly human-like bodies, walking upright, but had ape-like skulls and brains. We have little indication of their having any culture beyond that of chimpanzees, and no reason to believe they possessed language.

4.3.2 Growing brains

Contemporary with the later robust australopithecines, and slightly younger than australopithecus garhi, are the first fossils classified within our own genus, Homo. These early Homo fossils were originally believed to represent a single species, and named Homo habilis (Leakey et al., 1964). As more fossils were discovered, this simple picture became untenable, and the diverse assemblage of semi-human fossils from this period, is now commonly divided into at least two separate species, H habilis and rudolfensis, by many workers in this field. For convenience, and in order to sidestep the naming controversies, ‘habiline’ is often used as a generic term for any of them; I will follow that usage here. The oldest habiline fossils are a bit more than 2 million years old. The find of Kimbel et al. (1996) may be the oldest Homo yet, at 2.3 million years. The dating of another find, KNM-ER 1470, caused a major controversy (Lewin, 1987), but when the dust had settled the consensus view converged on a younger date (but see Fitch et al. (1996)).

Habilines generally have somewhat larger brains than australopithecines, and have more humanlike teeth. But the differences between habilines and australopithecines, which are all so similar that they hardly deserve genera of their own.

Wood & Collard (1999) and Wood & Richmond (2000), however, argue for placing them in australopithecus. On the other hand, Wood (1992) argued forcefully for retaining them in Homo. (Yes, it’s the same Wood in both cases; seems he changed his mind.) And Cela-Conde & Ayala (2003) want to extend the genus Homo instead, encompassing not only the habilines but also Kenyanthropus platyops. See Section 5.3.4, in particular Figure 5.5.
Table 4.1a. *Overview of the various hominid species discussed in the text, including alternate names for the same fossils. The usage varies in the literature; the names designated as primary in the Table are used in the text. Species belonging to the genus *Homo* are listed in Table 4.1b. (m.y.a. = millions of years ago)*

<table>
<thead>
<tr>
<th>Species</th>
<th>Significance</th>
<th>Time (mya)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Ardipithecus kadabba</em></td>
<td>Earliest reasonably unambiguous bipedal hominid</td>
<td>6</td>
</tr>
<tr>
<td><em>Ardipithecus ramidus</em></td>
<td>Early bipedal hominid</td>
<td>4</td>
</tr>
<tr>
<td><em>Australopithecus aethiopicus</em></td>
<td>Very robust, small brain, possible ancestor of the other robusts</td>
<td>2.6–2.3</td>
</tr>
<tr>
<td><em>Australopithecus afarensis</em></td>
<td>Early hominid with good fossil record</td>
<td>4–3</td>
</tr>
<tr>
<td><em>Australopithecus africanus</em></td>
<td>First real ape-human intermediate to be found, in 1925</td>
<td>3–2</td>
</tr>
<tr>
<td><em>Australopithecus anamensis</em></td>
<td>Likely ancestor of <em>A afarensis</em></td>
<td>4.5–4</td>
</tr>
<tr>
<td><em>Australopithecus bahrelghazali</em></td>
<td>Fragmentary fossils, may well be <em>Africanus</em> or <em>garhi</em></td>
<td>2.5</td>
</tr>
<tr>
<td><em>Australopithecus boisei</em></td>
<td>Robust, ‘Nutcracker man’, coexisting with early <em>Homo</em></td>
<td>2.1–1.1</td>
</tr>
<tr>
<td><em>Australopithecus garhi</em></td>
<td>Plausible intermediate between <em>Australopithecus</em> and <em>Homo</em></td>
<td>2.5</td>
</tr>
<tr>
<td><em>Australopithecus ramidus</em></td>
<td>(see <em>Ardipithecus ramidus</em>)</td>
<td></td>
</tr>
<tr>
<td><em>Australopithecus robustus</em></td>
<td>Robust. May have used tools</td>
<td>2–1.5</td>
</tr>
<tr>
<td><em>Kenyanthropus platyops</em></td>
<td>Recently discovered, possible <em>Homo</em> ancestor</td>
<td>3.5</td>
</tr>
<tr>
<td><em>Orrorin tugenensis</em></td>
<td>Close to the common ancestor of humans and chimps</td>
<td>6</td>
</tr>
<tr>
<td><em>Paranthropus</em></td>
<td>(see robust <em>Australopithecus</em>)</td>
<td></td>
</tr>
<tr>
<td><em>Praeanthropus africanus</em></td>
<td>(see <em>Australopithecus afarensis</em>)</td>
<td></td>
</tr>
<tr>
<td><em>Sahelanthropus tchadensis</em></td>
<td>Close to the common ancestor of humans and chimps</td>
<td>6–7</td>
</tr>
<tr>
<td><em>Zinjanthropus boisei</em></td>
<td>(see <em>Australopithecus boisei</em>)</td>
<td></td>
</tr>
</tbody>
</table>
Table 4.1b. *Overview of the various species of humans (genus Homo) discussed in the text, including alternate names for the same fossils. The usage varies in the literature; the names designated as primary in the Table are used in the text. (m.y.a. = millions of years ago)*

<table>
<thead>
<tr>
<th>Species</th>
<th>Significance</th>
<th>Time (mya)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homo antecessor</td>
<td>Likely Neanderthal ancestor, claimed to be the common ancestor of us and Neanderthals</td>
<td>0.7</td>
</tr>
<tr>
<td>Homo cepranensis</td>
<td>(see <em>H</em> <em>erectus</em>)</td>
<td>1.7</td>
</tr>
<tr>
<td>Homo erectus</td>
<td>Human body but slightly smaller brain</td>
<td>1.8–0.2</td>
</tr>
<tr>
<td>Homo ergaster</td>
<td>Intermediate between habilines and <em>erectus</em></td>
<td>2–1.5</td>
</tr>
<tr>
<td>Homo floresiensis</td>
<td>Odd pygmy <em>erectus</em> descendant</td>
<td>0.1–0.02</td>
</tr>
<tr>
<td>Homo georgicus</td>
<td>(see <em>H</em> <em>erectus</em>)</td>
<td>1.7</td>
</tr>
<tr>
<td>Homo habilis</td>
<td>Resembles <em>Australopithecus</em>, but slightly larger brain and indications of tool use</td>
<td>2.3–1.8</td>
</tr>
<tr>
<td>Homo heidelbergensis</td>
<td>Intermediate between <em>H</em> <em>erectus</em> and <em>sapiens</em></td>
<td>0.5–0.2</td>
</tr>
<tr>
<td>Homo helmei</td>
<td>(see <em>H</em> heidelbergensis)</td>
<td></td>
</tr>
<tr>
<td>Homo neanderthalensis</td>
<td>Large-bodied large-brained Ice Age hunter. Coexisted with modern humans for millennia</td>
<td>0.5–0.03</td>
</tr>
<tr>
<td>Homo rhodesiensis</td>
<td>(see <em>H</em> heidelbergensis)</td>
<td></td>
</tr>
<tr>
<td>Homo rudolfensis</td>
<td>Close relative of <em>H</em> <em>habilis</em>, formerly believed to belong to that species</td>
<td>2</td>
</tr>
<tr>
<td>Homo sapiens</td>
<td>Know thyself (Linnaeus, 1758)</td>
<td>0.2–0</td>
</tr>
<tr>
<td>Homo sapiens (archaic)</td>
<td>(see <em>H</em> heidelbergensis)</td>
<td></td>
</tr>
<tr>
<td>Homo sapiens neanderthalensis</td>
<td>name sometimes used to distinguish strictly modern humans from Neanderthals and archaic <em>sapiens</em></td>
<td></td>
</tr>
<tr>
<td>Homo sapiens sapiens</td>
<td>(see <em>H</em> <em>neanderthalensis</em>)</td>
<td></td>
</tr>
<tr>
<td>Pithecanthropus erectus</td>
<td>(see <em>Homo erectus</em>)</td>
<td></td>
</tr>
<tr>
<td>Sinanthropus pekinensis</td>
<td>(see <em>Homo erectus</em>)</td>
<td></td>
</tr>
</tbody>
</table>
ithecines are slight, and might not have been considered significant, were it not for two factors:

- Stone tools turn up in the archaeological record at about the same time as habilines (Wood, 1997b; Semaw et al., 1997; Kimbel et al., 1996; Semaw et al., 2003), and keep on looking much the same throughout the tenure of the habilines (Semaw, 2000). These tools are further discussed in Section 10.2.
- Habilines (some of them, at least) are plausible ancestors of later, more clearly human, fossils.

There is not complete consensus about which habilines should be assigned to which species — see e.g., Prat (2002) — and no agreement on which species is the most probable ancestor of later Homo. H rudolfensis has a larger brain than Homo habilis sensu stricto (McHenry & Coffing, 2000; Lieberman et al., 1996), which might mark rudolfensis as the prime ancestor candidate, but the situation is more complicated when the rest of the body is taken into account. Some habilines, notably OH62 (Johanson & Shreeve, 1989; McHenry & Berger, 1998, but see also Korey (1990) and Haeusler & McHenry (2004)), may have had rather primitive apelike body proportions, making them unlikely ancestors of later Homo species. According to e.g. McHenry & Coffing (2000) and Strait et al. (1997), H habilis s.s. is the apelike one and rudolfensis more humanlike, but Lieberman et al. (1996) argue the exact opposite. In a later paper (1999), however, the first author of Lieberman et al. (1996) leaves the issue open instead. Blumenschine et al. (2003) argue that H rudolfensis isn’t really a separate species, but should be sunk into H habilis — but then they go on to argue that two species are still needed for the habilines, but the other one shouldn’t be rudolfensis.

The question of whether two fossil groups belonged to the same species or not, which in effect is a question of whether they would interbreed or not, is of course quite impossible to answer by direct experiment, as breeding experiments with dry bones yield meager results. Morphology, which is usually all we have for fossils, is not very helpful — among modern animals, we sometimes have different species that you need a microscope to tell apart, and at the same time there can be gross morphological variation within a single species. Among not a few paleoanthropologists, there is a regrettable tendency to coin a new species name for every new fossil that differs in the slightest from previous finds, leading to a proliferation of names of dubious value (White, 2003; Tobias, 2003). It is highly doubtful whether the various fossil hominids actually couldn’t interbreed — it is not even self-evident that humans and chimpanzees can’t. Still, there are many

27 Semaw et al. (2003) have found tools that are perhaps 200,000 years older than the oldest habilines yet found, and suggest that A garhi may have been the toolmaker. But the fossil record of habilines is very sparse, and it is imprudent to draw far-reaching conclusions from a discrepancy that is barely beyond the error margin in the dating.

28 The genetic differences between humans and chimps are comparable to those between horses and donkeys, who can and do interbreed, though the offspring is sterile. Nevertheless, the one experiment with human-ape hybridization of which I am aware did not produce any offspring (Rossiianov, 2002). Further experiments are unlikely, due to ethical constraints.
examples of modern species that can interbreed, but choose not to — very likely we and the Neanderthals formed such a species pair (see page 71 below). Possibly there are more examples in the hominid fossil record, but it is nearly impossible to tell in most cases. It is also conceivable that hybridization and even reticulation may have played a role in hominid evolution (Hollliday, 2003), but again there is little evidence one way or the other.

White (2003) and Conroy (2002) caution against overestimating the diversity of species and genera among fossil hominids. I basically agree with them, but for clarity and compatibility with the literature, I have nevertheless chosen to follow common usage as much as possible in this book, even though I have doubts about the status of some of these ‘species’.

After the habilines, evolution towards a more human form is rapid. Homo erectus (Figure 4.8), present in the fossil record from 1.8 million years ago, is indisputably human, with a body below the neck that is barely distinguishable from ours. Early African erectus fossils, as well as transitional forms between habilines and erectus (Walker & Shipman, 1996; Leakey & Lewin, 1993) are often called Homo ergaster, a usage which will be adopted in this work, even though there is no clear dividing line between ergaster and erectus. The brain size, and associated skull shape, is the main character distinguishing erectus from sapiens. Other differences are minor, though some, like the prominent brow ridges above the eyes, are pretty obvious.

Both the origin of erectus and the origin of the habilines coincide with periods of unusually dry climate in East Africa. The habitat variability caused by these episodes of climatic instability may have contributed to the evolution of these species (Wynn, 2004); cf. the ‘variability selection’ of Potts (1998), discussed on page 97.

Homo erectus is also the first in the human family to spread beyond Africa, probably as much as 1.7 million years ago (Gabunia et al., 2000; Vekua et al., 2002; Antón & Swisher, 2004) and certainly beyond 1 million years (Gibbons, 2001b). Fossils of erectus are widespread in Asia (‘Peking Man’ (Black, 1929; Andersson, 1934), ‘Java Man’ (Dubois, 1894; Delson et al., 2000; Broadfield et al., 2000), etc.), and probably reached Europe as well (Ascenzi et al., 1996; Oms et al., 2000; Rightmire, 1995; Clarke, 2000; Roebroeks, 2001), though the evidence is not totally clear (Tattersall, 1997; Ascenzi et al., 2000; Manzi et al.,

29Conroy (2002) supports his argument by comparing with other genera of mammals. He finds that among human-sized mammals, the average diversity is very low, rarely more than one or two species in each genus.

30There are unconfirmed reports of fragmentary finds in Asia that may be attributable to Homo habilis (Etler, 1996), but so far nothing substantial.

31It should be noted that Goguitchaichvili & Pares (2000) argue for caution in interpreting that dating. Also, assignment of these fossils to Homo erectus is not self-evident; their finders have recently reassigned them to a new species, Homo georgicus, and both H ergaster and H habilis (Balter & Gibbons, 2002) have also been considered. However, recent finds of very similar fossils in Africa would appear to argue against a separate Georgian species (Gibbons, 2003f).
Human origins and evolution

— some European *erectus*-like fossils have recently been renamed *H. cepranensis* (Mallegni et al., 2003; Schwartz, 2004). But despite the evidence of *erectus* or *erectus*-like fossils on various continents, and despite computer simulations of hominin dispersal (Mithen & Reed, 2002), according to which it would be surprising if *H. erectus* hadn’t reached both Asia and Europe, it is not self-evident that the migrations of *erectus* led to permanent settlements there, as the record is quite sparse (Dennell, 2003). As far as we know *erectus* did not reach the New World,\(^\text{32}\) but there are traces of their activities on the island Flores in what is now Indonesia, that was unreachable without boats (Morwood et al., 1998).

A very odd apparent *erectus* descendant was discovered recently on the same island. Formally named *Homo floresiensis* (Brown et al., 2004), and nicknamed ‘hobbits’ after the diminutive heroes of Tolkien (1937), these hominids are tiny, just one meter tall, and with an ape-sized brain. Nevertheless, a strong case is made for placing this species in *Homo*, close to *erectus*, based on numerous similarities in the skull, teeth, pelvis, and legs. The dating of the fossils is also remarkable, as they are very recent, the most complete one just 18,000 years old (Morwood et al., 2004). Stone tools and hints of fire were found in the same cave, and may be made by *H. floresiensis* (Morwood et al., 2004), but the tools are of a type normally associated with *H. sapiens*, not *erectus* (Mirazón Lahr & Foley, 2004); attributing a fairly advanced technology to someone with a chimp-sized brain requires strong evidence, particularly since *H. sapiens* was already present in the region at the time, and may well have reached Flores.

The material culture of *Homo erectus* included a well-defined set of standardized stone tools, known as the Acheulean industry (see Section 10.2), which represents a significant step forward from the crude tools of earlier humans (Oldowan

\(^{32}\)But see Turtledove (1988) for some entertaining speculations in fictional form.
industry). These are used by *erectus* everywhere except\(^{33}\) possibly Western Europe, where Oldowan tools have been found (Carbonell et al., 1999; Roebroeks, 2001). It is possible that *Homo erectus* used fire, since traces of fire, interpreted as ancient hearths, have been found in caves where they lived (Black, 1931), but the original evidence has been increasingly questioned (Balter, 1995; Wuetrich, 1998; Wu, 1999). Recently, however, Goren-Inbar et al. (2004) have found much stronger evidence of 800,000-year-old controlled fires in Israel. The maker of the fire is unknown, but the timing as well as the Acheulean cultural context implicates *Homo erectus*, though *H. heidelbergensis* (see below) is also a possibility (Rightmire, 2004).

About a million years ago, *Homo erectus* was a single, if not very uniform, species, spread across much of the Old World (Asfaw et al., 2002; Sanders, 2002; Gibbons, 2002a; Gilbert et al., 2003; Potts et al., 2004), but after that the populations on different continents started diverging. By half a million years ago, African, Asian, and European *erectus* descendants were clearly distinct. The last *Homo erectus* may have persisted in Asia until quite recent times, possibly as late as 25,000 years ago, well after the arrival of modern *Homo sapiens* (Gibbons, 1996b; Swisher et al., 1996, but see also Grün & Thorne (1997)).

In conclusion, the period from 2.5 million years ago up to 0.5 million years ago marks the transition from bipedal apes with no remarkable cognitive abilities, to people with brains near the modern human norm, and with the cognitive skills to develop a material culture effective enough to enable their spread across half the world.

### 4.3.3 Ourselves and other humans

This brings us, finally, to the origin of our own species, *Homo sapiens*. But again, just like with the habilines, the fossil situation is messy and complex, with many different fossils displaying different mixtures of *erectus* and *sapiens* characters. It is likely that there were at least two separate branches on the family tree during the last 500,000 years, one leading to us and one leading to Neanderthals\(^{34}\) (Figure 4.8). But it is unclear how isolated these branches were from each other – more on that below.

It may be noted that the evolution of human brain size was largely finished by this time — already fossils from 600,000 years ago are well within the modern human range (Conroy et al., 2000). Considerable cultural and presumably cognitive change has taken place since then, but not accompanied by any significant

---

\(^{33}\)It was formerly believed that this toolset was lacking also in East Asia (Tattersall, 1997; Ciochon & Larick, 2000), but Acheulean tools have recently been found in China (Yamei et al., 2000).

\(^{34}\)Sometimes spelled Neandertal. The place where the first one was found has changed its spelling from Neanderthal to -tal since then, and some, but not all, paleoanthropologists have followed suit. The traditional -thal spelling is used in this book.
additional brain expansion — the explanation for the major cultural differences between us and the other *erectus* descendants must be sought elsewhere.

The entire assembly of *erectus-sapiens* transitional fossils (Figure 4.9) is sometimes labeled as *Homo sapiens*, in which case the phrase ‘archaic sapiens’ is used for the more *erectus*-like ones and ‘anatomically modern sapiens’ or *Homo sapiens sapiens* is used to specify the late fossils that are indistinguishable from us living humans. Likewise, some people regard Neanderthals as just a subspecies: *Homo sapiens neanderthalensis* (Foley, 2004). Others prefer to split this group into several species. *Homo heidelbergensis* has gained a fair amount of acceptance in recent years as a label for the clear *erectus-sapiens* intermediates, as discussed in Rightmire (1995; 1996), whereas McBrearty & Brooks (2000) favour the name *Homo helmei* for the African intermediates, and Bermudez de Castro et al. (2004) propose to call them *H rhodesiensis*. But the naming of some recent fragmentary Spanish finds as a separate species, *Homo antecessor* (Bermudez de Castro et al., 1997; Bermudez de Castro et al., 1999; Arsuaga et al., 1999; Falguères et al., 1999; Carretero et al., 1999; Lorenzo et al., 1999), has met with some skepticism (Gibbons, 1997b; Delson, 1997), though the species is accepted e.g., in the general human-fossil review of Wood & Richmond (2000).

In the following, *Homo sapiens* will be used exclusively for humans anatomically indistinguishable from living people, the *erectus-sapiens* intermediates will be called *H heidelbergensis*, and *Homo [sapiens] neanderthalensis* informally ‘Neanderthals’. *H antecessor* will be used specifically for the Spanish fossils from which the species is named.

The issue of how many species there were in the *erectus-sapiens* interface is intimately entwined with a major debate concerning the origin of *Homo sapiens* that has been ongoing for some decades (Treisman, 1995), the fight between the two hypotheses:

**Out of Africa:** We evolved from a small subpopulation of *Homo erectus* (*sensu lato*) in Africa, and spread from there all over the world, replacing the resident *erectus* and *neanderthalensis* populations without significant intermixing. (Stringer & McKie, 1996)

**Multiregional evolution:** The different races of modern humans evolved separately from different local populations of earlier humans: Asians evolved from ‘Peking Man’ (Etler, 1996; Wu, 2004), Europeans from Neanderthals, Africans from African *erectus*, and so on. (Keith, 1936; Wolpoff et al., 1984; Trinkaus & Shipman, 1993; Wolpoff & Caspari, 1996; Hawks et al., 2000)

The debate is now approaching a consensus, in favour of the Out-of-Africa hypothesis, largely based on the recent arrival of massive amounts of molecular evidence (Cavalli-Sforza, 1998), as well as on difficulties with the multiregional hypothesis — separate evolution would generally lead to separate species, whereas
modern *Homo sapiens* is most definitely one single, genetically rather homoge-
neous species\(^{35}\) (Barbujani et al., 1997; Gibbons, 1995a; Hickson & Cann, 1997). The molecular data strongly support a fairly recent common origin for all ex-
tant humans, somewhere around 100,000 – 200,000 years ago (Ayala & Escalante, 1996; Wood, 1997a; Bergström et al., 1998; Cavalli-Sforza & Feldman, 2003). The so-called ‘Mitochondrial Eve’\(^{36}\) (Cann et al., 1987; Saville et al., 1998; Cavalli-
Sforza, 1998),\(^{37}\) the putative common ancestress of all women, was at the fore-
front of this molecular wave, but she has since been joined by a corresponding ‘Y-
chromosome Adam’\(^{38}\) (Fu et al., 1996; Dorit et al., 1995; Pääbo, 1995), as well as by data from non-sex-linked genes (Fischman, 1996), and from X-
chromosomes (Disotell, 1999). All this supporting evidence makes the fact that problems (Gib-
bons, 1998a) have been found with the original Eve studies (Cann et al., 1987) rather uninteresting; the total case for a recent common origin\(^{39}\) for all humans is highly compelling.

Proponents of the multiregional hypothesis, e.g., Hawks & Wolpoff (2001) or Relethford (2001a), try to rescue their case in different ways, e.g., by postulating large-scale gene flow between regions — but gene flow sufficient to keep humanity united in a single species would likely be sufficient also to wash out regional differ-
cences. But the continuity of regional differences was the basis of multiregionalism in the first place, and is still invoked by its supporters, e.g., Wolpoff et al. (2001) and Wu (2004). Other rescue attempts dilute the multiregionalism hypothesis into untestable meaninglessness (Foley, 1998).

The discovery of *Homo floresiensis* does not add support for multiregionalism (Mirazón Lahr & Foley, 2004), as it clearly demonstrates that the last known sur-
vivors of the Asian *erectus* lineage had evolved in a totally different direction, and

\(^{35}\)The standard biological criterion for regarding two individuals as belonging to the same species is based on their ability to produce offspring together. By this criterion, *Homo sapiens* is definitely a single species. I can personally testify that people of European and South American origin are interfertile.

\(^{36}\)Mitochondria are small bodies inside each cell, which carry their own DNA, and which are inherited from mother to daughter exclusively (Ankel-Simons & Cummins, 1997; Awadalla et al., 1999; Bromham et al., 2003; Kraytsberg et al., 2004, but see also Kivisild et al. (2000))). This makes it possible to trace descent through the female line. See the review of Stoneking & Soodyall (1996) for more details.

\(^{37}\)Brown (1980) had presented similar results several years earlier, but received less attention.

\(^{38}\)The Y-chromosome is inherited exclusively from father to son, so it traces the male line, in the same way that mitochondria trace the female line. It is interesting that both the male and the female line of descent are pointing in roughly the same direction. Males and females can have very different patterns of migration and gene flow (Pennisi, 2001), so their concordance strengthens the case for this really pointing towards the last common ancestor population of modern humans.

\(^{39}\)But the Adam-and-Eve talk should not be taken too literally, as has been done by some Biblical literalists (and some journalists) — the results by no means imply that only a single pair of humans existed at any particular time, and the time frame is nowhere near the Biblical one, There is, however, a fair bit of evidence for a tight population bottleneck in our not-too-distant past, perhaps 100,000 years ago (Ambrose, 1998; Gibbons, 1995a; Zietkiewicz et al., 1998; Gagneux et al., 1999a; Takahata & Satta, 1997) or even more recently (Marth et al., 2003), when humans apparently came perilously close to extinction (but see also Ayala et al. (1994), Harris & Hey (1999a) and Hey & Harris (1999)). Chimps, on the other hand, haven’t gone through any bottleneck, and retain high genetic diversity (Kaessmann et al., 1999; Stone et al., 2002).
Figure 4.9. Fossils of intermediate forms between Homo erectus and sapiens, that may be attributed to Homo heidelbergensis, found at Steinheim (left) and Broken Hill (right).

were clearly distinct from Homo sapiens. Substantial gene flow between our ancestors and the ancestors of H. floresiensis would have prevented the emergence of such a deviant form of human.

The case for placing the recent common origin of modern humans in Africa, rather than elsewhere, rests on five types of evidence:

1. **Fossils.** The oldest modern human fossils are found in Africa (White et al., 2003; Clark et al., 2003; Gibbons, 2003; Stringer, 2003; Tattersall, 1997; Pickford & Senut, 1998). Until the recent finds of White et al. (2003), both the dating and the identification of all African fossils from the relevant period was uncertain and contentious, leaving room for doubt (Wolpoff & Caspari, 1996). But the solid dates from Clark et al. (2003) make the situation much clearer, establishing the presence of near-modern humans in Africa between 154,000 and 160,000 years ago, at least 50,000 years earlier than anywhere outside Africa (but see also Faupl et al. (2003)).

2. **Archaeology.** The oldest archaeological traces of modern human behaviour, with more complex tools and ornaments, are found in Africa (Brooks et al., 1995; Gibbons, 1995), and appear to arise gradually there (McBrearty & Brooks, 2000), further discussed in Section 9.4.1 below, whereas the same tools and ornaments turn up much later and quite suddenly elsewhere.

3. **Genes** (Templeton, 1997; Cavalli-Sforza, 1998). When constructing our family tree from the present-day distribution of genes, it appears that non-African populations share an even more recent common ancestor with each other, than Africans do. This implies that the human family tree is rooted in Africa (Reich

---

40 Named by White et al. (2003) as a separate subspecies, Homo sapiens idaltu.

41 A fossil Homo sapiens skull from China has recently been dated to at least 68,000 years old, most likely 111,000-139,000, and possibly even older (Shen et al., 2002). There is, however, considerable doubt as to the provenance of the skull — was it really found in the layer that was dated?
Origins of language (Goldstein, 1998; Pääbo, 2003). This has been checked with various types of genes:

- Mitochondria: (Cann et al., 1987; Stoneking & Soodyall, 1996; Saville et al., 1998; Cavalli-Sforza, 1998; Parks, 1999; Disotell, 1999; Maca-Meyer et al., 2001; Huelsenbeck & Imennov, 2002), but see also Adcock et al. (2001) and Relethford (2001b).
- Different kinds of small repeated DNA pieces: (Batzer et al., 1994; Goldstein et al., 1995; Tishkoff et al., 1996; Ayub et al., 2003), but see also Pritchard & Feldman (1996).
- Y chromosomes: (Gibbons, 2001a; Ke et al., 2001; Thomson et al., 2000), but see also Brookfield (2000).
- X chromosomes: (Disotell, 1999), but see also Harris & Hey (1999b).
- More general methods with different types of genes (Ruvolo, 1996; Takahata et al., 2001)

Population genetic data also indicate that Africa had a larger population throughout the relevant period (Relethford & Jorde, 1999), and that non-African people may have passed through a bottleneck with very small population (Alonso & Armour, 2001; Rogers, 2001). There are recent indications, though, that the exact pattern of expansion from Africa may be more complex than just a single exodus (Templeton, 2002; Pavesi, 2003, but see also Cann (2002) and Pearson (2004)).


5. Parasites. The louse Pediculus humanus exclusively inhabits the human body surface, and spreads only from person to person. This means that louse populations follow human populations, so that the evolution of our lice tracks our own evolution and migrations. The patterns found when comparing DNA from lice found on different humans, resemble the patterns from the DNA of the humans themselves, and indicate that our lice also share a recent common ancestor in Africa (Kittler et al., 2003) — except for the odd fact that a second strain of lice is inhabiting some people, a strain that has been separated from our main tenants since the time of H. erectus (Pennisi, 2004a). This can be interpreted as evidence for intimate bodily contact between H sapiens and H erectus, before the latter went extinct.

A recent common origin for Homo sapiens is further supported by the recent recovery of DNA from eight different Neanderthal fossils (Kahn & Gibbons, 1997;

---

To prevent misunderstandings, perhaps it should be emphasized here that this does not imply that some living people are more ‘primitive’ than others, in the pejorative sense. All modern people are equally human as far as the important species-defining characters are concerned, specifically including our large brain. The minor details referred to here are things like the exact pattern of ridges and knobs on the teeth, details that can serve to disentangle the family tree but have absolutely no impact on what makes us human.
Ward & Stringer, 1997; Krings et al., 1999; Höss, 2000; Ovchinnikov et al., 2000; O’Rourke et al., 2000; Knight, 2003), showing that Neanderthals had diverged from *Homo sapiens* somewhere around half a million years ago or even as much as 800,000 years ago (Beerli & Edwards, 2002). Such an early divergence date is consistent with the Out-of-Africa hypothesis, but hardly with multiregionalism. The DNA data also show that the Neanderthals were not more closely related to present-day Europeans than to other modern humans, as the multiregional hypothesis would predict. Nor are there any genetic similarities between Neanderthals and the *Homo sapiens* living in Europe immediately after their demise (Caramelli et al., 2003; Serre et al., 2004; Cooper et al., 2004). But multiregionalism supporters, like Wolpoff (1998), Hawks & Wolpoff (2001), and Relethford (2001a), are not entirely convinced, and claim support from the DNA of an ancient Australian fossil, Mungo Man (Holden, 2001). And there is also a recent report of a fossil find of a possible Neanderthal-sapiens hybrid child, in Lagar Velho, Portugal (Duarte et al., 1999; Zilhão, 2002), which would argue for a closer relationship between the two, if its hybrid status were confirmed. On the other hand, the last surviving Neanderthals appear ‘pure’, with no admixture of *sapiens* genes (Hublin et al., 1996), and the earliest *sapiens* in Europe more resemble Africans than Neanderthals, as observed in body proportions by Holliday (1997; 1999) and in general morphology by Tyrrell & Chamberlain (1998). The preponderance of the evidence supports the wholesale replacement of Neanderthals by modern humans in Europe (Klein, 2003).

Whether Neanderthals and modern humans belonged to the same species is not an easy question to answer. The presence or absence of signs of hybridization is a good indicator, with the preponderance of the evidence, both morphological and genetic, here arguing for separateness (Tattersall, 1998a), but with the Lagar Velho child keeping the issue open. Another possible indicator is the degree of morphological and genetic differentiation. Neanderthals are clearly morphologically (Lieberman et al., 2002; Bruner et al., 2003) and genetically (Knight, 2003) distinct, but determining whether the difference is large enough to warrant placing them in a separate species requires comparison with how much difference there is between living species who do and do not interbreed. Schillaci & Froehlich (2001) present such a study, gathering data on several closely related monkey species all living on the same island (Sulawesi), some interbreeding and some not, and comparing with the Neanderthal-sapiens degree of difference. Their conclusion is that Neanderthals and modern humans are more different from each other than are even the monkey species that do not interbreed, and much more different than the monkey species that do interbreed. Harvati (2003) has performed a similar comparison with the two extant species of chimps, with similar results, extended with more

---

43 This article led to emotional outbursts that are remarkable even by the standards of paleoanthropology (Holden, 1999); see Tattersall & Schwartz (1999) and particularly the response of Trinkaus & Zilhão (1999). On a not totally unrelated note, this find was also regarded as of sufficient general interest to be spread by a major news service (Hatton, 1999), which is quite rare for any scientific result.
species in Harvati et al. (2004). If these monkey and ape results can be generalized to humans, this would argue against interbreeding between Neanderthals and *Homo sapiens*.

Neanderthals appear to have had a rather stable existence in Europe and western Asia for several hundred thousand years. They may have coexisted with *Homo sapiens* for quite some time in the Middle East, around 100,000 years ago, but recent studies indicate that the territory that is now Israel shifted hands back and forth in synchrony with climate variations (Gibbons, 2003d) — Neanderthals appear to have been better adapted to cold climates. Their body form was cold-adapted (Weaver, 2003), to the extent that it may have affected their locomotor efficiency substantially (Steadel-Numbers & Tilkens, 2004). Even so, during the height of the Ice Age in Europe they must have worn some form of clothing (Balter, 2004), as they inhabited sites with median winter temperatures around \(-16^\circ C\).

About 35,000-40,000 years ago the first *Homo sapiens* fossils turn up in Europe (Gibbons, 2003a), and not long thereafter the Neanderthals were driven to extinction, apart from the possible hybrid mentioned above. The last Neanderthal finds are from southwestern Europe, about 30,000 years ago (Hublin et al., 1996), which also may have been the area in which they evolved, as rather clear *erectus*-Neanderthal transitional fossils have been found there (Arsuaga et al., 1997; Dean et al., 1998; Rosas, 2001). For most of their existence, the Neanderthals used the so-called Mousterian set of stone tools, but the very last Neanderthal populations showed some technological and cultural progress, with more complex and diverse tools, as well as personal ornaments (Hublin et al., 1996; Simek & Smith, 1997; Gibbons, 2001c). This progress may have been influenced by the encroaching *sapiens* hordes, with which these Neanderthals must have been in contact for some time, though others argue in favour of independent invention (Bahn, 1998).

The exact cause of the demise of the Neanderthals remains uncertain and contentious (Balter, 2001d; Stringer & Davies, 2001), though interaction with *Homo sapiens* is high on the list (but see Finlayson (2004) for an alternative perspective). But this interaction may have been anything from out-and-out genocide to simple competition for food — and Neanderthals coexisted with modern humans in Europe for at least 2,000 years, and possibly as long as 10,000 years\(^{44}\) (Churchill & Smith, 2000; Gibbons, 2001c). There is some evidence of Neanderthals having led a much more strenuous life than contemporary *sapiens* (Gibbons, 1996a; Churchill, 1998), which may indicate their being less efficient food-gatherers, or may simply reflect the harsh periglacial environment which they inhabited — or may just be an artefact; see Holliday (2002) for some words of caution. But the cause of the difference in efficiency, if any (Sorensen & Leonard, 2001), is far from clear. Everything from the domestication of dogs (Sejnowski, 1999; Morell, 1997; Vila et al., 1997) to the origin of language (Maddox, 1998) has been invoked.

\(^{44}\)There are anomalies in the \(^{14}C\) record in the relevant period that makes the exact length of the coexistence period quite uncertain; it may have been rather short (Conard & Bolus, 2003).
assuming in both cases that the Neanderthals lacked dogs or language, which is far from obvious. Any straightforward physical takeover by modern humans is unlikely, since the Neanderthals were larger and stronger than we are (Gibbons, 1997a; Holliday, 1997; Kappelman, 1997; Ruff et al., 1997). There is some evidence that modern humans had a more varied diet than the mainly meat-eating Neanderthals (Richards et al., 2001), and also evidence from hand anatomy indicating that Neanderthals and modern humans used their tools differently, a difference that can be attributed to modern humans, but not Neanderthals, hafting their tools (Niewoehner, 2001). On the other hand, Briggs (2002) and Hardy et al. (2001) report what is interpreted as glue and other traces of hafting on tools associated with Neanderthals, and d’Errico (2003) reviews other evidence of Neanderthal hafting, as well as other ‘modern’ skills.

What is quite clear from the archaeological record is that the arrival of modern humans was accompanied by a marked increase in cultural diversity and inventiveness. The oldest traces of such a ‘modern’ culture are from Africa (Brooks et al., 1995; Gibbons, 1995b; McBrearty & Brooks, 2000), thus lending more support to the ‘Out of Africa’ scenario — cf. Section 9.4.1.

The extent to which Neanderthals had a human mind and human culture remains controversial — see Wynn & Coolidge (2004) for some ideas about the Neanderthal personality. Their brain is, if anything, larger than ours, but differently proportioned with smaller frontal and parietal lobes and larger occipital lobes. There is some evidence of ceremonial burial of (and presumably by) Neanderthals (Trinkaus & Shipman, 1993; d’Errico et al., 2003), which would imply a human level of awareness of self and others, but the evidence is still open to other interpretations (Gargett, 1999; Davidson, 2003) — whereas among Homo sapiens from the same time frame ceremonial burials are both common and unambiguous (Formicola et al., 2001). There is also evidence of cannibalism among Neanderthals (Defleur et al., 1999; Culotta, 1999b) and Homo antecessor (Fernández-Jalvo et al., 1999a), though the significance of that is unclear — there is clear evidence of cannibalism also in Homo sapiens (Marlar et al., 2000),45 and genetic hints that cannibalism may have been quite widespread among our recent ancestors (Stoneking, 2003)

45 Anecdotal reports of cannibalism have long been common in the anthropological literature, but have been dismissed by many as just signs of our wish to portray other people as primitive barbarians, a discussion reviewed in Lindenbaum (2004). The site excavated by Marlar et al. (2000) can, however, not be so lightly dismissed — finding traces of human flesh in human excrement demonstrates cannibalism beyond reasonable doubt.
Neanderthals apparently cared for their elderly and crippled members, as fossils have been found of Neanderthals with healed injuries and chronic diseases, who must have lived for years unable to fend for themselves (Trinkaus & Shipman, 1993; Onion, 2001; Lebel et al., 2001, but see also DeGusta (2003) and Cuozzo & Sauther (2004)). Fragments of what may have been musical instruments have also been found associated with Neanderthals (Zhang et al., 1999), though d’Errico et al. (2003) argue that Neanderthal ‘flutes’ are most likely just carnivore-damaged bones.

After the demise of the Neanderthals, and the roughly simultaneous disappearance of Asian erectus/floresiensis (Swisher et al., 1996; Morwood et al., 2004), the only humans left on Earth was a homogeneous population of modern Homo sapiens, which was fruitful and multiplied and rapidly spread to all continents except Antarctica. By 10,000 BC, essentially all inhabitable land was the territory of some hunter-gatherer tribe or another. Since their first appearance in Africa, modern humans have remained morphologically unchanged, apart from the minor superficial details that we use to tell apart people from different continents.

The uniform language capacities of all human populations today prove that all adaptations for language, biological and otherwise, must have been in place in the last common ancestor of all living people, who most likely lived more than 100,000 years ago, or at the very latest by the time different populations of modern humans parted company on their way to different continents. This happened 60,000 years ago or more (Cann, 2001; Cavalli-Sforza & Feldman, 2003; Coupé & Hombert, 2001), so the modern human language capacity, to whatever extent it is biologically based, cannot possibly be younger than that.

\[46\] While the Neanderthals may have been the oldest musicians, Lineweaver (2001) would place the oldest music some 13 billion years before their time, echoes of which are still detectable today.
4.4 Summary

Humans have evolved from ape-like ancestors, sharing a common ancestor with chimpanzees, who lived perhaps five million years ago or a bit more. Some details of the family tree are still obscure, and there is a wild variety of species and genus names of dubious validity in use, but the major patterns are quite clear. Some key figures are shown together in Figure 4.10.

Four main periods can be distinguished:

- 5–2 million years ago: Our ancestors were australopithecines, bipedal creatures with ape-sized brains. No clear sign of tool use or any other indication of cognitive capacities beyond that of apes.
- 2–0.5 million years ago: Our ancestors were *Homo ergaster/erectus*, people with human bodies but not quite human-sized brains. These people used tools and fire, and spread across Africa, Asia, and likely Europe.
- 0.5–0.1 million years ago: Two types of humans with full-sized brains evolved and lived in parallel, Neanderthals in Europe and *Homo sapiens* in Africa. They were most likely separate species who did not mix when they met, sharing a common ancestor more than 0.5 million years ago.
- 0.1 million years ago – today: Modern humans, *Homo sapiens*, spread out from Africa and conquered the world. Other types of humans were replaced by us, and rapidly disappeared.

The order of appearance of key human features:

- Bipedal walking, \( \sim \) 5 million years ago
- Stone tool making, 2.5 million years ago
- Human body below the neck, 1.8 million years ago
- Human-sized brain, 0.5 million years ago
- People indistinguishable from us, 160,000 years ago

Contentious issues:

- Who is our nearest living relative? *Resolved:* Chimpanzees (including bonobos)
— Who was the last common ancestor of us and chimps? Unresolved: Possible candidates include *Orrorin*, *Sahelanthropus*, and *Ardipithecus*, but very likely the right fossils are still undiscovered.

— What went on at the origin of the genus *Homo*? Unresolved: There is no consensus on the relations between habilines, *A. garhi*, *K. platyops*, and the various other fossils from the relevant time frame around 2.5 – 2 million years ago.


— Who were the Neanderthals? Resolved: Separate species of humans with their own culture, hunters in Ice Age Europe. Died out without issue around 30,000 years ago.

Further reading

CHAPTER 5

ANATOMICAL AND NEUROLOGICAL PREREQUISITES FOR LANGUAGE

The use of language is possible only for beings possessing a number of specific capabilities. Spoken language requires the physical ability both to produce and to perceive complex strings of sounds, making great demands on the performance of the vocal and hearing organs, discussed in Section 5.1 and 5.2. Other possible modalities of language, such as sign language, make corresponding demands within their own domains, e.g., manual dexterity and visual acuity. The emphasis here will be on speech both because it is our dominant modality today, and because there is considerable evidence that our vocal organs are specifically adapted for speech. Sign language will be discussed further in Section 9.6.

The greatest language-specific demands are, however, those placed on the processing capacity of the brain. A substantial fraction of the human brain is used in language processing, though there is no consensus on to what extent language is processed in physically separate, dedicated modules, nor any consensus on to what extent our language processing machinery is genetically determined. Evidence addressing these issues is reviewed in Section 5.3

5.1 Sound production

All mammals, with rare exceptions, are capable of producing a variety of sounds, and many of them use sounds for communicative purposes (Hauser, 1997). But human language, at least in the forms spoken by people today, requires a highly sophisticated vocal apparatus, capable of rapid sequences of precisely modulated sounds (Deacon, 1997), that are out of reach of most other mammals, specifically including our closest relatives (Savage-Rumbaugh et al., 1993). Apes can produce approximations of several human vowels, but not all, and have severe difficulties with consonants, particularly stops. They are, furthermore, unable to string sounds together in any patterns but their traditional calls. There are three main factors contributing to this difference between human and non-human primates:
The shape and flexibility of the vocal tract.

Voluntary neural control of the vocal apparatus.

Control of rapid sequencing of motor actions.

5.1.1 Vocal tract differences

In adult humans, the larynx has descended further down the throat, compared with its position in other primates, increasing the space available for articulation. Furthermore, the supralaryngeal tract in humans is bent almost 90 degrees, with the bend dividing it into two almost-equal pieces, unlike the fairly straight tube of most other mammals. This shape, together with our abilities to manipulate the shape further, broadens the spectrum of articulatory possibilities (Lieberman & McCarthy, 1999; Aiello, 1998). The larynx itself has not changed in any notable way (Hauser & Fitch, 2003), apart from the minor differences in the vocal chords discussed below. Figure 5.1 compares the vocal tract of a human and a chimpanzee.

In this particular case, ontogeny does recapitulate phylogeny (Haeckel, 1897), as human babies are born with the larynx in the ape position (Lieberman & McCarthy, 1999), which is better suited for breast-feeding. The larynx then descends as the baby graduates from breast-feeding to talking, gradually from age 3 months to 3 years or so (Fitch, 2000b). The larynx of a chimpanzee infant also descends a little bit, but not as far as in humans (Nishimura et al., 2003). It appears that the phylogenetic descent of the larynx was a two-step process — the laryngeal skeleton descended relative to the hyoid bone sometime before the last common ancestor of humans and chimps, and the hyoid bone itself descended in the hominid branch after separating from chimps.

Having the larynx permanently lowered is commonly believed to be unique to adult humans. However, Fitch (2000a) urges some caution in the interpretation of the higher resting position of the larynx in other mammals, as many of them are capable of lowering the larynx temporarily, during loud vocalizations, and Fitch (2002a) presents examples of other mammals with a permanently lowered larynx. Nevertheless, the risks associated with the permanent lowering in humans must be compensated by evolutionary advantages not attainable with the standard mammalian larynx position, however flexible it may be.

The inability of apes to produce stop consonants is largely due to the anatomical impossibility of their achieving rapid velopharyngeal closure (closing off the connection between nose and mouth), without which air isn’t forced towards the

---

1Haeckel’s recapitulation law is of course totally discredited as a general biological law, for both theoretical and empirical reasons. But this does not mean that recapitulation-like ontogenetic trajectories never happen.

2A bone which is situated in the floor of the mouth, providing attachment for muscles moving the tongue and larynx (Aiello & Dean, 1990), thus playing a crucial role for speech.
constrictions that we use to make consonants (Savage-Rumbaugh & Lewin, 1994).

There have been some controversies concerning at what point in human evolution did the descent of the larynx take place. It has been argued that it did not take place until the advent of *Homo sapiens*, and that Neanderthals retained an apelike vocal tract (Lieberman & Crelin, 1971), but others would place it earlier (Falk, 1975; Le May, 1975), possibly as early as *Homo ergaster* (Aiello, 1998). The position that the Neanderthal vocal tract differed substantially from ours still has its defenders (Schwartz & Tattersall, 1996; Laitman et al., 1996, but see Franciscus (1999)), but appears hardly tenable after the discovery of a Neanderthal hyoid bone (Arensburg et al., 1989). This bone was very similar to a *sapiens* hyoid, leading to the conclusion that Neanderthals had a human vocal tract (Houghton, 1993) or near enough (Boë et al., 1999), and were anatomically capable of human speech (Arensburg et al., 1990; Arensburg & Tillier, 1991; Boë et al., 1999; Heim et al., 2002; d’Errico et al., 2003).

---

3 But see also Kennedy & Faumuina (2001), who question how unique this morphology really is.

4 Also Philip Lieberman (of Lieberman & Crelin (1971), the seminal paper arguing for Neanderthal muteness) appears to have changed his mind, at least partially, according to Holden (1998) and Aitchison (2001). (Note that Lieberman & Crelin (1971) and Lieberman & McCarthy (1999), both cited in this section, are two different Liebermans.)
The stretching and bending of the supralaryngeal tract is associated with various changes in the cranial base, reviewed in Lieberman et al. (2000). Hints of the cranial base changes can be seen already in *Homo ergaster* (Tattersall, 1998b), and *H. erectus* has a near-modern cranial base (Baba et al., 2003; Gibbons, 2003c), though Fitch (2000b) cautions against drawing far-reaching conclusions. The cranial base is affected by our increased brain size, by the change to bipedal posture, and by the shrinking of our jaws (Spoor, 2000), all of which will confound any effect from speech adaptations.

It would appear then that the human vocal tract has an ancient history, dating back at least to our common ancestor with Neanderthals, perhaps half a million years ago. This common ancestor would be *Homo erectus* or a close relative, thus extending the age of the human vocal tract into the time of *erectus*, half a million years ago or bit more. Reeves et al. (1996) argue for a larynx move around two million years ago, among the first *Homo*, not primarily for speech, but as an adaptation to a drier habitat, and Pearl (2000) places it even earlier, as a byproduct of the transition to bipedalism. Its later utility for speech can in these cases be regarded as an exaptation.

Another proposed reason for the descent of the larynx is sexual selection favoring a deeper voice in males (Randerson, 2001). This is a plausible explanation for the additional descent taking place in males at puberty — human females do find a deeper male voice more attractive, according to Collins (2000); see also Section 10.3 — but hardly for the whole process. Larynx descent is also observed in deer, and may in their case indeed be interpreted in terms of selection for a deeper voice rather than speech, as stags do use their deep voices to reproductive advantage (Fitch & Reby, 2001).

A puzzling aspect of larynx evolution in humans is the loss of the laryngeal air sacs that most of our relatives have. The sacs almost certainly play a role in the vocal communication of apes, but little is known of the details. According to Hewitt et al. (2002), the enhanced breathing control that has evolved in humans (see Section 5.1.2 below) may have made the sacs superfluous. Nevertheless, the loss of a vocal adaptation in such a vocal species as ours is odd (Fitch, 2000b).

The human vocal tract in its current form is exquisitely adapted for speech, at considerable evolutionary expense (Pinker & Bloom, 1990; Spangle & Menzel, 1991; Pinker, 1995; Fitch, 2000b; Deacon, 2000). This makes it somewhat problematic to argue for its origin as a rather marginal improvement for some other purpose, or just a byproduct of some other adaptation. If the modifications of the vocal tract did evolve for the purpose of facilitating speech, as appears highly likely, this implies that our ancestors had some form of spoken language before they had a human vocal tract — evolution does not plan ahead!

The vocal chords in humans also differ in several respects from those of other animals, that can reasonably be interpreted as speech adaptations (Sataloff, 1993), though Sataloff’s claim that the vocal ligament is found only in humans is contra-
dicted by both Suzuki et al. (1988) and Genack et al. (1993), who report on studies of vocal ligaments in rodents. Still, the multiple layers with fine-tuned elastic properties in the human vocal chords, as well as the shock-absorbing ligament attachment structures, do appear adapted to heavy use and complex vibrations (Sataloff, 1993; Lindblad, 1992). It is a bit odd, though, that the full development of the chord structure isn’t attained until after puberty (Sataloff, 1993), something which may indicate that the chords are at least partly shaped by sexual selection rather than adaptation for speech.

Humans possess a more mobile and flexible tongue as well, allowing even more speed and flexibility in articulation.

### 5.1.2 The neural wiring of the vocal apparatus

Humans have direct conscious control of their vocal chords, whereas apes have access only through indirect means, that cannot be easily modulated voluntarily, similar to human laughter and other non-voluntary sounds (Deacon, 1992). The different level of control is connected with different neural wiring; humans have pathways from the neocortex to the vocal chords, whereas the vocalizations of other primates are controlled from older levels in the brain. This limits the range of sounds that apes and monkeys can produce to minor variations around largely innate templates. Apes do make good use of the sound pattern variation they can produce (King, 2003), and with extensive training, some modest progress can be made, but cross-fostering experiments between different primate species (ape-human and monkey-monkey) have been near-total failures — even monkeys cross-fostered with another monkey species whose vocalizations ought to be within reach anatomically, failed (Fitch, 2000b). The easy vocal learning, vocal imitation, and vocal innovation of humans appears to be unique among primates, signifying a new level of neural control of the vocal apparatus. The vocal apparatus must come under conscious control — but at the same time the innate automaticity of primate vocalizations must be relaxed. What is first needed here is actually not primarily selection for speech, but the removal of selection for innate automatic vocalizations, leaving room for vocal learning to develop (Deacon, 2003a). The expansion of the neocortex in humans compared to other mammals, discussed in Section 5.3 below, may provide a mechanism for building the new pattern of connections — due to their larger numbers, cortical neurons may simply out-compete neurons from the old system in the neural Darwinism (cf. page 31) of the growing brain (Deacon, 2000).

But both pathways are still active in humans. We do have a few vocalizations that are controlled the old way, e.g., laughter, and we all know how difficult vol-

---

5A similar process has been observed in a domestic bird (Kazuo Okanoya, cited by Deacon (2004b)).
untary control of such vocalizations can be if we are aroused, and how they can interfere with speech, overriding the new control system (Deacon, 2000).

The standard mammal neural wiring of the vocal chords includes tight connections between the procedures for swallowing and breathing, that need to be precisely synchronized. This system has apparently been re-wired in humans in evolutionarily recent times — as everybody who has choked on a piece of food knows, some debugging remains to be done (Deacon, 1997). The situation in other primates is, however, not entirely clearcut, and may be interpreted as intermediate between the human and the standard mammalian condition (Mueller, 1996). Humans also have significantly more motor neurons going down the spine to the thorax (presumably for breathing control), which may well be an adaptation for speech (MacLarnon & Hewitt, 1999). This would be visible in well-preserved fossils, but as recently as Homo ergaster it is absent (Walker & Shipman, 1996), as it is in australopithecines. Neanderthals, on the other hand, are similar to modern humans in this respect (MacLaron & Hewitt, 1999), indicating that the enhanced breathing control evolved sometime between 1.6 and 0.5 million years ago. If language was the driving force behind this evolutionary step, then late Homo erectus must have possessed language in some form.

This enhanced breathing control has also been linked to bipedal walking, as effective quadrupedal walking requires close coordination of breathing pattern and stride. The independent breathing control necessary for speech wasn’t possible until we were walking upright (Melichar, 1999).

A similar case of nerve canal enlargement is the hypoglossal canal, which is a conduit for the nerves controlling the tongue. The hypoglossal canal is significantly larger in humans than in most other apes, again presumably because of the finer control of tongue movements necessary for speech. Concerning the fossil situation here, Kay et al. (1998) and DeGusta et al. (1999) are in direct contradiction. According to Kay et al. (1998) australopithecines have ape-sized hypoglossal canals, and later Homo have modern ones, consistent with what is found for vocal tract shape, whereas DeGusta et al. (1999) state that human-sized hypoglossal canals are found both in some australopithecines and in various nonhuman primates, making any conclusions concerning speech effectively impossible.

Concerning neural control of articulation, MacNeilage (1998) proposes a detailed theory of its evolution, invoking the rhythmic jaw motions of chewing and sucking (still visible in infant babbling) as exaptations, a suggestion that is further discussed in Section 11.1.

---

6 It has been suggested by Verhaegen (1998), a proponent of the Aquatic Ape Theory (Morgan, 1982, but see also Langdon (1997)) that the breathing control is instead an aquatic adaptation, later co-opted as an exaptation for language. But MacLaron & Hewitt (1999) reject both this and other alternative explanations for the nerve canal enlargement.
5.1.3 Control of rapid sequencing

Humans have the ability to produce rapid complex sequences of actions, an ability which is obviously vital to speech (Studdert-Kennedy, 2000), but which is also used for e.g., accurate throwing (Calvin, 1993; Calvin, 2003). The action sequences performed by humans are too rapid for sensory feedback, so the sequences must be pre-planned, and executed with split-second timing. Other animals also use rapid action sequences, but these are generally shorter and more stereotyped than what humans are capable of (Ott et al., 1994).

The timing precision requirements, both in speech and in throwing, is a matter of milliseconds. For example, throwing a ball and hitting a 20 cm target 8 meters away, which most people are capable of, requires the ball to be released within a time window on the order of 2-3 milliseconds, with the various muscle actions involved coordinated at a similar level of precision (Chowdhary & Challis, 1999). It takes around 100 times longer for the brain to receive and act upon any return signals from the throwing arm, so the throw has to be entirely pre-planned. A further complication is that individual neurons in the brain do not have that kind of timing accuracy — the intrinsic timing jitter in neuron firing is on the order of 10 ms. Higher precision can be achieved by averaging the signals from multiple neurons (Ivry & Richardson, 2001), which is both mathematically and neurologically straightforward, but the number of required neurons grows exponentially with the precision requirements (Calvin, 1993). Ivry & Richardson (2001) have shown experimentally that averaging the timing signal from several timing circuits in the human brain does improve timing precision in rhythmic tasks.

In comparison with humans, monkeys perform miserably in throwing tasks — they need to be ten times closer to the target to achieve comparable accuracy (Westergaard et al., 2000, cited in Watson (2001)) — supporting Calvin’s (1993) case for precision throwing as a human adaptation.

Some speech features, such as voice onset time, also need to be produced with a timing accuracy much better than any conceivable feedback loop time (Ladefoged, 1971; Cho & Ladefoged, 1999). Before we receive any feedback on our speech production, we are several phonemes further ahead, so the feedback comes much too late to do any good. Nevertheless, we manage to produce streams of dozens of phonemes, each of which requires several separate motor actions (tongue, lips, vocal chords, lungs, ...) to be coordinated to within 10 ms or so.

The importance of sequencing for language is supported by studies of SLI (Specific Language Impairment; see page 103). One family of SLI sufferers “have a severe impairment in the selection and sequencing of fine orofacial movements, which are necessary for articulation ...” combined with deficits in several other aspects of language processing (Lai et al., 2001, p. 519, emphasis added) that can plausibly be interpreted as sequencing-related. Similar cases are reported also by Joanisse & Seidenberg (1998).
Sequencing may be important not only in production. The decoding and segmenting of speech must work at speeds comparable to those of speech production, detecting features in incoming sounds with millisecond precision, and going through all the multiple steps in the decoding of a full sentence in a time not significantly longer than the time it takes to say the sentence. This requires a different kind of rapid-sequence handler, the origins of which is unclear. But similar, if less complex, sequence decoding is part of the communication processing of many animals, as reviewed in Pollack (2001), with precision timing at the neural level discussed in Carr et al. (2001).

The learning of sequences is also implicated in language, particularly acquisition but also other aspects, according to Stowe & Haverkort (2003). In this context, it is worth noting that while non-human primates are quite capable of learning straight sequences, there are limitations on their performance when it comes to hierarchical structure in sequences (Conway & Christiansen, 2001).

There are multiple proposals for the brain location handling this rapid sequencing of both speech and other actions. Broca’s area (Broca, 1861, further discussed in Section 5.3.3) is one candidate (Passingham, 1981; Wilkins & Wakefield, 1995), which may provide an alternative explanation for the nonfluent speech of Broca’s aphasia. In further support, Calvin (1993) cites cases of patients with aphasia who also had trouble with motor sequences in general (but without general paralysis; the actions could be performed one by one, but not in rapid sequence). Similarly, Heim et al. (2003) cite evidence that Broca is involved in the processing of rhythm and musical sequence.

Calvin argues that a large chunk of the ‘language cortex’ on the left side of the brain is actually a general ‘sequence handler’, which handles language but also other sequencing tasks. Hickok et al. (1998a) disagree, citing evidence from sign-language aphasia, but Corina (1999) identifies gaps in Hickok et al’s arguments, and labels their conclusion premature.

But other parts of the brain are also involved in sequencing. Richards & Chiarello (1997) propose a right-hemisphere role in higher-level action planning. Lieberman (2001b) suggests that the subcortical basal ganglia have a key role in sequencing, a suggestion also receiving some support from SLI studies (Balter, 2001a) and from patients with damaged basal ganglia (Holden, 2004b), and Desmond & Fiez (1998) cite evidence of the actual timing circuits being located in the cerebellum. This proposed role for the cerebellum is supported by both lesion and functional imaging studies reported in Ackermann & Hertrich (2000), but Mauk et al. (2000) interprets the cerebellum as the seat of more generalized temporally specific learning. Lesions of the cerebellum commonly cause minor impairments of a variety of language components, something which Fabbro et al. (2000) interprets as:

---

7A simple and pertinent experiment would be to compare the throwing accuracy of patients with different types of aphasia, but to my knowledge this has not been done, though Richards & Chiarello (1997) do refer to deficits in motor action planning among aphasia patients.
cerebellar lesions [...] may be related to altered language control processes (such as the control of the temporal and sequential organization [...]) rather than to impairment of specific modules of the language system [...]. (p. 185-186)

The evolution of a sequencing module may possibly, as argued by Calvin (1993), have been a crucial step in human evolution. It may originally have been an adaptation for improved throwing or other motor sequences, driven by straightforward natural selection for better hunting skills, something which may be supported by the persistent sex differences in throwing accuracy among modern humans (Watson, 2001). But once established, for whatever reason, such a sequencing module would be a vital exaptation for language, making the eventual evolution of speech much easier (Wilkins & Wakefield, 1995; Calvin, 1993). Concerning the timing of its evolution, the first unequivocal throwing weapons in the archeological record are about 400,000 years old (Thieme, 1997), but, as Calvin (1993, but see also Wynn (2002)) notes, Acheulean ‘hand’ axes (see Section 10.2), a million years older, would make excellent throwing weapons as well.

5.1.4 When did we start to speak?

The evidence concerning the evolution of humanlike speech capabilities is not conclusive; we do not know for sure when it evolved, though it may well have been evolutionarily fairly recent, most likely sometime during the past two million years or so. Some type of speech must have been present in our last common ancestor with Neanderthals, 500,000 years ago or so, though fully human speech with all our articulatory capacity need not be much older than 100,000 years.

But this does not necessarily mean that language as such evolved at the same recent time. To begin with, it is conceivable that language first evolved in some other modality than oral speech, a possibility further discussed in Section 9.6. This means that the emergence of speech adaptations does not date the origin of language, but does set a lower limit on the age of language, unless one assumes that complex oral vocalizations were used for some other purpose than language.

Finding an upper limit on the age of language is more difficult. The emergence of speech adaptations doesn’t even set a firm upper limit on the age of spoken language. Speech does not strictly require the full modern human set of speech sounds, but could well make do with a smaller number of phonemes, emitted by a simpler and more primitive vocal apparatus. In fact, no modern language uses the full capacity of our speech organs — there are languages that use as few as 11 sounds out of the hundreds that we are capable of producing (Vaneechoutte & Skoyles, 1998). It is thus not difficult to imagine proto-speech, that could be spoken with a proto-human larynx with a much smaller repertoire of sounds. And the vocal apparatus of apes and monkeys does appear capable of producing a non-negligible number of phonemes that are usable for communication (if subtle to human ears), with a richer structure than has often been believed (Rendall et al.,
1999; Fischer, 1998). This structure includes features like fundamental frequency declination (Hauser & Fowler, 1992), which has sometimes been claimed as a human language universal, as well as others (Maurus et al., 1988).

But once our ancestors did possess speech in any form, they would likely have been under a fair bit of selective pressure for improving the vocal apparatus (Donald, 1997). This is the plausible time (whenever it was) for the descent of the larynx. There is also some evidence for the co-evolution of the shape of the vocal tract, its motor control system, and the human vowel system (Honda & Kusakawa, 1997), further strengthening the case for speech driving the evolution of the vocal apparatus.

### 5.2 Sound perception

The auditory apparatus as such is essentially identical in humans and other primates, and very similar in all mammals. Measured auditory abilities of humans and other apes do not differ significantly (Spoor & Zonneveld, 1998), apart from the tuning (Martínez et al., 2004) discussed on page 88 below. The processing of sounds in the brain also appears similar, as described in detail by Kaas & Hackett (2000), at least up to the point where the processing becomes linguistic in a narrower sense, at which point human processing of language sounds shows a pattern of neural activation different from that of other sounds (Mueller, 1996).

The evolution of the mammalian ear is described in many textbooks, such as Cowen (1995), and also in e.g., Rowe (1996), with Manley (2000) and Fritzsch et al. (2002) covering the history of the inner ear in more detail. Clack (1998; 2002) covers the earliest history of our ears. The advantages of the mammalian ear, compared with its predecessors, consists mainly in a wider frequency range, extended towards higher frequencies (Meng & Fox, 1995; Manley, 2000).

It would appear then that the hearing of most mammals ought to be sufficient for speech perception, as is demonstrated by our everyday experience of e.g., dogs and horses responding to voice commands. As for the finer details of phonetic analysis, this does not appear to be a uniquely human skill either (Zoloth & Green, 1979). Monkeys are quite able to discriminate human phonemes (Sinnott, 1989; Sinnott & Adams, 1987; Sinnott & Kreiter, 1991), as are even guinea pigs (McGee et al., 1996) and some birds (Mueller, 1996). As noted by Savage-Rumbaugh et al. (1993), this is counterevidence against the fairly popular view that humans have a unique ‘phonetics module’ that is needed to decode our phonemes.

---

8Wray (1998) postulates an early phase of language evolution, without grammar. This could be called a one-word stage (see page 235), but Wray (1998; 2000) prefers to talk about holistic unanalyzed phrases (see also page 234). A system with a large number of wordlike units but no grammar may well, as argued by Wray (1998) and also from a different perspective by Nowak and associates (1999; 2000), place even heavier demands on phonetic differentiation than modern speech. This would strengthen the selective pressure towards articulatory richness and a complex vocal tract during early speech evolution.
Both monkeys (Rendall et al., 1996; Ghazanfar & Hauser, 2001) and various other mammals, such as elephants (Poole, 1999), seals (Insley, 2000), hyaenas (Holekamp et al., 1999) and dolphins (Sayigh et al., 1998; McComb et al., 2000), have sufficient aural discrimination to recognize individual voices of conspecifics (and have the social intelligence needed to use this information). In the case of dolphins, they appear to be using a kind of ‘signature whistle’, different for each individual, and functioning very much like a name (Tyack, 2000; Janik, 2000; Watwood et al., 2004, but see also McCowan & Reiss (2001)).

Likewise, the time resolution necessary for speech perception (around 25 ms) is present in mammals in general, and appears to be the basis even for such an apparently language-specific feature as the human categorical perception of voice onset time (Mueller, 1996; Simos et al., 1997). Categorical perception of entire calls is observed in the processing of vocalizations among several species of apes (Hauser, 2000), though the results of Fischer et al. (2001) and Prell et al. (2002) are less clearcut. Monkeys also show categorical perception of human phonemes, with perceptual boundaries close to where humans have them (Kuhl, 2000). The general auditory processing in primates is reviewed in Kaas et al. (1999), and the processing of vocal communication in Wang (2000).

It is well established that humans, even pre-verbal infants, are very good at perceiving speech, and segmenting it into language-relevant units (Kuhl, 2000; Saffran et al., 2001; Jusczyk, 1999), a capability which has been invoked as an innate language specialization, notably by Chomsky (1957, cited in Kuhl (2000)). Interestingly enough, neuroimaging performed on human infants during the course of this type of experiments show that they use largely the same brain areas as adults when perceiving speech (Dehaene-Lambertz et al., 2002).

However, the sound-pattern analysis abilities that infants use for this purpose have been found to be usable also for the analysis of non-speech sounds, such as music (Saffran et al., 2001), which indicates that the abilities are not language-specific. The experiments establishing speech-analysis capabilities in infants have also been replicated in monkeys, who performed at the same level as the human infants (Hauser et al., 2001), from which it can be concluded that this segmentation capability has not evolved for the purpose of speech perception. Similarly, Ramus et al. (2000) have found that monkeys and human infants are both good at discriminating between different unfamiliar human languages, when played forwards but not when played backwards, and Toro et al. (2003) have even replicated this result with rats.

One may thus conclude that human hearing in general, as well as phonetic discrimination and segmentation in particular, has not evolved especially for speech perception, but was available as an exaptation among our ancestors. This makes more sense than the converse, as it is evolutionarily more reasonable that the

---

9Newborn infants can even do this in their sleep, learning to discriminate sounds while asleep (Cheour et al., 2002). In adults, learning of any kind during sleep has been notoriously difficult to establish.
Origins of language ‘choice’ of phonetic distinctions to be used in communication was guided by the pre-existing discriminatory capacity of our hearing system, rather than having our hearing evolve in order to perceive speech.

Some fine-tuning appears to have taken place to optimize speech perception (Pinker & Bloom, 1990; Moore, 2000a; Martínez et al., 2004) but the overall system is basically unchanged from our ancestors. One aspect of this fine-tuning that is particularly interesting, as it is diagnosable in fossils, is our improved perception of sounds in the 2–4 kHz range. The sensitivity of ape ears has a minimum in this range, but human ears do not, mainly due to minor changes in the ear ossicles, the tiny bones that conduct sound from the eardrum to the inner ear. This difference is very likely an adaptation to speech perception, as key features of some speech sounds are in this region. The adaptation interpretation is strengthened by the discovery that an inner-ear structural gene has been the subject of strong natural selection in the human lineage (Olson & Varki, 2004).

According to Martínez et al. (2004), these changes in the ossicles were present already in the 400,000-year-old fossils from Sima de los Huesos in Spain, well before the advent of modern Homo sapiens. These fossils are most likely Neanderthal ancestors, that Martínez et al. (2004) attribute to Homo heidelbergensis.¹⁰

The presence of speech adaptations in the form of speech-tuned hearing implies that the possessors had speech in some form. The discovery of Martínez et al. (2004) thus sets a firm lower limit on the age of speech at around 400,000 years. Furthermore, as these fossils are on the Neanderthal lineage, the origin of speech is pushed back at least to the common ancestor of us and Neanderthals, 500,000 years ago or more, consistent with the evidence from speech production discussed in the previous section.

5.3 Brain anatomy, modularity, and lateralization

Our brain is our most complex organ, by far, and it is also the seat of the most important language adaptations. Unfortunately, our knowledge of this complex organ is far from complete, and is not really sufficient to determine how the brain handles language, other than at a gross anatomical level.¹¹ The study of brain function has nevertheless made great strides forward in the past decade or so, due to new non-invasive techniques, such as PET, MEG, ERP, and MRI, making possible imaging

¹⁰I disagree with this attribution. As discussed in Section 4.3.3 I prefer to use H Heidelbergensis for the fossils that are intermediate between erectus and sapiens. I would call these Spanish finds Homo neanderthalensis, as they are probably on the lineage leading to classical Neanderthals, and almost certainly more close related to Neanderthals than to us.

¹¹There is only one animal whose brain anatomy is known in adequate detail to determine its exact function, and that is the nematode Caenorhabditis elegans. We can follow its growth, cell by cell (De Pomerai, 1990) and have mapped every single neuron (all 302 of them) in its nervous system, including the brain (Bargmann, 1998; Buller & Hardcastle, 2000). The linguistic talents of this tiny worm are, however, modest, so it is of no great relevance in this context.
of the functioning brain (Horwitz et al., 1999; Lounasmaa et al., 1996; Osterhout
et al., 1997; Menon, 2001).

Brain evolution starts very early indeed in our history, in Precambrian times. The common ancestor of all bilaterian\textsuperscript{12} animals most likely possessed something that might be called a brain (Nielsen, 1999), though a true brain, clearly homologous with ours, is a vertebrate characteristic (Pendleton et al., 1993; Shimeld & Holland, 2000).

One of the few even semi-consistent trends in evolution is towards larger and more complex brains, towards increasing encephalization, a trend that can be related to the argument of Ng (1996), that increasing complexity in nature favors the evolution of what Ng calls ‘rational species’ (1996, p. 303) with larger brains — but the presence of rational species increases the complexity of the environment for other organisms, creating a complexity ⇔ brain feedback loop.

Even so, this trend towards larger brains is really clear only within the vertebrate lineage. Fish had (and have) fairly small and simple brains, but significant modification took place with each major evolutionary step, starting when the first fishes acquired legs and went ashore (Ahlberg et al., 1996; Coates, 1998). Evolutionary brain growth has taken place both through enlargement of already existing structures, and through the addition of new structures on top of the old ones. People sometimes refer to the ‘reptilian’ part of our brain (Lieberman, 2001b) — this expression is quite literally true, since that part of the brain does correspond to structures that we have in common with reptiles. On top of the reptilian structures, mammals and birds have independently evolved an additional layer, with considerably larger processing capacity. The bird brain, present even in the earliest bird fossils (Domínguez Alonso et al., 2004), is basically an expanded reptile brain, anatomically quite different from the mammalian brain, but functionally comparable, at least up to a point. Figure 5.2 shows the relation between brain size and body size for different vertebrate classes. The distinction between fish and reptiles on one hand, and mammals on the other, is evident. Primates cluster along the upper edge of the mammalian distribution, with humans together with dolphins as outliers well above the others.

The mammalian top layer is known as the neocortex, or isocortex. It evolved during the reptile-mammal transition 200 million years ago, in parallel with the ear changes mentioned in the previous section (Rowe, 1996; Karten, 1997; Kaas & Reiner, 1999; Aboitiz et al., 2003). It is not just an additional layer of brain tissue, but adds a novel organizational principle (Nishikawa, 1997; Karten, 1997), opening the way for increasing complexity and processing power. The details of its origins are not quite clear, with at least two competing hypotheses (Kaas & Reiner, 1999; Nishikawa, 1997), but embryological studies with molecular tech-

\textsuperscript{12}Bilaterians are those animals that are bilaterally symmetric (or nearly so; see Section 5.3.2). This includes pretty much everything that we see as animals in our everyday world, from flatworms to fruit flies, the only prominent exceptions being jellyfish and their relatives.
Figure 5.2. Brain size plotted against body size for various species of vertebrates, with different symbols for different classes, as shown above.
Techniques are approaching an answer (Karten, 1997; Deacon, 1997; Chenn & Walsh, 2002; Aboitiz et al., 2003).

The basic brain structure of all different mammals is similar, with mainly quantitative differences (Nishikawa, 1997). The human brain is not obviously different from the brain of other primates, apart from being much larger and with a more convoluted surface. The general shape of a human brain, with the different lobes labeled, is shown in Figure 5.3, and some interior parts of the brain are shown in Figure 5.4. See Roth (2002) for more on the similarities and differences between our brains and those of other animals.

Dolphin brains are comparable to ours in size and even more convoluted, in a similar-sized body, but their shape is quite different. Much of the differences in brain size between different mammals can be attributed to body-size allometry. But not all of the differences; primates in general, and humans in particular, have larger brains than allometry would predict, as do dolphins (Deacon, 1997; Martin, 1998b; Roth, 2002).

The size differences are, however, not evenly distributed throughout the brain. The neocortex, and particularly the prefrontal lobes, have traditionally been regarded as a human specialization, even larger than would be predicted from a general enlargement, and constituting a larger fraction of the brain in humans than among other primates (Rilling & Insel, 1999; Deacon, 1997). But some recent studies of ape brains have failed to confirm this, and instead appear to show that all the main parts of the human brain have increased in tandem (Nishikawa, 1997; Semendeferi et al., 1997; Semendeferi & Damasio, 2000; Clark et al., 2001). On the other hand, both Barton & Harvey (2000) and Clark et al. (2001) find a neocortex enlargement instead when comparing primates with insectivores, whereas MacLeod et al. (2003) report that the cerebellum, rather than the neocortex, has increased in size in hominoids. Oxnard (2004) in a multivariate analysis finds multiple differences between the brains of humans and other primates, with both neocortex and cerebellum enlarged in humans, as well as striatum and diencephalon.

There is a notable lack of consensus, which may be due to the presence of several methodological pitfalls in this type of analysis. There are statistical problems involved in analyzing part-whole ratios, and the presence of humans as an extreme outlier in the dataset may bias the result. Furthermore, the question of homology between brain parts is a non-trivial issue, even between so close relatives as hu-

---

13 'Allometry' is about the correlations between the sizes of different body parts, or between some part and the whole. ‘How much larger brain (or heart or whatever) does twice as big an animal have? Twice as large, or more, or perhaps less?’ The answer to that question is a matter of allometry (Moore, 2000b; Rifkin, 1995; Deacon, 1997).

14 Note, however, that the methodology used by Semendeferi and colleagues cannot distinguish prefrontal cortex from the rest of the frontal lobe, and is thus not informative on the fraction of prefrontal cortex (Deacon, 2004a).

15 The use of insectivores for comparison is based on the assumption that insectivores are primitive basal mammals. This assumption is dubious, not least because ‘insectivores’ do not form a monophyletic group.
mans and apes (Deacon, 2004a), and different homology identifications may yield divergent results.

Dolphins, with human-sized brains, have a different expansion pattern, with the cerebellum expanded instead of the neocortex (Marino et al., 2000), similar to that reported by MacLeod et al. (2003) for hominoids but quite different from the apparent neocortex enlargement in humans. This shows that not all mammalian brain expansion is a matter of simple allometry. Dolphins and humans differ also in cellular and neurochemical details, with humans more resembling other primates, and dolphins resembling their artiodactyl relatives (Hof et al., 2000).

Embryologically, the amount of additional brain tissue in humans is largely controlled by the pattern and timing of fetal brain growth versus body growth, with expression patterns of homeotic genes playing a central role (Deacon, 1997; Karten, 1997). In mammals in general, the brain grows rapidly during early fetal development, and slows down later. In humans (and to some extent other primates) the rapid-growth phase is prolonged, into early childhood in the case of humans, leading to a larger brain-to-body size ratio, a growth pattern sometimes labeled as neoteny (McKinney, 1998; Penin et al., 2002). At birth, human babies are not very remarkable, with a brain size not a lot larger than that of other ape babies — but they keep growing like fetuses until their brains are four times the size of a chimp brain. If it is a straightforward matter of prolonging growth, the expected result would be a straight allometric increase — but note that this does not automatically
mean that all parts increase equally. The allometric relations may be different for different parts. In particular, neoteny would tend to favor those parts that develop late in ontogeny, such as the neocortex (Finlay et al., 2001). But detailed data on brain growth rates over time in different primates show a more complex pattern (Leigh, 2004). There is also evidence from studies of fetal gene expression (Pennisi, 2002) indicating that the difference between ape and human brains is more than just a matter of size or general growth prolongation — the pattern of gene expression in the brain, but not in other organs, differs markedly between humans and other apes.

The pattern of connections in the brain will be directly affected by the size changes. As discussed on page 31, neurons in the growing brain compete with each other for connectivity in a process of ‘neural Darwinism’. Changing the brain size relative to the body size will affect the competitive balance between the central and peripheral nervous system, increasing the internal connectivity of the brain at the expense of its external connections. Changing the relative numbers of neurons in different parts of the brain will likewise affect the patterns of neural connections, in a process of displacement, with growing brain parts gaining increased connectivity due to their numerical superiority (Deacon, 1997).

It should also be kept in mind that increasing the size of the brain poses a variety of design problems — simple scaling cannot be the whole story in any case (Kaas, 2000).
On top of the size differences, it has also been discovered recently that humans and apes share a unique new type of neurons, not found in other mammals. These spindle-shaped projection neurons are located in the anterior cingulate cortex. In humans, these cells are highly susceptible to Alzheimer and other degenerative diseases, hinting at a role in the higher cognitive functions hardest hit by these diseases (Nimchinsky et al., 1999).

But regardless of which part of the brain is significant here, any postulated selective advantages of our large brains have to be enormous, in order to offset both the high metabolic cost (Nilsson, 1999; Raichle & Gusnard, 2002) — we need something like 20-25% more food just for our brain — and the high risks associated with large-brained childbirth (Byrne, 2000). This high cost renders somewhat problematic the concept of ‘Cognitive reserve’ (Li & Hombert, 2002), spare cognitive capacity that can be recruited for new functions.\(^{16}\) We certainly start our lives with cognitive capacities that can be recruited for a variety of purposes — cf. page 112 — but why would this capacity be larger than what we’d normally use in a lifetime?

Apart from the massive quantitative growth of the brain in humans as compared with other primates, only minor qualitative differences are known. Humans have cross-connections between the thalamus and cortex that are absent in monkeys (Helmuth, 2001b) and humans lack a certain enzyme, CMP-sialic acid hydroxylase, the distribution of which in other animals indicates that it may hamper brain function (Alper, 2001).

5.3.1 Why did large brains evolve?

The reasons why larger relative brain size\(^ {17}\) evolved in primates are not totally obvious. A larger brain presumably implies improved cognition,\(^ {18}\) which in most contexts would be an advantage, but this advantage is at least partially offset by the brain being a metabolically very expensive organ to maintain (Nilsson, 1999; Dunbar, 1998; Lennie, 2003). Any hypothesis of brain evolution needs to explain both why primates have grown larger brains, and why most other mammals haven’t.

\(^{16}\) As an extreme example, the commonly heard assertion that ‘we only use 10% of our brains’ is, apart from being demonstrably wrong for many other reasons, from an evolutionary perspective completely ridiculous.

\(^{17}\) When talking about ‘relative brain size’, or ‘brain-to-body size ratio’, it would be more proper to use the more stringent but mathematically messier concept of ‘encephalization quotient’ (EQ). The EQ (no relation of the ‘emotional intelligence’ of recent fame) is an attempt to take into account how brain size normally varies with body size, and measure how large a brain an animal has compared with other animals of the same size.

\(^{18}\) The correlation between brain size and cognition (or intelligence) is not well established within *Homo sapiens* (Schoenemann et al., 2000), but is fairly clear when comparing between different species (Reader & Laland, 2002) — but see also Deacon (2003a) for some words of caution. Some measure of brain complexity may be relevant here, but such measures that exist, e.g., Tononi et al. (1994), are far from being operationally useful in this context.
As for the additional brain growth that has taken place in the human lineage, the reasons for this are intertwined with the reasons for language evolution, which will be addressed in Chapter 10. Arguments for a causal link between language and brain growth have been proposed in both directions, either with the needs of language driving the brain enlargement, or with the enlarged brain as a vital expectation for language. Deacon (1997), for example, regards the growth of the pre-frontal lobes as instrumental in the evolution of human symbolic abilities, which of course are vital for language.

Several types of hypotheses have been proposed as explanations for primate brain growth (Dunbar, 1998; Rifkin, 1995; Martin, 1998b):

**Epiphenomenal**

The large brain (or large brain parts) did not evolve for its own sake, but is a mere byproduct of the evolution of some other feature, e.g., large body size (Sacher & Staffeldt, 1974). Considering how expensive the brain is, this appears highly unlikely, but has its defenders (critically reviewed in Dunbar (1998)). Ragir (2001) suggests the idea that encephalization is a byproduct of bipedalism and the concomitant narrowing of the pelvis. But it appears rather counter-intuitive that forcing babies to be born smaller would drive the evolution of larger brains.

A new version of the epiphenomenal hypothesis emerged recently, with the discovery of a genetic difference between humans and other apes, which causes humans to have much weaker jaw muscles (Stedman et al., 2004). This is supposed to have allowed the brain to grow larger, as the skull was released from the role of buttressing heavy-duty chewing (Currie, 2004; Pennisi, 2004b). The reconstructed timing of the genetic change, 2.4 million years ago, is a fairly good match for the beginning of the fossil transition from ape-sized to human-sized brains. But there is a difference between allowing and causing — removing a constraint on brain size may have allowed its growth, but does not provide a driving force.

**For non-cognitive purposes**

The large brain evolved for some other purpose than to think with. The only serious proposal here that I am aware of is that of Falk (1990), who proposed that the brain evolved as a cooling device for our blood, much like the radiator of a car engine. It makes some sense for us to have acquired a radiator when we left the jungle for the open savannah — but why such an expensive one? It’s not even very efficient as a radiator, insulated by bone and hair. It would make more sense to do like African elephants, whose large ears perform the same function, cheaper and more efficiently.
Energy limited

The brain is an energy-expensive organ, and only animals with surplus energy could evolve larger brains. Fruit-eating primates presumably had more ‘spare’ energy than e.g., grass-eating ruminants. A variant of this hypothesis is the ‘maternal energy hypothesis’ (Martin, 1998b), in which the energy supply from the mother to the growing fetus is the limiting factor. Not very compelling either — if an organism has surplus energy, why spend it specifically on brains, unless a large brain provides some selective advantage? Fish & Lockwood (2003) do present some empirical support for a correlation between diet quality and brain size, but the direction of causation is not self-evident — see the next point below.

Another aspect of energy limitation is the reduction of the gastrointestinal tract that has taken place in the human lineage (Li & Hombert, 2002). The G.I. tract is also very energy-consuming, and may compete with the brain for energy — but the G.I. tract is of course also vital for supplying the body with energy. Its size can be reduced only if the diet is improved, with more easily digested and energy-rich food. Again, the direction of cause and effect between brain growth, diet, and G.I. tract is not obvious. A coevolutionary scenario may be more plausible, with mutual feedback between changes in brain, food, and belly. There is some fossil data on diet changes among our ancestors — see Section 10.1 — but little or no data on changes in the G.I. tract.

Environmentally driven

- Diet driven. A fruit eater needs more brain power in order to keep track of when and where ripe fruit is available, and the picking of fruit is more complex than grazing or browsing (Potts, 2004). It is, however, not obvious that fruit is harder to keep track of and catch than the prey of carnivores is. Carnivores do have fairly large brains as well, larger than most herbivores, but not as large as primates. Furthermore, far from all primates are frugivores; the group also includes e.g., gorillas, leaf-eating but still pretty smart. Variations on the diet-driven hypothesis concern e.g., the extraction of technically difficult food, or reliance on temporally and spatially highly variable food sources (Potts, 2004).

- Navigation driven. This is related to the cognitive demands of fruit eating above, but focuses in one version specifically on the demands on spatial cognition and mental map making, and in another version on the cognitive demands associated with arboreal life — climbing trees safely is not trivial for a heavy ape (Byrne, 2000). Possibly the hypothesis of Barton (2004), in which the image processing demands associated with binocular vision drives brain evolution in primates, may be placed here as well.

Patterson & Cohn (1990) report that they have administered a variety of standard IQ tests to a gorilla in their care (see page 135). She is said to have achieved test scores ranging between 70 and 95, well within the human range.
Anatomical and neurological prerequisites

**Tool driven.** The suggestion that the evolution of human intelligence was driven by tool making also belongs in this category; see Section 10.2 for more on this. On a related note, Reader & Laland (2002) find significant correlations between brain size, tool use, innovation, and social learning, across a number of primate species.

Potts (1998) regards environmentally driven brain expansion as part of a larger package of adaptations for coping with environmental **variability**, rather than adaptations for any specific lifestyle. According to Heinrich & McElreath (2003) this hypothesis receives some support from the parallel increase in brain size and climatic variability over the past 14 million years. A significant part of the package for coping with variability may be the capacity for social learning, which can be advantageous in such environments (Heinrich & McElreath, 2003; Boyd & Richerson, 1995), connecting it with the next point below. Wildgen (2004) connects variability selection with the origin of creativity and innovation in the human lineage. But on the other hand *Australopithecus afarensis* was apparently able to cope with rapid climate variability without visible effects on their brain size (Bonnefille et al., 2004).

**Socially driven**

Most primates live in complex social groups, where relations with (and manipulation of) conspecifics play a major role. Unlike most social mammals, where a rather straightforward dominance hierarchy is the rule, politics and coalition-building are important for a primate’s success (Seyfarth & Cheney, 2002), driving the evolution of social intelligence (de Waal & Tyack, 2003). This has been called the social brain hypothesis (Dunbar, 1998) or, more colorfully, ‘Machiavellian intelligence’, a label introduced by Byrne & Whiten (1988, cited in Tomasello (2000c)), though the general concept was first proposed by Alison Jolly (Small, 2000). And humans live in larger and more complex societies than other primates, so this selective force would be unusually strong for us (Flinn, 1997). This can indeed lead to a runaway ‘arms race’ of brain power, but begs the question of where the social complexity came from in the first place. An interesting parallel is the evolution of similarly large brains in dolphins and related whales, who have comparable social complexity (Connor et al., 1998; Lusseau, 2003). Dunbar (1996) has found a strong correlation in primates, not between group size and brain size as such, but between group size and neocortex fraction of the brain, which is not inconsistent with the Machiavellian hypothesis, as the neocortex is the part of the brain that has increased the most in size among primates, and particularly humans. It can also be noted that social knowledge among elephants (McComb et

---

20 But note that coalitions have been observed even among crabs (Backwell & Jennions, 2004), without any notable effect on crab brains.
al., 2001) and social networks among female baboons (Silk et al., 2003) directly translate into Darwinian fitness.

5.3.2 Lateralization

One notable aspect of language processing in the brain is its lateralization. It was noted at least as early as the 19th century, by Broca (1861) and others, that language is mainly handled by the left hemisphere of the brain (Radick, 2000a). Damage on the left side of the head commonly led to aphasia, whereas right-sided damage had little effect on language. The same lateralization pattern is observed regardless of modality, with sign language being as left-lateralized as spoken language (Hickok et al., 2002).

There is some disagreement on how absolute this left-sided language specialization is, and what effect it has on our general cognitive faculties. Lately, some evidence has surfaced for a right-hemisphere role in some aspects of language, notably prosody, but also phonology (Simos et al., 1997), word learning (Sabbagh, 1999) and others (Bonvillian et al., 1997; Locke, 1997; Richards & Chiarello, 1997; Copland et al., 2001; Poeppel & Hickok, 2004). Syntax, however, does appear to be the near-exclusive domain of the left hemisphere (Grodzinsky, 2000), even though Vigliocco (2000) presents neuroimaging studies where the right hemisphere is activated in syntax processing. And language is not totally absent (though it is severely impaired) even in patients whose left hemisphere has been completely removed (Code, 1997).

If one hemisphere is damaged very early in life, well before language acquisition, nearly normal language development can still follow, regardless of which hemisphere is damaged (Feldman et al., 2002). If the left hemisphere is damaged early enough, or even if it is entirely removed, language can still be acquired normally, but handled largely by the right hemisphere (Müller et al., 1998; Stowe & Haverkort, 2003), showing that the laterality of language is not strictly hardwired, but is handled with considerable plasticity (Mueller, 1996). Some plasticity may remain as late as age eleven, as shown by the left-hemispherectomy case presented by Immordino-Yang (2004). Nevertheless, the left hemisphere does dominate language perception and production, in the vast majority of people.

The brain is by no means unique in its lateral specialization. Various other organs in the body have a distinct left-right asymmetry, and the asymmetry is initiated very early in fetal development (Izpisua Belmonte, 1999), by molecular (Ramsdell & Yost, 1998; Levin & Mercola, 1998; Boorman & Shimeld, 2002) and genetic (Isaac et al., 1997) mechanisms that are beginning to be understood. Left-right asymmetry also has an ancient evolutionary history (Boorman & Shimeld, 2002).

The right hemisphere has other functions, such as spatial and synthetic, ‘holistic’ (as opposed to analytic) thinking (Walsh, 2000). The latter aspect has led to a large New-Age-flavored literature surrounding right-hemispheric cognition, of dubious scientific value.

21The right hemisphere has other functions, such as spatial and synthetic, ‘holistic’ (as opposed to analytic) thinking (Walsh, 2000). The latter aspect has led to a large New-Age-flavored literature surrounding right-hemispheric cognition, of dubious scientific value.
likewise, language is far from the only brain activity that is laterally localized. The vast majority of humans preferentially use one hand rather than the other for precision work, and for most of us this is the right hand (and thus the left hemisphere of the brain, since it’s cross-connected). This asymmetric pattern of behaviour can be observed very early in fetal development (Hepper et al., 1998). Natsopoulos et al. (1998) argue for a connection between handedness and language lateralization, based on evidence for weaker language development in left-handed children, but this conclusion is disputed by Annett (1998). This issue is further discussed by Halpern (1996), Natsopoulos et al. (2002), and Ross & Bever (2004), but without clear resolution.

during the acquisition of sign language, there is a marked preference for using the right hand for signing. this might be partially explained by the general preference for using the right hand for tasks requiring precision, but Bonvillian et al. (1997) found that right-hand dominance was even stronger for linguistic signing than for non-linguistic tasks.

concerning the asymmetry of speech processing, it is interesting to note that many other animals also process species-specific communications preferentially on one side of the brain (Bisazza et al., 1998). Songbirds do (Cynx et al., 1992; Mueller, 1996), though they are so phylogenetically distant from us that this can hardly be relevant. That the same kind of asymmetry is found in monkeys (Hauser et al., 1998; Locke, 1997; Hauser & Andersson, 1994; Bisazza et al., 1998; Ghazanfar & Hauser, 1999; Hauser & Akre, 2001) is more interesting, and even more so the signs of language-related lateralization in chimpanzees (Hopkins & Leavens, 1998; Hopkins et al., 1991), though the evidence here is ambiguous (Hopkins et al., 1992; Hopkins & Morris, 1989). Particularly interesting is the observation (Neergaard, 1998) that bonobos use almost exclusively their right hands for gesturing. It should be noted that the asymmetry observed in monkeys is not a matter of processing all sounds in the left hemisphere — sounds that are not recognized as conspecific vocalizations produce the opposite pattern (Ghazanfar et al., 2001).

there are anatomical asymmetries in the human brain, with enlargement on the left side in locations corresponding to the classical ‘language areas’, and asymmetries at the neuron level have recently been reported as well (Hutsler, 2003), with very large pyramidal neurons more common in the language areas than in the corresponding areas on the right side of the brain. These large neurons are thought to be involved in the long-range connectivity between different sectors of the brain.

Corresponding anatomical asymmetries have been observed also in chimpanzees (Gannon et al., 1998; Neergaard, 1998), as well as other apes (Hopkins et al., 2000).
5.3.3 Modularity

Paul Broca (1861) is famous in the history of neurolinguistics not only for discovering the laterality of language processing mentioned in the previous section, but also for pinpointing a specific part of the brain, now known as Broca’s area, where injuries would impair language. Wernicke (1874) and others found more such apparently language-specific brain pieces, which led to the conjecture that language (and for that matter other brain functions as well) is handled by dedicated brain modules, so that each section of the brain has a specific purpose. For low-level input and output functions, the modularity is well established; there is ample evidence that e.g., hearing is handled by one dedicated part of the brain, and vision by another dedicated part. What remains controversial is the modularity of higher-level functions like language, and also the properties of the putative modules and to what extent they are innate. It is, however, quite clear that the anatomical organization of language is nowhere near as simple as in the traditional Broca/Wernicke picture (Poeppel & Hickok, 2004).

This anatomical modularity is closely related to, but distinct from, the functional modularity hypothesis of language organization, in which the language faculty (as well as the rest of the mind) is composed of several autonomous computational modules. Modularity in this sense is a hypothesis developed by Fodor (1983), a highly influential book despite being “notably uninformed by evolutionary thought” according to Lloyd (1999, p. 220). Modules in the Fodorian sense are characterized by information encapsulation and domain specificity, and typically contain innate databases pertaining to the specific domain (Mameli, 2001).

Chomsky (e.g. 1988), has supported a related modularity hypothesis, but this is less evident in his most recent works (Uriagereka, 1999; Stemmer, 1999).

The hypothesized modularity of language is relevant in the context of language evolution, because an autonomous, encapsulated, monolithic language module must have an evolutionary history that is quite different from (and more difficult to trace than) that of language abilities based on a combination of pre-existing general cognitive systems; cf. the evolvability of language in the generative vs. cognitive perspectives, discussed briefly at the end of Chapter 2.

Language is not the only aspect of brain function for which modularity is proposed — e.g., Duchaine et al. (2001) and Tooby & Cosmides (1992, cited in Buller & Hardcastle (2000)) argue for a system with separate innate brain modules for everything from sex to semantics, hundreds or even thousands of modules in total.

A functional module may, but need not, correspond to a distinct piece of the brain, an anatomical module. Chomsky emphasizes the logic with little interest in the hardware. Fodor is commonly interpreted as equating functional and anatomical modularity, but according to Coltheart (1999), Fodor explicitly acknowledges the possibility of modules being distributed, or arising in an equipotential brain.
The scope of the proposed modules is not well-defined. The language faculty as a whole is sometimes regarded as a single module, but there is also much talk of smaller submodules that are supposed to handle specific components of language, such as syntax or phonology. Levy (1996) introduces such a distinction between ‘Big Modularity’ and ‘small modularity’, and compares the empirical evidence for different-sized modules.

The hypothesized ‘Language Acquisition Device’ of children would be a Big Module in this sense. It should be noted that it is not self-evident (though often taken for granted) that this would be the same module(s) as adults use for normal language processing. The sensory and processing demands in acquisition, are quite different from those in routine use of a known language (Bates, 1993).

Furthermore, there is some evidence from lesion studies that language acquisition is affected by a pattern of lesions quite distinct from the lesions of the left temporal area that typically cause adult aphasia (Plunkett, 1997). Notably right-hemisphere damage in children, unlike adults, may affect language acquisition (Selnes, 2000; Curtiss et al., 2001), and Sabbagh (1999) proceeds to propose a specific right-hemisphere role in acquisition. On the other hand, Holowka & Petitto (2002) found that the prelinguistic babbling of infants shows signs of left hemisphere lateralization.

There is evidence of such a dissociation between acquisition and mature competence also in Williams syndrome; see page 105 below. Neuroimaging studies also show that children and adults process language differently, even after differences in performance have been taken into account (Schlaggar et al., 2002).

It is sometimes erroneously assumed, e.g., by Cosmides (1989, cited in Lloyd (1999)), that modularity and innateness are necessary postulates for any evolutionary explanation of any cognitive system, including language. But the connection between modularity and evolution is tenuous at best (Lloyd, 1999; Mameli, 2001; Shapiro & Epstein, 1998; Atkinson & Wheeler, 2004), and as shown by Cummins & Cummins (1999), much weaker postulates are sufficient for evolutionary explanations to be tenable. Actually, the opposite argument can be made — evolutionary processes have a penchant for re-using preexisting equipment, instead of building something new from scratch (see page 16), and applying this to cognition we’d expect to find that:

...natural selection had recruited extant cognitive capacities for new purposes rather than going to the trouble of developing new capacities every time a novel problem comes along. (Shapiro & Epstein, 1998, p. 176).

Either a modular or a distributed language faculty could evolve, but innateness is required only in the weaker form of predispositions and developmental biases and canalization, and the argument of Shapiro & Epstein (1998) against strict modularity is fairly compelling.

Furthermore, as argued by Karmiloff-Smith (1994) and Buller & Hardcastle (2000), modules need not be innate and genetically pre-programmed. They propose instead a gradual process of modularization as the infant mind develops, in
which environmental input is instrumental in shaping the brain from its original plastic state.

Empirical evidence that is relevant to the issue of language modularity includes answers to the following questions:

- **Normal child development.** What correlations are there between language acquisition and other cognitive development?
- **Developmental defects and delays in children.** Are there any defects that exclusively affect language, or are language acquisition problems always accompanied by non-language deficits?
- **Brain lesions in adults.** Same questions as for the developmental problems: do any brain lesions damage nothing but language?
- **Brain activity.** Do brain scans show any areas of the brain that are used for language processing and nothing else? Does everybody use the same brain areas for language? How about other language modalities?

A rich array of experimental and clinical evidence surrounding these issues is available, some of which is reviewed below:

**Normal child development**

What correlations are there between language acquisition and other cognitive development?

According to Bates (1993), such correlations have been observed between major steps in language acquisition, and the acquisition of various non-linguistic behaviours, mostly related to communicative and social skills; see Table 5.1.

There are some intriguing similarities between the language and non-language columns in Table 5.1, suggestive of a common underlying mechanism. This would argue against any language acquisition device being strictly language-exclusive.

### Table 5.1. Ontogenetic stages in linguistic and non-linguistic cognitive and social development.

<table>
<thead>
<tr>
<th>Age</th>
<th>Language</th>
<th>Non-language</th>
</tr>
</thead>
<tbody>
<tr>
<td>8-10 mo</td>
<td>Word comprehension</td>
<td>Communicative gestures and routines, causal analysis, tool use</td>
</tr>
<tr>
<td>12-13 mo</td>
<td>Naming</td>
<td>Gestures symbolizing an action or object</td>
</tr>
<tr>
<td>16-20 mo</td>
<td>Multi-word utterances</td>
<td>Multi-action sequencing and planning</td>
</tr>
<tr>
<td>20-30 mo</td>
<td>Grammar</td>
<td>Scripts in playing</td>
</tr>
</tbody>
</table>
Also, as noted by Levy & Kave (1999), “normal language development is seen as a major predictor of neurological intactness in children” (p. 115). This would hardly be the case if language acquisition were totally separate.

Developmental defects and delays in children

Are there any defects that exclusively affect language, or are language acquisition problems always accompanied by non-language deficits? Conversely, are there defects that affect general cognition, but leaves language unscathed? Are there different defects that affect different subcomponents of language? Are any of these defects caused by any identifiable ‘language genes’?

‘Specific language impairment’ (SLI) is a label used for children with significant discrepancies between their linguistic and non-linguistic development, without any obvious reason such as hearing problems (Levy & Kave, 1999). It is, however, a rather vague diagnosis for a heterogeneous group of children, and it is not well established that SLI is a single well-defined disorder (Bishop, 1994). It is never a matter of language being completely lost with other cognitive functions spared (Pinker, 2003). Instead, there may well exist several different disorders of language, as listed in Joanisse & Seidenberg (1998) or van der Lely & Stollwerck (1997), with different causes. Furthermore, a significant fraction of children with an SLI diagnosis do turn out to have general cognitive problems, which only happened to be most apparent in the language sector (Bishop, 1994). Several specific non-language problems commonly associated with SLI are also known (Joanisse & Seidenberg, 1998), notably an impairment of timing and rhythmic ability (Alcock et al., 2000), that might be interpreted in the context of the sequencing hypothesis of Calvin (1993) discussed on page 83.

Grammatical SLI appears to be the best established variety (Levy & Kave, 1999; van der Lely & Stollwerck, 1997), and a considerable amount of work (reviewed by Levy & Kave (1999)) has been done concerning the precise deficits of grammatical-SLI-children. Much of the work is done in terms of the absence of specific structural grammatical principles, such as Binding and coreference as proposed by van der Lely & Stollwerck (1997), RDDR (Representational Deficit for Dependent Relations) as proposed by van der Lely & Battell (2003), or Procrastinate as proposed by Davies (2002)). On the other hand, Tomblin & Pandich (1999, but see also responses by van der Lely (1999) and Marcus (1999a)) argue that the children with grammatical SLI just populate the extreme tails of the normal distribution of language abilities. And explanations for ‘grammatical’ SLI that are based on processing deficits rather than grammar per se may remain tenable (Joanisse & Seidenberg, 1998; Levy & Kave, 1999), though Gopnik & Goad (1997) rule out low general IQ as well as auditory and articulatory processing deficits as explanations. And SLI is rarely if ever characterized by a total absence of grammar, as would be the case if a grammar module were totally non-functional — rather, the SLI children are aware that syntactical rules exist, and apply many of them cor-
rectly, but miss a few. If SLI means that a module is knocked out, it must be a very small and highly specific module. It appears premature to conclude, as van der Lely & Stollwerck (1997) do, that “[t]he data provide additional evidence for an innate syntactic module...” (p. 283, emphasis added).

There exists a fair amount of evidence for grammatical SLI having a genetic cause (van der Lely & Stollwerck, 1996; Stromswold, 2001), including twin studies (Palmour, 1997; Gopnik, 1997) as well as numerous family studies (Gopnik & Goad, 1997; Palmour, 1997), but identifying individual genes associated with complex traits like language is a non-trivial problem (Felsenfeld, 2002). Pre- and perinatal problems (which commonly cause general neurological problems, with associated language deficits) are not a significant cause of SLI (Bishop, 1997), further supporting a genetic basis.

Recently a specific gene, known as FOXP2, has been identified as defective both in a family with multiple SLI sufferers, and in an unrelated boy with SLI (Lai et al., 2001). The gene is involved in the regulation of embryonic brain development, and may be related to observed defects in the basal ganglia (Balter, 2001a), specifically in the caudate nucleus (Watkins et al., 2002), of the brain, as well as underactivation of Broca’s area in brain scans (Liégeois et al., 2003). FOXP2 belongs to a family of regulatory genes, and is active in the fetal brain, among other places, so presumably it is regulating some aspect of neural growth and development, controlling the expression of other genes (Marcus & Fisher, 2003). It is, however, unclear exactly what this gene does (Lai et al., 2001; Balter, 2001a; Pinker, 2001), and it does not appear to lie behind many other cases of SLI (Newbury et al., 2002) or less severe language impairment (Meaburn et al., 2002).

Comparison with the corresponding gene in other mammals shows that the FOXP2 gene is identical between chimps, gorillas and monkeys, and nearly identical even between chimps and mice. The human version, on the other hand, contains two functional changes that appear to have been strongly selected in the human lineage (Enard et al., 2002), possibly quite recently in our evolution, around the time Homo sapiens first appeared (Marcus & Fisher, 2003; Balter, 2002b).

Gardner (1991) argues for the existence of a language gene on chromosome 15, but he bases his argument on syndromes that produce a wide variety of behavioral anomalies, of which language deficit is one. Thus, his finds are not compelling evidence of language-specific genes. Two other loci associated with SLI have been reported (Pinker, 2003).

In general we should be careful about concluding that a gene is ‘for’ some specific trait, unless both the causal link from the gene to the trait, and the evolutionary link from the trait to the gene, have been established (Kaplan & Pigliucci, 2001), neither of which is the case for any proposed ‘language gene’, not even the widely publicized FOXP2. The meta-analysis of Stromswold (2001), reviewing more than 100 studies, indicates significant heritability of both normal language abilities and
language defects, but concludes that “[t]he relationship between the genotype and phenotype (linguistic ability) is not simple, however.” (2001, p 705).

Interestingly enough, Palmour (1997) reports a specific anatomical abnormality in the structure and asymmetry of the brain of many SLI patients, in the perisylvian fissure (in the vicinity of the traditional ‘language areas’ of the brain), though, as noted above, abnormalities in the basal ganglia have also been associated with SLI.

In the opposite direction from SLI, there are several syndromes that cause severe general retardation but with little effect on language, such as the Cocktail Party Syndrome, the Savants\textsuperscript{22} (Levy & Kave, 1999), Spina Bifida, and Williams Syndrome (Pinker, 1995).

Williams Syndrome (Williams et al., 1961), reviewed in Bartke & Siegmüller (2004), is a genetic disorder where a contiguous sequence of genes along a chromosome has been deleted (Meng et al., 1998) with consequent malformations in numerous places, notably the posterior cortex, with the frontal lobes and cerebellum spared (Deacon, 1997). It has attracted particular attention because the dissociation between linguistic and other skills appeared clear (Clahsen & Almazan, 1998; Deacon, 1997; Pinker, 1999). But recent investigations have revealed some subtle deficits of language also in Williams-patients whose language is superficially normal (Karmiloff-Smith et al., 1998; Grant et al., 2002; Reilly et al., 2004; Levy, 2004). Clahsen & Almazan (1998) interpret the pattern of language deficit in Williams-patients as stemming from difficulties with lexical access, with an intact syntactic-computational system, in contrast with the opposite pattern in SLI patients,\textsuperscript{23} whereas the deficits that Grant et al. (2002) identify are clearly syntactic. It is thus hardly tenable to interpret Williams Syndrome as the result of an intact language module with some other cognitive modules knocked out. Instead, this kind of pattern with subtle deficits in several areas is more reasonable if the affected functions are distributed in the brain, but with different parts of the brain contributing unequally to different aspects of cognition. Deacon (1997) interprets Williams Syndrome along these lines, in the context of his model where the prefrontal lobes are important for combinatorial and attentional aspects of cognition, leading to the relatively spared capacity for symbolic processing in Williams Syndrome, with its mainly posterior malformations.

Paterson et al. (1999) find that, while Williams patients do well enough on linguistic tests in adulthood, their language acquisition is significantly delayed, indicating a dissociation between the brain machinery used for acquisition and for mature language processing. Paterson et al. (1999) also report that the Williams patients, oddly enough, display the opposite pattern for mathematics, with good acquisition but poor performance in adulthood. To the extent that mathematics is

\textsuperscript{22}But note that many Savants also have symptoms of autism, which does have associated language deficits (Frith, 1997).

\textsuperscript{23}This is suggestively similar to the opposite loss patterns in agrammatic and anomic aphasia.
a matter of symbolic processing, one might have expected a pattern similar to that of language.

**Brain lesions in adults**

Same questions as for the developmental problems: do any brain lesions damage nothing but language?

Broca’s aphasia is the classic example of a brain lesion that mainly damages language. There are many other kinds of aphasia, but Broca’s is noteworthy in that grammar is severely impaired, but other cognitive functions are largely spared — unlike many other aphasics, Broca patients retain their previous IQ (Grodzinsky, 2000). And Broca’s aphasia is commonly associated with damage in the vicinity of Broca’s area in the brain, which used to be taken as evidence of Broca’s area being a grammar module. But more detailed studies show that the situation is more complex (Deacon, 1997; Mueller, 1996; Lieberman, 2002; Deacon, 2004a). Other areas of the brain, notably the basal ganglia, are implicated in Broca’s aphasia, as well as, or possibly even instead of, Broca’s area (Lieberman, 2002).

And the pattern of grammar loss in Broca’s aphasia (or agrammatic aphasia in general) does not lend itself to any straightforward explanation in terms of a disabled grammar module, which ought to affect all modes of language equally. Instead, there is a pattern of partial aphasia in all possible dimensions, with grammar lost for one aspect but not another, different for different patients (Mueller, 1996):

- Production vs. reception
- Oral vs. written
- Oral vs. sign language
- Different languages (for bilinguals)

There is also no really clean dissociation between Broca’s aphasia and other types, with both word-finding deficits and subtle grammatical deficits present in varying proportions in all different forms of aphasia (Bates, 2003).

Furthermore, parts of the syntactic loss can be compensated using heuristic strategies, partially restoring comprehension, in patterns that may mask the pattern of the original aphasia. Training can restore language performance in some but not all aphasia patients (Weinreich et al., 2001).

There is also some evidence of non-language abilities residing in Broca’s area, notably some motor functions (Binkofski & Buccino, 2004). Verbal oral apraxia is commonly associated with Broca’s aphasia. Sequencing was discussed on page 84, and may be implicated in the motor control deficit observed in some patients.

---

24 But it is not obvious how IQ is reliably measured in aphasia patients, or how IQ loss is distinguished from comprehension loss.

25 A heuristic strategy here is basically a matter of applying a combination of semantic knowledge and normal (non-language) intelligence to sentence analysis.
Anatomical and neurological prerequisites

(Vigliocco, 2000). Another interesting discovery is that damage to Broca’s area impairs the ability to handle non-verbal hierarchical structures (Greenfield, 1993, cited in Savage-Rumbaugh & Lewin (1994)) — highly significant in view of the hierarchical nature of grammar. If Broca’s area contained a general hierarchy handler, this would explain much of Broca’s aphasia. Non-language deficits are not commonly reported for Broca patients — though they were present even in Broca’s original patient (Lieberman, 2002) — but Schweiger & Brown (2000) may have a point:

The apparent uniqueness and dissociability of language from other skills, and specifically grammatical deficits, reported in the aphasia literature, reflect the narrow focus of researchers rather than the scope of symptoms presented by the patients. (p. 214).

Just as for the SLI children, there are two classes of hypotheses concerning the deficits of agrammatics (Levy & Kave, 1999; Crain et al., 2001):

- Structural deficits, where some specific principle of grammar is disabled. The ‘Trace Deletion hypothesis’ is a typical example, in which the trace that is left after movement (in a transformational-grammar framework) is lost, so that the patient cannot reconstruct the pre-movement position (and thus theta-role) of a moved noun phrase. Further structural hypotheses are reviewed in Levy & Kave (1999), Berwick (1997), and Grodzinsky (2000).

- Processing deficits, where syntactic knowledge as such is unharmed, but where computational or working-memory resources are diminished, so that processing limitations are exceeded even in fairly short sentences26 (just like it is quite possible to construct a sentence with fifty nested embedded clauses which is perfectly grammatical but which cannot be parsed by any normal person without paper and pencil, because of processing limitations). Miera & Cuetos (1998) test one processing hypothesis, by measuring the working memory of agrammatics, comparing it with that of anomics and normal controls. They find that both groups of aphasics have poorer memory than controls, but also that the performance of agrammatics is not significantly affected by sentence length, from which they conclude that a working-memory deficit is not sufficient to explain agrammatism. Both Hartsuiker et al. (1999) and Crain et al. (2001) on the other hand do find evidence of a shortage of processing resources in Broca’s aphasia. Processing hypotheses thus remain tenable (Levy & Kave, 1999; Stowe & Havercort, 2003), though the normal nonverbal intelligence of Broca-patients argues against any significant general computational deficits.

The conclusion of Levy & Kave (1999) is that we have insufficient evidence to choose between structural and processing hypotheses. Structural hypotheses have an inherent theoretical elegance, appealing to many linguists, but that in itself is

26 Cf. the hypothesis of Wynn & Coolidge (2004) that the main cognitive difference between us and Neanderthals is precisely this kind of working memory; see Gruber (2002) for more on the possible role of working memory in human evolution.
no guarantee of truth\textsuperscript{27} — the messier, but possibly more psychologically realistic, processing hypotheses cannot be discounted. Beretta et al. (2001) present an empirical comparison of a processing hypothesis with several structural hypotheses,\textsuperscript{28} claiming to find solid evidence against the structural hypotheses.

Here again, no clear case for the modularity of language can be made. Only agrammatism has been discussed at length here, but the other varieties of aphasia do not change this non-conclusion. There is, for example, some evidence of bilinguals with selective aphasia in one of their languages (Gomez-Tortosa et al., 1996; Daroff, 1998), but this remains controversial\textsuperscript{29} (Paradis, 1996). But even if it were true that different languages occupy different modules, this would tell us little about central language modules for syntax and grammar, which, according to the Universal Grammar hypothesis, the two languages would have in common.

An interesting result is that of Caramazza et al. (2000), where it is shown that consonant and vowel production can be affected separately in aphasia, indicating that they are processed by distinct neural mechanisms. Another is that reported in Pinker (1997) concerning selective aphasias affecting either regular or irregular verbs. Both of these results support the existence of small and highly specific modules. But the complex patterns of deficits exhibited by most aphasia patients, commonly affecting what linguistic theory regards as separate components (Blumstein & Milberg, 2000) does not indicate any simple modular structure. The existence of Big Modules is not supported by aphasia data.

\textit{Brain activity}

Do brain scans show any areas of the brain that are used for language processing and nothing else? Does everybody use the same brain areas for language? How about other language modalities?

The development of non-invasive high-resolution techniques for studying the activities of living conscious brains has led to a flood of data. Unfortunately, it is not trivial to interpret these data, as it is very difficult to ascertain that a brain is doing one thing and one thing only — no matter how hard you concentrate on the experimental task, extraneous thoughts will be flitting through your mind every now and then, adding noise to the scan results. Experiments are necessarily

\textsuperscript{27}A parallel can be made with the issue of theoretical elegance within physics. It is a historical fact that physical theories that have the same kind of inherent theoretical elegance, have indeed turned out to be true much more often than might reasonably be expected. Does this tell us something fundamental about the universe — or about our sense of elegance? In any case, the success of elegant theories has been so striking that elegance has become a major heuristic criterion in recent theoretical physics — see for example \textit{The elegant universe} by Greene (2000). This issue is critically reviewed by Kosso (1999). In linguistics, a similar quest for elegance is quite explicit in Chomsky (1995), according to the review of Uriagereka (1999).

\textsuperscript{28}Note that Beretta et al. (2001) use the label ‘structural’ for what is here called ‘processing’, and ‘linear’ for what is here called ‘structural’. This discrepancy is due to their labeling hypotheses after what capacities remain, instead of naming hypotheses after what is damaged.

\textsuperscript{29}It may be noted that MRI scans of bilinguals show that in early bilinguals, the two languages share the same brain structures, but not in late bilinguals (Kim et al., 1997).
comparative, measuring the difference in brain activity between the task of interest and some control task.

Despite these difficulties, some interesting results have been obtained. To begin with, the auditory cortex contains an area specifically sensitive to human voices (Belin et al., 2000), which is the first step towards decoding speech.

Lounasmaa et al. (1996) present results from picture naming exercises, where a clear temporal progression could be seen from the visual cortex at the back of the head, towards more frontal areas, ending up in Broca’s area which was activated when the actual naming took place (passive viewing of pictures without naming them caused very little Broca activity). It can be noted that Lounasmaa et al. (1996) observed only a modest difference between the left and right hemisphere, in contrast both with other studies of word generation such as Klein et al. (1995), and with the general evidence of language lateralization reviewed in Section 5.3.2. In a similar exercise (Caramazza, 1996b) it was found that the semantic system used for recognizing picture or word stimuli appears to be distributed through large areas of the left hemisphere. On the other hand, Thompson-Schill et al. (1997) report that a number of experiments have identified a certain piece of the brain (the left inferior frontal gyrus (IFG), in the vicinity of Broca’s area) as consistently activated in semantic tasks, but Thompson-Schill et al. (1997) go on to argue that the IFG is not a semantics module, but instead a selection module, activated when choosing between semantic alternatives. On the other hand, Klein et al. (1995) find IFG activation also in purely phonological rhyming tasks. Price (1998) and Klein et al. (1995) have also found brain activation in additional areas in a variety of semantic tasks.

Lexical knowledge can with some modest degree of confidence be pinpointed to the left temporal lobe of the brain, apparently with different semantic categories of words stored separately at some level, with one small piece of brain handling animal names, another small piece of brain handling words for tools, and so on (Damasio et al., 1996). This is also indicated by the existence of similarly selective aphasia (Caramazza, 1996a). However, Martin & Chao (2001) present evidence that the neural organization of semantic knowledge is more complex and distributed in the brain. The results of Damasio et al. (2004) may clear up the picture, with their finding that lexical and conceptual knowledge are stored separately, with lexical knowledge in the left temporal lobe and conceptual knowledge in the right hemisphere, both sorted by category, supported by both imaging and lesion studies. Concept and word knowledge may have not have been cleanly separated in earlier studies.

In speech processing studies, there is evidence of interaction between phonetic and lexical processing (Bowers & Davis, 2004), and between lexical processing and world knowledge (Hagoort et al., 2004), which is strong evidence against the informational encapsulation postulated in strict modularity.
Neuroimaging studies of syntax processing have been rare, but are now accumulating at a rapid rate. Syntax encoding in speech production is located by Inefrey et al. (2001) to an area adjacent to ‘classical’ Broca, the Rolandic operculum. In a study of syntax parsing, Embick et al. (2000) find that sentences with syntax errors (with spelling errors used as control) specifically activates Broca’s area. Broca is activated also in the processing of grammatical information about words, such as word category and gender, according to Heim et al. (2003), who conclude that Broca is the common denominator in syntax processing, even though specific tasks may be partially handled elsewhere. Dogil et al. (2004) reach the same conclusion. Vigliocco (2000) reviews several studies, most of which point towards the vicinity of Broca’s area in syntax tasks, and slightly forward of Broca for semantic tasks. These results strengthen the case for Broca as a syntax module, though the non-linguistic deficits in Broca’s aphasia mentioned on page 107, as well as the use of Broca in some sequencing tasks, raises the question of whether such a module is really language-specific. Müller & Basho (2004) propose that the key role of Broca in language is due to its “afferent convergence of audio-visual-motor processing streams, ...” (p. 334), giving it a role in many multimodal coordination tasks of which language is one.

In one very interesting study, Musso et al. (2003) had adult humans learn new languages while having their brains imaged. With real languages, Broca was active, but with fake languages with non-hierarchical grammatical rules (violating the Universal Grammar of linguistic theory), Broca wasn’t used. Musso et al. (2003) take this as evidence that Broca is “specialized for the acquisition and processing of hierarchical (...) structures, ...” (p. 778); cf. page 107. In a similar study, Petersson et al. (2004) found that syntax violations in an artificial language activated Broca.

Iacoboni et al. (1999) report activation of Broca’s area, as well as its homolog in monkeys, in imitation tasks; cf. Section 9.6.2.

In general, meta-analyses of brain scan experiments have shown that the same piece of brain is generally used in many different tasks, and any given task typically activates several disjoint parts of the brain (Stowe & HaVERkort, 2003; Lloyd, 2000), not supporting any simple modular model.

Modularity conclusions

In conclusion, there is on one hand ample evidence for the existence of small specific modules, but it is difficult to pinpoint the exact purpose of any given module. On the other hand, there is no sign of any monolithic Big Modules encapsulating any major brain functions, apart from low-level sensory processing. Specifically, language appears to be handled by a large number of separate subsystems, distributed over a significant fraction of the brain (Blumstein & Milberg, 2000; Poeppel & Hickok, 2004) — as Lieberman (2002) bluntly puts it “The traditional theory equating the brain bases of language with Broca’s and Wernicke’s neocortical ar-
Anatomical and neurological prerequisites

It is also difficult to find a clear correspondence between the functional subsystems identified in the brain, and the logical components of the language faculty identified in linguistic theory (Poeppel & Hickok, 2004). Some of the systems may be language-specific but others may not — not all aspects of language need to be handled in the same manner (Bloom & Markson, 2001) — and it is difficult to find compelling evidence either way. Corina (1999) appears correct in his assessment: “All higher cognitive systems (i.e., language, [...] etc.) exhibit a range of domain-specific and domain-general processes.” (p 231). The right question to ask may not be whether our brains are modular, but rather how much modularity there is (Atkinson & Wheeler, 2004). But Atkinson & Wheeler go on to show that even that question is tricky: “there is no way of establishing a principled and robust distinction between domain-specific and domain-general features.” (2004, p. 147).

Plasticity

Another problem for the concept of innate modules is the fact that most of the brain displays considerable plasticity in development. The same part of the brain can be recruited for totally different purposes depending on what sensory input the growing brain receives at a sensitive age (Wong, 1995; Buller & Hardcastle, 2000).

Experimental re-routing of the optical nerve to the auditory cortex in new-born ferrets, results in the auditory cortex being transformed into visual cortex, and the signals coming to it being perceived as vision and not sound (Merzenich, 1999; Swindale, 2000). Transplanting a piece of brain works just as well, with the transplanted piece forming connections appropriate for its new site (Deacon, 2000). Even without surgery, substantial changes in the layout of the auditory cortex can be achieved by sound exposure during a critical period — additional cortex is assigned to the type of sounds heard (Nakahara et al., 2004), which may have some implications for the importance of infant exposure to language. The molecular-level mechanisms behind one such case of input-driven neural wiring is reviewed by Fox (1999), with more studies of neural plasticity at different levels reviewed in Sur et al. (2002), King (2002), and Schnupp & Kacelnik (2002). In humans there is similar but less drastic evidence of developmental plasticity, mainly in response to sensory deficits (Dufour & Gerard, 2000).

This plasticity in the face of damage should not be regarded as primarily a repair mechanism — instead, plasticity is a major factor in normal ontogeny, with the neural connections in the growing brain shaped and pruned depending on patterns of use (Deacon, 2000), in the ‘neural Darwinism’ process described on page 31. The outcome of this process in the adult brain is fairly uniform between individuals, but this is most likely not due to pre-programming, but simply due to similar patterns of use in different individuals of the same species.

In embryology, there is good evidence for large-scale innate patterning of the brain, with the major sections and basic wiring laid down using genetically deter-
mined molecular cues (Bishop et al., 2000; Krubitzer & Huffman, 2000; Deacon, 2000; Fukuchi-Shimogori & Grove, 2001; Rakic, 2001; Ragsdale & Grove, 2001; Monuki & Walsh, 2001), some of which are functionally identical across all animals studied, being interchangeable even between mammals and insects (Hirth & Reichert, 1999). There is, however, little sign of genetic determination of the fine structure needed for more specific modules — active recruitment of neurons for different tasks during a critical period of plasticity appears more plausible, given what we know about neural patterning in the brain (Berardi et al., 2000; Knudsen et al., 2000), contra Marcus (2004a). This way we can build a vastly more complex brain than our relatives, without significant amounts of additional genetic information — cf. the ‘poverty of the genes’ argument on page 186. The information is instead constructed during ontogeny, ‘rediscovered’ by each new generation. Quantitative changes in the large scale patterning that is under genetic control may be enough to bias the information construction, so that species-specific functions are consistently reconstructed (Deacon, 2000).

Furthermore, it is obvious that a large fraction of our cognitive capacity is not located in encapsulated genetically determined task-specific modules — such genetically determined modules can only have evolved for tasks that were relevant for fitness in the EEA where we evolved, but today we have no problem recruiting cognitive capacity in our brains to tasks like flying fighter jets or analyzing quark-antiquark annihilation processes, tasks that brain modules cannot possibly have evolved for. This demonstrates a cognitive versatility that is a strong argument against having any large fraction of the brain devoted to encapsulated ‘Big Modules’, and an argument for the possibility of task-switching even for those smaller modules that we quite possibly do have:

Exaptation of intelligence does appear routine for at least modern humans, and it seems safer to leave the issue of modularity to empirical enquiry (...) rather than assuming that cognitive mechanisms are immune from recruitment to other functions. (Byrne, 2000, p. 556)

5.3.4 Brain and language in other species, fossil and extant

Detailed information on the history of human brain anatomy is difficult to extract from fossils; what is available is largely limited to the gross size and shape of the inside of the skull, giving a rough estimate of the size and shape of the brain. At best, endocasts, and impressions from brain features on the inside of the skull, give some insights to the surface anatomy of the brain (Rilling & Insel, 1999; Brandt, 1993). Sometimes the blood supply to the brain can also be deduced from

30EEA = Environment of Evolutionary Adaptedness; see page 194.
31An endocast is a ‘petrified brain’, either a natural copy of the brain formed when the braincase is filled with sediment after the brain itself has rotted away, or an artificial mold of the interior shape of the skull. The most famous proto-human endocast belongs to the australopithecine ‘Taung baby’ (Dart, 1925).
the size of blood-vessel conduits through the skulls of fossils (Braga & Boesch, 1997), but few firm conclusions can be drawn from this.

The overall size of the brain does not show any significant increase until fairly late in human evolution. Australopithecines have brains comparable in size to those of chimpanzees or gorillas. Australopithecines, at least the gracile varieties, were, however, quite small creatures (around 30-50 kg), slightly smaller than modern chimpanzees, and much smaller than gorillas, which implies a somewhat larger brain-to-body-size ratio (Kappelman, 1996; Ragir, 2001). On the other hand, there is little indication of any humanlike features on australopithecine endocasts according to Brandt (1993), whereas Culotta (1999a) and Gibbons (2002c) both quote Dean Falk as claiming the presence of a humanlike brain shape in A. africanus. Holloway et al. (2004) also present evidence of humanlike features in A. africanus, notably in the position of the lunate sulcus, which marks the extent of the primary visual cortex in the occipital lobe.

There have also been some recent controversies surrounding australopithecine brain sizes, primarily involving one recently found A. africanus skull (Conroy et al., 1998), but calling into question a variety of other australopithecine skull measurements as well (Falk, 1998; Conroy et al., 1998; Lockwood & Kimbel, 1999; Holloway, 1999). The outcome is still uncertain, but there appears to be no reason to believe that the conclusion from Kappelman (1996) referred to above will change.

The brain status of the habilines is unclear, for several reasons. Only a few skulls are sufficiently undamaged to permit a reasonable estimate of cranial capacity, and they are widely different from each other, spanning the full range from australopithecines to more clearly human erectus specimens (Kappelman, 1996). Likewise, available endocasts indicate that some specimens have a very apelike brain surface, whereas others have a more human-like pattern (Brandt, 1993). This spread of data may well be connected with the taxonomic uncertainty surrounding this group, discussed on page 60.

Homo erectus presents a clearer picture, with brain capacities just below the modern human range. There are hints of a gradual increase in brain size throughout the erectus period (Rightmire, 2004), as seen in Figure 5.5. The transitional forms between erectus and modern humans are, as expected, intermediate in brain size. Neanderthals actually had slightly larger brains than we do, on average, but with a larger body size as well, their brain-to-body size ratio was not significantly different from ours. The shape of the Neanderthal brain, however, was quite different from ours, lower and longer. According to Bruner et al. (2003), the Neanderthal brain represents a straight extrapolation of earlier hominid patterns of phylogenetic brain growth, whereas Homo sapiens sets a new trend, with parietal expansion.

Concerning signs of specific language adaptations in the brain, the only ones that may conceivably be detectable in fossils are an increase in lateral asymmetry,
Figure 5.5. Brain size of various fossil and extant hominids, plotted against the time when they lived. Data mainly from Aiello & Dean (1990), with post-1990 finds added. Note that absolute skull volume is plotted, not normalized to body size.
and increased size of those brain modules, if any, that are dedicated to linguistic processing.

Outside of *Homo sapiens*, Wernicke’s area can be anatomically identified also in chimpanzees (Gannon et al., 1998), though Marshall (2000) cautions against over-interpreting this discovery. Broca’s area is seen in both chimpanzees and gorillas (Cantalupo & Hopkins, 2001). Likewise, some fossil skulls, notably that of *Homo ergaster*, have a bulge in the right position to accommodate an enlarged Broca’s area (Walker & Shipman, 1996; Holloway, 1983). But since we have little or no data on their use in chimps and ancient hominids, and since they have non-linguistic functions even in modern humans, the significance of their presence in our relatives is unclear. And in an overall comparison between chimp and human brain shape, more differences are actually observed in the right hemisphere than in the left (Gibbons, 2002c).

If there is a connection between language lateralization and hand preference, as discussed on page 99, then historical and comparative data on hand preferences may be relevant. Human hand preferences have an ancient history, as the oldest stone tools show tell-tale asymmetries revealing that they were knapped by right-handed tool makers (Westergaard & Suomi, 1996; Wilkins & Wakefield, 1995). Similar behavioral asymmetries can be found among whales (Clapham et al., 1995), monkeys (Westergaard & Suomi, 1996; Miller & Paciulli, 2002), orangutans (Rogers & Kaplan, 1996), and chimpanzees (Hopkins & Bard, 1993; Morris et al., 1993), but mainly on an individual level — each individual prefers either left or right, but there is no strong population-level bias (McGrew & Marchant, 1997) comparable to the human pattern with a 10:1 ratio of right:left preference. The meta-analysis of Palmer (2002) revealed statistical anomalies that cast doubt on the positive results of some earlier studies of chimps, but Hopkins et al. (2003) argue that their new data show a population-level bias even after taking into account the criticism of Palmer (2002). The possibility of a small population-level effect in chimps remains an unresolved issue — see the continued exchange between Hopkins & Cantalupo (2003) and Palmer (2003).

Kaas (2000) raises the interesting idea that lateralization of function may be an adaptive response to brain growth — having the same function in two places on opposite sides of the brain is suboptimal, but viable as long as the distance between the two sides is small. But in a large brain, this becomes untenable, and lateralization follows. This would explain why lateralization is so prominent in human brains, and less clear in apes.
5.4 Summary

Concerning the bodily equipment needed for language, the following conclusions can be drawn:

- **Speech production:**
  1. Ape equipment inadequate on the following points:
     - Vocal tract shape
     - Cortical control of vocalizations
     - Breathing control
     - Sequencing control
  2. The items above are probably speech adaptations, and are almost certainly adaptations for some kind of complex vocalizations.
  3. Fossil evidence indicates that Neanderthals had human-style vocal tract and breathing control.
  4. 2 & 3 ⇒ The last common ancestor of us and the Neanderthals, 500,000 years ago, had some form of speech.

- **Speech perception:** The equipment of apes, and for that matter most mammals, is perfectly adequate. This includes features like categorical perception of phonemes, present in non-human primates, and thus not a speech adaptation. Human hearing is apparently fine-tuned for better perception around a few kHz, very likely a speech adaptation. This tuning is present also in Neanderthals, thus strengthening the conclusion above that the last common ancestor of us and the Neanderthals, 500,000 years ago, had some form of speech.

- **Brain size:**
  1. The human brain has increased tremendously in size during evolution, at considerable cost, but the reasons for this are not well established. Social pressures driving the evolution of ‘Machiavellian intelligence’ appears to be the most plausible hypothesis.
  2. Brain growth in the human lineage took place mainly during two periods (see Figure 5.5):
     (a) ~ 2–1.5 million years ago, around the origin of the genus *Homo*, from ape-sized to around 900 cc.
     (b) ~ 0.5 million years ago, around the time of the last common ancestor of us and Neanderthals, from 900 cc to modern size.
  3. The connection between brain size and language is insufficient for any firm conclusions from fossil cranial volumes to language capacities.

- **Lateralization:**
  1. Language in humans is strongly lateralized, mainly (but not exclusively) handled by the left hemisphere.
  2. The lateralization is *not* genetically hardwired — children with early left-hemisphere damage can acquire normal language in the right hemisphere.
  3. The brains of monkeys and apes are not symmetric either, so the appearance of asymmetries in fossil skulls carries little weight as evidence of language capacities.
Brain modularity:

1. Good evidence for modularity only in the case of low-level sensory processing.

2. Even these modules are not innate and genetically hardwired, but display considerable developmental plasticity. The circuitry of the brain is largely shaped by early inputs.

3. In the case of language, there are some signs of small-scale modularity, but not anything resembling a monolithic encapsulated dedicated language module. The evidence rather points towards language being handled by several subsystems, many of which also have nonlinguistic functions.

Further reading

CHAPTER 6

ANIMAL COMMUNICATION IN THE WILD

While language in the strict sense may be uniquely human, numerous other species have their own means of communication, many of which appear to share at least some, if not all, of the properties of language.

‘The infinite use of finite means’ is a crucial property of human language (Humboldt, 1836), but it is also a property of the songs of certain birds and whales, at least in the limited sense of their combining a set of elements in an infinite variety of permutations. Recursivity may be found in some songs (Li & Hombert, 2002), and the coining of new elements has been reported for some birds (Clemmons, 1991). The majority of animal communicative acts may be non-symbolic signals, but there are examples of vocalizations where this is not self-evident (Marler, 1998, further discussed below). Hauser (1997) gives a thorough review of animal communication, in an evolutionary perspective, and Håkansson (1995) provides a popular overview.

Birdsong, just like human language, can be culturally transmitted, and geographical ‘dialects’ are common (Wiener, 1986; Baker, 1996). Some birds appear to have an innate ‘song acquisition device’ (Whaling et al., 1997), in analogy with the proposed language acquisition device of humans. Birds of many species need to hear the songs of other birds of the same species during a sensitive period while they grow up, or they will not develop normal singing abilities (Wiener, 1986; White, 2001). And Okanoya (2002) reports that there are patterns in the song of a Bengalese finch that can be modeled by grammar-like rules.

But whatever similarities there may be between birdsong and human speech, they have to be produced by parallel evolution — the underlying hardware is completely different, both in the vocal apparatus (Goller, 1998) and in the brain. It is interesting for comparative studies (Wiener, 1986), but if we are looking for possible precursors of human language, we had better turn to mammals.

Whale songs and calls are culturally transmitted as well (Whitehead, 1998; Noad et al., 2000; Yurk et al., 2002), and some researchers argue that whale songs have something resembling a hierarchical grammar (Seife, 1999), though this remains controversial. The overall style of whale songs more resembles bird-
song than speech, so their relevance for language is not totally obvious. Gibbon songs also resemble birdsong more than they resemble language, even though gibbons are our fairly close relatives (Hauser, 2000; Geissmann, 2002). Still, Cowley (2002) and Ujhelyi & Buk (2001) suggest that there may be links between gibbon song and human language.

The sounds produced by dolphins have more language-like features, but their quantitative study is still in its infancy. McCowan et al. (1999) find promise in an information-theoretical approach, akin to Zipf’s classical work on human language (1935; 1949), but the available ‘dolphinese’ corpus is still insufficient for any firm conclusions. The classification and discrimination of dolphin ‘words’ is also a non-trivial task for humans (or human-built computers), since the ‘phonology’ of dolphins is so different from ours (Janik, 1999).

6.1 Do animal calls mean anything?

“What do animal sounds mean” asks Marler (1998, p. 2). The traditional view, among biologists as well as linguists, has been that they have only affective meaning, expressing only the emotions of the ‘speaker’, without symbolic referents. This view is likely to be accurate in the case of most animal communication. But numerous studies in recent decades, starting with Seyfarth et al. (1980), have demonstrated that many animals use alarm calls and/or food calls, that for all practical purposes function as if they carried symbolic referential meaning.

The original study of Seyfarth et al. (1980) concerned the alarm calls of vervet monkeys. These monkeys have a set of three distinct alarm calls, used for three different predators (snakes, leopards, and eagles). When a vervet monkey hears one of these calls, he or she takes appropriate action, different for each alarm call. They run for cover in bushes when hearing the eagle call, climb up into the treetops when hearing the leopard call, and stand up to scan the grass when hearing the snake call. A purely affective call, basically conveying only that the caller was scared by a predator, could not reasonably have led to such appropriate actions. Marler (1998) also reviews some interesting data on the call acquisition of young monkeys — the calls as such appear to be innate,¹ but the association of a particular call with a particular predator is learned, and the young monkeys start out by over-generalizing, using the eagle call for anything from falling leaves to actual eagles, and then gradually learn when it is appropriate to use the call. Adults use the call almost² exclusively when a monkey-eating bird is around.

¹Though there exist other monkey and ape calls, where the calls themselves appear to be learned, since ‘dialect’ differences between groups have been observed (Mitani et al., 1999; Fischer et al., 1998; Hauser, 1992; Marshall et al., 1999).
²The rare ‘mistakes’ concern birds that resemble dangerous birds, but are actually safe.
It should also be emphasized that these monkey calls are not, as far as we can
tell, iconic (cf. Section 11.3). They do not resemble any sounds of the predator
they’re referring to (Carstairs-McCarthy, 1996). In Diana monkeys, the alarm calls
of males and females are acoustically different, both from each other and from any
predator sounds, but nevertheless engender the same response in both genders (Zu-
berbühler, 2000d). This non-iconicity is a central property of human languages as
well, and has been invoked as a defining and uniquely distinguishing property of
human languages. Finding similar non-iconicity in animal communication under-
cuts the human claim to uniqueness.

But in order to regard calls as truly symbolic, referentiality and intentionality
are crucial diagnostic features, that are difficult to operationalize in wild animals.
The phrase ‘functionally referential’ is often used in animal communication con-
texts, basically in order to sidestep the contentious issue of whether animals intend
to refer to an eagle or whatever, but retaining the implication that these calls for all
practical purposes function as if the caller intended to warn his fellows. We shall
return to the issue of communicative intent below.

Since the original work by Seyfarth et al. (1980), similarly ‘functionally refer-
ential’ calls have been observed in numerous species:
1. **Birds:**
   - Chicken, both domestic and their wild relatives (Evans & Evans, 1999; Mar-
     ler, 1998; Hauser, 1997).
   - Several species of passerines (Marler, 1998).
2. **Rodents:**
   - Alpine marmots (Marler, 1998), but oddly enough not the closely related
     yellow-bellied marmot (Blumstein & Armitage, 1997).
   - Several species of squirrels (Greene & Meagher, 1998; Marler, 1998).
3. **Suricates** (Manser et al., 2002).
4. **Primates:** too many to list here. Reviewed in both Marler (1998) and Hauser
   (1997). The list includes a few lemurs (Fichtel, 2004), numerous monkey
   species, and chimpanzees.

It should be noted, however, that the functional referentiality is not always along
the dimension of different predator species. According to Zuberbühler (2000d),
squirrels use different calls depending on the urgency of the alarm, and chicken
use different calls depending on predator elevation — a hawk on the ground elic-
its the ground-attack call normally used for foxes, rather than the air-attack call

---

3 The non-iconicity of human language is not, however, totally self-evident and universally accepted.
Nuckolls (1999), Langacker (2001), and Wilcox (2004) argue for non-negligible iconic components in
human speech, and in sign languages the iconicity of many signs is obvious.
4 Shriner (1998) found that squirrels and marmots also respond to each other’s alarm calls.
5 Hauser (1997) and Marler (1998) disagree on how solid the chimpanzee evidence is, which is rather
remarkable since Marler’s sole reference on this issue is to a study by Hauser. Byrne (2000) and
Crockford & Boesch (2003) add some more evidence of chimpanzee call referentiality, which may
help to resolve the issue.
normally used for hawks, and researchers have even elicited the air-attack call from chicken by having a stuffed raccoon flying overhead. This makes it dubious whether anything remotely resembling semantic labeling is involved in either squirrel or chicken calls. Monkey calls, however, clearly label the specific predator, even when it approaches from an unusual direction (Zuberbühler, 2000d). The suricate, a kind of small mongoose from southern Africa, encodes two kinds of information in its alarm calls, both about predator type and urgency (Manser et al., 2002).

The studies discussed above concern predator-alarm calls and food calls, where it is experimentally feasible to demonstrate functional referentiality. Another field of candidate symbolic thought in animals that is experimentally tractable is that of mathematics, where numerous studies of animal counting capabilities have been performed, e.g., Matsuzawa (1985), Boysen & Berntson (1989) Maliukova & Molotova (1995), Carey (1998), Brannon & Terrace (1998), Boysen & Hallberg (2000) or Hauser et al. (1996), but this is less relevant for language. Somewhat more relevant may be the acquisition of numerical symbols, the digits from 0 to 9, by a chimpanzee (Biro & Matsuzawa, 2001).

Vocalizations that are used socially are at least as numerous as alarm calls, and may be more relevant for the origin of language; see e.g., Dunbar (1993; 1996) or Section 10.5 below. But here it is much more difficult to disentangle affective and referential uses. Cheney and Seyfarth and associates have done extensive research on the social vocalizations of baboons, finding that the baboons do extract information from the calls of their peers (more information than is evident to human ears), but that it is difficult to establish communicative intent. The grunts that are used by baboons for various purposes are analyzed by Cheney et al. (1995) and Rendall et al. (1999), and their ‘contact barks’, with the apparent function of maintaining contact between dispersed members of a group, by Cheney et al. (1996). One odd form of social vocalization is the ‘comment’ calls of Barbary macaques (Brumm et al., 2004), issued by bystanders witnessing interactions of other group members.

6.2 Mental states of communicating animals?

‘Communicative intent’ is a central and thorny issue here. A vocalization can hardly be regarded as having anything at all in common with language unless the ‘speaker’ intends to communicate. But what does it take for us to say that an organism intends to communicate?

— At one extreme, consider a plant in a pot. If the leaves of the plant hang down, this informs us that the plant needs water — but does that mean that the plant has communicated its need for water? Hardly. Or intended to communicate? Certainly not.
At the other extreme, when I am exercising my perceived free will to write this paragraph, my communicative intent is self-evident. Animal calls typically fall somewhere between those extremes. Close to the plant we have e.g., the insect that emits pheromones in order to attract the opposite sex. This might reasonably be called communication, but there is nothing resembling intent involved.

More interesting are the various animal calls discussed in the previous section. In principle, an alarm call could be issued purely automatically whenever a predator was seen, in which case it would not be appropriate to speak of intent. Intent presupposes choice, in effect it presupposes free will in the animal. Communicative intent is present only if an animal chooses voluntarily to communicate. This becomes closely entwined with the perennial issue of mind and consciousness, since it makes little sense to speak of the intent and free will of a being without a mind. The questions concerning animal minds are discussed in chapter 8 below; the main part of this issue will be left until then.

But a few aspects of intent are vital in this context. One possibly diagnostic feature that may distinguish between affective and communicatively intentional vocalizations is whether the caller cares who (if anybody) is listening. Operationally, this might be measured as an observed difference in calling patterns, correlated with a difference in the potential audience. This has been found to be the case with some, but not all, of the functionally referential calls listed above, as well as with chimpanzee and orangutan communicative use of gaze and gestures (Leavens et al., 1996; Leavens & Hopkins, 1998; Savage-Rumbaugh & Lewin, 1994; Bard, 1992; Call & Tomasello, 1994), but has been very difficult to establish for the social calls. For example, when chimps communicate they use sounds regardless of whether anybody is looking at them or not, but they use visual gestures only if the recipient is looking (Call, 2001; Leavens et al., 2004), and they also adapt the location of their gestures to where their partner is looking (Povinelli et al., 2003).

Is it possible to explain such an audience effect without communicative intent? A particularly clear example may be chimpanzee alarm calls, where Byrne (2000) reports that chimps make alarm calls only when the predator is hidden from the threatened individual, not when the danger is plainly visible. Some linguists, such as Bickerton (1995), remain skeptical of any claims that animal vocalization are anything but affective. The evidence remains strongly suggestive, but not conclusive.

---

6Including even the calls of domestic hens (Wauters et al., 1999)
7Call et al. (1998) and Tomasello et al. (1999) found intriguing but ambiguous results on whether chimps can use information from the gaze of others. Monkeys failed similar tests (Anderson et al., 1996).
8These studies are of captive human-raised chimps, but Vea & Sabater-Pi (1998) found that wild bonobos also use gestures, and Jucquois (1991) claims that gestures are “un moyen de communication privilégié” (p. 22) for wild chimps.
9Dolphins also monitor and adapt to the attentional state of the recipient of their visual gestures (Xitco et al., 2004).
Martin (1998a) addresses the issue of communication and referentiality from the point of view of the ‘listener’ rather than the ‘speaker’. He asks “What does it mean to claim that a word (or ... alarm call) is representational?” (1998a, p. 72), and answers the question in terms of neurological processes. His answer, in short, is that the monkey’s ‘mental image’ of an eagle should be activated not only by seeing an eagle, but equally by hearing the appropriate alarm call, in the same way that hearing a word for us evokes a mental image of the word’s referent. This is the listener’s side of the ‘joint reference’ of Sinha (2001).

Activating a mental image is here to be contrasted with simply triggering the appropriate action, something which the monkey could do ‘instinctively’, or due to simple conditioning, without symbolic processing. A scheme for experimentally distinguishing these possibilities is proposed by Martin (1998a), based on neural imaging results from human word processing, but the appropriate experiments remain to be done with monkeys.

Zuberbühler (2000a; 2000b; 2000c) has performed a different series of experiments with Diana monkeys that appear to show that these monkeys do not just have the appropriate action automatically triggered by an alarm call. To begin with, Diana monkeys respond not only to their own alarm calls, but also to the alarm calls of other species living in the same area, notably guinea fowl (Zuberbühler, 2000a), chimpanzees (Zuberbühler, 2000b), and Campbell monkeys (Zuberbühler, 2000c).

Campbell monkeys produce distinct calls for leopards and eagles, to which Diana monkeys respond with appropriate action. So far, this is no different from the vervet monkeys of Seyfarth et al. (1980). But if an alarm call is followed after a brief interval by playback of predator vocalizations, the reaction of the monkeys differed depending on whether the predator vocalization matched the alarm call or not. This clearly indicates that something beyond the triggering of an action goes on in the monkey’s mind, and may indicate that a mental image was triggered by the alarm call and compared with the subsequent predator call (Zuberbühler, 2000c).

Guinea fowl, unlike the monkeys, do not produce differentiated alarm calls, but use the same call for both leopards and humans (both of which eat both fowl and monkeys). If the Diana monkeys hear the guinea fowl call, they react as if a leopard were present — unless they already had independent reason to suspect that humans were in the vicinity, in which case they took the fowl alarm as a sign that the humans were getting too close. This means that the monkeys are capable of reacting differentially to the same call, depending on which other information they have available, appearing to apply causal reasoning to infer the probable cause of the fowl call (Zuberbühler, 2000a).

More complex is the interaction of Diana monkeys with chimpanzees. Leopards hunt both chimps and Diana monkeys, so both have a leopard alarm call which they use whenever they notice a leopard. But chimps also hunt Diana monkeys and eat them. The Diana monkeys, however, do not use alarm calls to warn for chimps.
— instead they hide quietly when chimps are around, implying that chimps (but not leopards) are smart enough to take advantage of an alarm call revealing the location of their dinner. The interesting part comes when chimps encounter a leopard and issue their leopard alarm call, within earshot of Dianas — should the Diana monkeys then adopt their anti-chimp tactic, because they hear chimps, or adopt their anti-leopard tactic, because the chimp calls tell them that a leopard is around? The results of playback experiments indicate that some groups of Diana monkeys behave as if they understood that the chimp call means ‘Leopard!’, whereas other groups only heard it as indicating the presence of chimpanzees. Interestingly enough, the Diana groups that had the most experience of being hunted by chimps were more likely to adopt anti-leopard tactics when hearing the chimp call. The interpretation of these results is difficult, but it does hint at non-trivial cognitive processes in the Diana monkeys (Zuberbühler, 2000b).

The neural basis for primate communication has been studied in a few lesion experiments on monkeys. One notable result is that damage to Broca’s area does not affect their vocalizations, nor their gestures, implying that the human use of Broca’s area for language processing is a later development, and that speech is not a direct descendant of monkey vocalizations. Instead, part of Broca’s area in monkeys control chewing and other mouth motions, according to MacNeilage and associates (MacNeilage, 1998; MacNeilage & Davis, 2001), who invoke this functional change in support of their hypothesis that speech evolved from chewing motions. Other monkey homologs of human language areas are used for auditory processing (Poremba et al., 2003) including conspecific vocalizations (Ghazanfar & Hauser, 2001), and control of facial muscles, which may also be potential exaptations for language evolution (Nowak, 2000b).

6.3 The evolution of animal communication

Another issue to consider in the context of animal food calls and alarm calls is why they have evolved. The receiver of an alarm call or food call ought to benefit from reacting appropriately, so the evolution of the ability to perceive and respond to calls is less problematic. But what about the caller — shouldn’t alarm calling impose a considerable risk on the caller, who attracts the attention of the predator? As noted in Section 3.3, behaviours do not normally evolve for the benefit of others, if they are detrimental to the survival of the one performing the behaviour in question. In the specific case of communication, Noble (1999) has shown that under reasonable assumptions communicative behaviours will evolve only if both sender and receiver benefit from it. Kin selection may be invoked in the case of alarm calls — a behaviour may be evolutionarily favored even though it harms you, if the benefit to your close relatives is large enough so that the net effect is to enhance the spread of your genes, copies of which are carried by your rela-
tives (Silk, 2002). A monkey surrounded by siblings and cousins may well find it profitable to its genes to raise a predator alarm, despite some personal risk. Still, the evolutionary pressures behind communicative behaviours remains a non-trivial issue.

Nicastro (2001) proposes that the habitat plays an important role in the evolution of alarm calls, in that the visual affordance of open environments invites more nuanced alarm calls, as well as decreasing the risks to the caller. The vervet monkeys discussed earlier do live in such an open habitat, but the Diana monkeys of Zuberbühler (2000b) do not, nullifying the support Nicastro (2001) invokes from the vervets.

Another problem in the evolution of communication concerns how to keep callers honest (Gouzoules & Gouzoules, 2002). Why not cry ‘Leopard!’ when the other guy is sitting in a rich food patch, and then steal his food after he’s run for cover? Was Kipling (1928) right:

> What was the first practical use Man made of [speech]? Remember he was, by that time, past-master in all arts of camouflage known to the beasts. ... In short he could act any kind of lie then extant. I submit, therefore, that the first use Man made of his new power of expression was to tell a lie — a frigid and calculated lie. (Kipling, 1928, p. 233), quoted in Clark (2000, p. 407)

This problem gets even worse in the case of ‘advertising’ signals used in sexual displays and similar competitive contexts (Hsu et al., 2002). A common solution in the animal (and human) world is to rely on signals that are expensive to send — any man can say to his date that he’s rich, but she won’t be impressed until she sees tangible evidence of his wealth. However, human language is not based on this costly-signal solution — talk is proverbially cheap. Lachmann et al. (2001) and Noble (2000) review this issue, and propose possible solutions for the evolution of cheap talk. Furthermore, even if we did talk, why would we use our powers of speech for informative communication, rather than for the deception suggested by Kipling above? Dessalles (2000) proposes the interesting idea that while speech in itself may be cheap, relevant information-laden speech is not — and talk containing valuable information, unlike lies or empty chattering, can gain you status in the group; cf. Section 10.5. Dessalles (2003b) shows in computer simulations that altruistic information-sharing may arise this way. But the issue still remains problematic.

A related issue is raised by Baker (2003), who discusses why human universal grammar has parameters. Why can languages be either head-first or head-last, and so on? Baker’s proposed solution is that these parameters are an adaptation not to aid but to hinder understanding between different tribes, an easy way to make languages mutually unintelligible. The purpose of this adaptation would be to conceal information from outsiders. This is an intriguing suggestion, but the empirical support for it is at present rather meager.
6.4 Animal syntax?

Concerning language beyond the level of single words, it is generally believed that the only species to have that in the wild is *H. sapiens*. But, as Savage-Rumbaugh et al. (1993) point out, the empirical support for this belief is not overwhelming. There are major methodological problems involved in testing whether a species uses a language-like communication system, particularly as we cannot take for granted that it would resemble our own vocal speech (Hauser, 2000). Both chimps and dolphins in the wild do manage to coordinate their behaviour in quite sophisticated ways, which is difficult to explain in the absence of a non-trivial communication system. On the other hand, the few investigations of chimp vocal behaviour in the wild that have been done (Arcadi, 2000, and references therein) have not been able to find any signs of complex language, though their compound calls do have some syntax-like structure, according to Ujhelyi & Buk (2001).

There is also, for what it’s worth, a considerable body of anecdotal evidence of chimps conveying fair amounts of information by vocalizations and gestures alone, including information that couldn’t have been indicated by situational cues (Savage-Rumbaugh & Lewin, 1994).

Zuberbühler (2002) reports that he has found evidence that Diana monkeys react to certain combinations of calls as if one call were a modifier of the other. The monkeys reacted to the ‘main’ call quite differently (and in both cases appropriately) depending on whether it had been preceded by the modifier call. Zuberbühler interprets these results as evidence “*that nonhuman primates are able to generate and comprehend simple syntactic rules...*” (2002, p. 298). That may be going too far, but his results are nevertheless intriguing.

In the light of the results on captive apes and dolphins (reviewed in the next chapter), further empirical exploration of this topic is clearly warranted. At present, we have insufficient data for any kind of conclusion.

6.5 Summary

- Most animals communicate in some way.
- In most cases, there is no reason to believe that animal communication is anything but affective displays, telling the world ‘I am angry!’ or ‘I am horny!’.
- There are nevertheless a number of examples of animal calls that appear to have a referential function:
  - *Alarm calls* in a number of species, most famously vervet monkeys. When a monkey sees a leopard it issues the ‘leopard call’, and the others in the flock run to the treetops where they are safe from leopards — but when a monkey sees an eagle it issues the ‘eagle call’ instead, whereupon the others run from the treetops where they are exposed to eagles. For all practical purposes, these calls function as if they were words referring to the respective predators.
Origins of language

– *Food calls*, informing the rest of the flock that there’s food to be had.
– *Social calls*, used for a wide variety of social functions.

It is still an open issue to what extent these functionally referential calls carry symbolic meaning like human words, and in what sense, if any, the caller and receiver understand what they are ‘saying’ (rather than just acting from unconscious instinct or conditioning).

For some types of calls, it has been established that the caller modifies its calling behavior depending on who is listening. There are hints also of non-trivial cognitive processes behind the interpretation of some calls. But in general, there is little evidence concerning the mental states of communicating animals.

Some animal calls, such as bird songs, are composite, made up of smaller units that can be recombined in different ways. But there is little evidence of syntax in any interesting language-like sense beyond *Homo sapiens*.

Thus, there is some evidence that the natural communication systems of some animals contain units that functionally resemble the words of human language. Whether the calls are actually symbolic, proto-symbolic, or mere association-based signals that functionally resemble words, remains an open issue. But even if the ability to use and understand single symbol-like calls were indeed present in some non-humans, this would be a far cry from full human-style language capabilities. It is a necessary precondition for the evolution of true language, but both syntax and true symbolicty remain as major hurdles.

The selective advantages behind the evolution of communication is a non-trivial issue. The advantage of better communication would appear to be self-evident — but most advantages of e.g., alarm calls or food calls go to the receiver, not to the caller. The caller may well incur a cost instead, so why would a tendency to issue calls evolve? There have been some theoretical and simulation work done on this point, but the issue remains open. This is an unsolved problem in the case of human language as well.

Further reading


CHAPTER 7

CAN NON-HUMANS BE TAUGHT LANGUAGE?

It is well-known that many animals can be taught behaviour that they do not normally display in the wild, even some typically human behavioral patterns. Attempts to teach language to animals has a long history, with mixed results. Many birds can learn to mimic human speech, but in most cases show little sign of this being anything but rote learning without understanding. Likewise, we have little reason to believe that dogs responding to verbal commands show evidence of anything beyond conditioning.

More interesting are the recent systematic attempts to teach language in a fuller sense to apes, dolphins, and parrots. These are especially relevant in connection with the Chomskian linguistic paradigm, in which language, and particularly grammar, is supposed to be uniquely human, and to require a ‘language organ’ found only in humans. If language, with grammar, could be taught to non-humans, this would falsify the strong Chomskian claims of human uniqueness, and would cast strong doubt on the existence of a dedicated ‘language organ’. Not unexpectedly, Noam Chomsky displays a negative attitude towards ape language research, dismissing it out of hand: “But the question whether it [an ape] has a language faculty is a meaningless question and therefore nobody should talk about it.” (Chomsky, quoted in Belsack et al. (1999, p. 35)).

That an ape might be taught to speak or sign was first suggested by Samuel Pepys in 1661 (Limber, 1982; Fouts, 1997), and again by La Mettrie (1742), but the first actual systematic attempt took place in the early years of the 20th century with an orangutan, who unfortunately died young before much progress had been made. Several unsuccessful attempts to teach chimpanzees to speak also took place around the same time (Furness, 1916; Yerkes & Yerkes, 1929, cited in Limber (1982)). Later, during the 1930s, the Kellogg family tested ape learning in the environment in which humans learn language, by co-rearing a young chimpanzee, Gua, together with their own son (Savage-Rumbaugh et al., 1993). Gua kept pace

---

1 Mammals generally do not mimic sounds like many birds do, but there is at least one case reported of a mimic seal (Deacon, 1997).
with, or even outpaced, their son Donald on many tasks, including human language comprehension (but not production). Rumor has it that Donald also acquired a fair number of chimp calls from Gua. But the experiment was terminated before the age of two, before either participant had really got into grammar (Desmond, 1980; Fouts, 1997). A similar experiment a few years later again employed a chimp baby, Viki\textsuperscript{2} (but this time no human baby). Viki, like Gua, learned to respond to a number of sentences as if she understood them, but it is unclear how much of this ‘understanding’ was based on language, and how much on situational cues. Language production was a failure, again. Viki could articulate only a few words, with great difficulty (Savage-Rumbaugh et al., 1993; Desmond, 1980; Deacon, 1997). This limitation can be explained by the different vocal-tract anatomy and neural wiring of apes, as discussed in Chapter 5.

In non-spoken language modalities, where the animals do possess the requisite hardware and control, teaching experiments have been much more successful, though the significance of that success is still a contentious issue. In a series of experiments, starting in the late 1960s, several apes of four different species, as well as dolphins and parrots, have been taught to use appropriate language modalities that they are anatomically equipped for:

- Chimpanzees: Rumbaugh et al. (1975), Gardner & Gardner (1984), Terrace et al. (1979), and many others
- Bonobos: Savage-Rumbaugh and associates (1986)
- Gorillas: Patterson and associates (1981; 1990)
- Orangutans: Miles (1990, cited in King (1996) and Belsack et al. (1999))
- Dolphins: Herman et al. (1984)

### 7.1 Apes

The pioneer here is a chimp named Washoe, with her ‘adoptive parents’, the Gardners (1984; 1985). Washoe was taught sign language from an early age, and learned to reproduce a large number of signs, and appeared to use them communicatively. A very interesting development with Washoe was when her adoptive son, Loulis, appeared to acquire sign language from her, without human intervention or training (Fouts, 1997).

A different modality, with abstract tokens invented for this purpose, either on physical plastic chips, or on a computer keyboard, was taught to three chimpanzees at the same time as the Washoe project (Rumbaugh et al., 1975; Savage-Rumbaugh et al., 1993), with a similar degree of success. Lana and the other apes learnt to

\textsuperscript{2}Or Vicki (Deacon, 1997); the spelling varies in different sources.
produce large amounts of more-or-less appropriate strings of tokens, in their respective modalities, which was interpreted as evidence of their linguistic prowess.

But these early experiments were severely criticized by Terrace et al. (1979) who attempted to replicate the work of Gardner & Gardner (1984) with a different chimp, named Nim Chimpsky.\textsuperscript{3} Terrace et al. (1979) concluded that all the ‘utterances’ of Nim (and by extension those of Washoe and the others as well, even though the experimental conditions were far from identical) could be explained as simple imitation, ‘parroting’, of the human teachers, reinforced by the rewards given to the apes for producing appropriate ‘words’.

Terrace et al. (1979) did have a point in that the early experiments were sometimes rather lax in their methods and optimistic in their interpretations. The experimental protocols used were insufficient to distinguish between actual language learning, and the null hypotheses of parroting or the ‘Clever Hans’ effect.\textsuperscript{4} Subsequent ape language experiments were performed with modified methods and tightened controls, using blind tests\textsuperscript{5} and other devices, in order to circumvent the valid criticisms of Terrace et al. (1979) and others.

An interesting experiment is the one with the chimps Sherman and Austin, performed by Sue Savage-Rumbaugh and associates (1978; 1980; 1994). For the first time, the chimp pupils lived in a social setting with other chimps, and the emphasis was on learning and using a set of tokens, rather than on producing strings of signs. Communication was in focus, and for the first time an ape experiment put as much weight on comprehension as on production. The experiment was quite successful, giving strong evidence that chimps are capable of symbolic thought and symbolic communication (Deacon, 1997), as well as communicative intent (Savage-Rumbaugh & Lewin, 1994). According to Greenfield & Savage-Rumbaugh (1984), there are some interesting similarities with the way young children handle symbols.

Sherman and Austin definitely achieved the joint-reference proto-symbolic level of Sinha (2001) discussed in Section 11.3. This is shown most clearly in their ability to invest new (or previously unused) tokens with meaning, e.g., in the experiment described on p. 79 in Savage-Rumbaugh & Lewin (1994), where Sherman and Austin are given a novel food item, and one of them chooses a token to use for this item. Pointing back and forth between the token (on a keyboard) and the item, they establish jointly the correspondence between referent and symbol, to all appearances making sure that they both knew that the other knew the correspondence, and from then on both of them used the agreed-upon token for this food

\textsuperscript{3}Distantly related to a certain famous linguist.

\textsuperscript{4}Clever Hans was a German horse, who displayed remarkable apparent mathematical abilities. But his prowess did not derive from an ability to calculate, but instead from an ability to read subtle nonverbal cues from people. It has been demonstrated that apes are similarly adept at using experimenter cues (Itakura & Tanaka, 1998).

\textsuperscript{5}According to Fouts (1997) blind testing was used also with Washoe, with fair success, but the Washoe project was nevertheless discredited, possibly unjustly so, by Terrace et al. (1979).
item. It is difficult to interpret this event as anything but them jointly assigning that token as a symbol to ‘mean’ the new food item.

Whether construal *sensu* Sinha (2001) was also present is a matter of interpretation, both of Sinha and of the chimps. Operationalizing construal in a controlled experiment is non-trivial, and as far as I can tell none of the controlled experiments performed with Sherman and Austin gives a clear answer on this issue.

There is also evidence from later ape studies that chimps can handle another important aspect of symbolicity as well, detached representation (see page 7), initiating communication about absent objects, including objects not recently seen, outside the context where they were last seen (Menzel, 1999), clearly distinguishing their representations from the cued representations more typical for non-humans.

The Sherman-Austin experiment escaped the devastating criticisms that earlier efforts had received, but instead it was argued that the abilities involved were too far removed from language to be of any relevance, because syntax was not emphasized. In the absence of syntax, the clear symbolic communication at the one- and two-word level, not only between chimps and experimenters, but also between the two chimps, was dismissed (Savage-Rumbaugh & Lewin, 1994), even though symbols are as important as syntax for language, and the clear evidence of joint reference in non-humans is an important breakthrough.

Another important breakthrough came about quite accidentally in the 1980s, again in the lab of Savage-Rumbaugh et al (1985a; 1993). A teaching experiment, basically similar to those with Lana et al, was attempted with a bonobo female named Matata. Matata herself failed to learn anything notable, but her infant, Kanzi, who had spent the language lessons clinging to his mother’s fur or playing in the ‘classroom’, spontaneously started to display signs of apparent language acquisition.

This was initially attributed to species differences; possibly bonobos were more adept at language learning than common chimps (Savage-Rumbaugh et al., 1985a; Savage-Rumbaugh et al., 1985b; Sevcik & Savage-Rumbaugh, 1994), but this does not explain the failure of Matata, who was after all also a bonobo. An alternative explanation lies in the learning framework of Kanzi, who was not deliberately taught anything, but was instead immersed in a symbol-using environment from an early age. This is unlike most of the other laboratory ape experiments (including Matata), which have not started with infants, and where the learning has been much more structured and formal. Kanzi’s case has some interesting parallels to the language learning of human infants, where early exposure to language is vital, and where formal teaching is not a prominent part of the acquisition process.

---

6Sources differ on whether he is her biological or adopted child. The official Kanzi biography, which can be found at http://www.gsu.edu/~wwlrc/biographies/kanzi.html, says he’s adopted. (The same website also has biographies of the other apes involved with Savage-Rumbaugh et al.) As the story is told by Savage-Rumbaugh & Lewin (1994), he was practically kidnapped by Matata from his biological mother.
These two explanations were later contrasted in a controlled experiment, where both a bonobo (Kanzi’s kid sister, in fact) and a common chimp were exposed to the same kind of learning environment as Kanzi (Brakke & Savage-Rumbaugh, 1996). This experiment did show some species differences, but nothing resembling a clearcut with/without language ability contrast. And the chimp learned language significantly better in this environment than her conspecifics had done in more formal settings, so the environmental effect is clearly present.

The informal social learning environment of Kanzi and his friends is apparently very effective. Another aspect of it is that Kanzi interacts socially and communicatively with his human teachers in a much more natural manner than in previous ape experiments. He is an active participant in activities and conversation, and according to both Savage-Rumbaugh & Lewin (1994) and Shanker & King (2002) it is difficult to avoid assigning mental states to Kanzi, in the same way and on the same basis as we infer mental states in our fellow humans. Kanzi displays exactly the same behavior that would cause us to infer that a human had understood what we said. But this type of subjective impressions is very difficult to quantify in controlled experiments, making the social communicative aspects vulnerable to the criticism that they are based entirely on anecdotes and anthropomorphizing (Miklósi, 2002).

Some, notably Deacon (1997), argue that the experience with Kanzi is evidence for a critical period in the language acquisition of apes, similar to that found in humans (Grimshaw et al., 1998; Batali, 1994). Savage-Rumbaugh et al. (1993) further support the existence of a critical period in apes by noting that of the nine apes that they had up to then reared in a similar environment, the four who were exposed to language before the age of 2 1/2 learned the use of symbols easily and naturally, whereas the five who weren’t exposed to language until later in life required extensive training to make even modest progress.

But, as further argued by Deacon (1997), this raises a perplexing problem. The critical period is generally taken to be an attribute of the innate language acquisition device of the Chomskian paradigm, so if apes display a critical period, this would be evidence of their having a language acquisition device. But why would apes, who do not, as far as we can tell, acquire language in the wild, possess a language acquisition device, dormant for millions of years until human scientists came along? Parallel evolution of a complex device that isn’t used is evolutionarily ridiculous. More reasonable is the hypothesis that the common ancestor of humans, chimps, and bonobos already possessed the early enhanced learning abilities that we to all appearances have in common. This implies one of the following four possibilities:

1. Chimps and bonobos do use and acquire language in the wild, with a language acquisition device that’s a shared inheritance from our common ancestor. This cannot be totally excluded, but there is no real evidence in favor of it either. Of
course, if this possibility is correct, then the whole controversy over Kanzi and his friends is moot.

2. This common ancestor already had language, which was then subsequently lost in the chimp/bonobo lineage, with some vestiges of the language acquisition device retained. But why would they lose something so obviously useful?

3. The ‘language’ acquisition device isn’t language-specific, but is a more general learning device, used by the common ancestor, and still by the chimps, for some purpose other than language.

4. The apparent critical period in language acquisition is simply a consequence of the general early brain plasticity discussed on page 111, which is certainly shared between us and the other apes.

As long as we have no evidence of language acquisition among apes in the wild, the fourth possibility appears most likely. This means that if the existence of critical periods in apes is confirmed, the case for a uniquely human language acquisition device is weakened. But the data available so far on ape language acquisition at different ages would be statistically insufficient even if it were uncontested.

As interpreted by Benson et al. (2002), Kanzi has some notion of the pragmatics of language as well, being able to participate in the give-and-take of a conversation, and participate in the construction of a social world of discourse. It is not totally obvious how much of the construction was done by Kanzi, and how much by his human partner, but the conversation analyzed by Benson et al. (2002) is nevertheless interesting.

Another aspect of language where Kanzi and his friends excel, by ape standards, is the comprehension of spoken English. Comprehension experiments are often vulnerable to the Clever Hans hypothesis, but experiments with Kanzi and a few others have been done under circumstances where it is hard to escape the conclusion that they actually understand even syntactically non-trivial spoken sentences. Kanzi has, for example, demonstrated in blind tests comprehension (as evidenced by correct action) at the 70%–80% level of sentences like ‘Give the lighter to Rose’ or ‘Go get the carrot that’s in the microwave’. This is marginally

---

7 The methodology is described in detail in Savage-Rumbaugh et al. (1993), and appears as watertight as is reasonably possible, consistent with ethical and practical considerations (Bates, 1993).

8 A non-negligible fraction of the remaining 20% - 30% may simply be due to him not being in the mood for yet another silly request — in experiments with apes, as with young children, getting and keeping their attention and cooperation is a non-trivial problem. As Bates (1993) puts it: “I sometimes wonder whether we are working towards a rich theory of language comprehension in the docile child — a theory that may not extend to their more rambunctious peers.” (p. 230), a quote that applies even more to ape studies. Other methodological concerns from child language studies are discussed in Guasti (2004), concerns that may be applicable to apes as well.

9 It might be worth noting that this sentence contains an example of recursivity, albeit minimal. Kanzi’s successful parsing of this structure is interesting particularly in the context of the arguments of two proponents of language as innate and uniquely human, Hoekstra & Kooij (1988). After grudgingly admitting that some progress has been made in teaching apes to use symbols, they go on to argue that the recursivity of human language is forever out of reach for apes:

Both systems [language and counting] are recursive, and neither is accessible to apes. ... This difference in the capacity for handling recursive systems is not a quantitative matter: one either has it or one doesn’t.
better than the performance of a two-year-old human child in the same experiment (Savage-Rumbaugh et al., 1993; Belsack et al., 1999). Kanzi’s sensitivity to word order reversals (‘Put the juice in the egg’ versus ‘Put the egg in the juice’) was likewise significantly better than chance (as well as better than his human competitor). He has also shown his ability to carry out multi-step instructions received through a telephone, with no person physically present aware of the details.

Comparing details in the performance of Kanzi with that of Alia, his human co-subject, some significant patterns can be observed. Contrary to expectations, and contrary to what is known about ape performance on sequential-learning tasks (Conway & Christiansen, 2001), Kanzi is the one to excel on sentences where syntax is the key, whereas Alia outperforms Kanzi on those sentences where the syntax is simple but short-term memory may be a limiting factor (Savage-Rumbaugh et al., 1993).

According to Belsack et al. (1999) and Savage-Rumbaugh & Lewin (1994), Kanzi’s own production also displays hints of syntactical structure, with consistent word order, on the same level as a child at the two-word stage of language acquisition. The word order is determined by constituent roles, rather than by the specific words, as in human grammars — Agent–Action and Action–Patient, rather than e.g., consistently putting object names first regardless of their thematic role (Savage-Rumbaugh & Lewin, 1994).

The rate of word-learning by Kanzi and his sister was measured by Lyn & Savage-Rumbaugh (2000), and was found to be quite competitive with that of young human children. The apes were capable of mapping new (invented) words onto objects with only a modest number of trials, even without visual contact with the objects.

Very impressive results appear to have been achieved with the gorilla Koko (Patterson & Cohn, 1990; Patterson, 1981). Unfortunately, the story of Koko’s apparent language acquisition is not stringently documented and controlled, so the earlier criticisms of Terrace et al. (1979) still apply. This lack of stringency is unfortunate, since remarkable abilities are reported for Koko, that would lead to very interesting conclusions if they could be corroborated. These abilities include communicative competence close to that of Kanzi, but also things like the production of representational art and the reporting of what a work of art is supposed to represent, and an IQ well within the human range as measured by standard IQ tests.

There are also reports of gorilla groups in captivity spontaneously developing a system of gestural communication (Tanner & Byrne, 1999; Pika et al., 2003).

---

The fact that all languages are characterized by this recursive property reflects a predetermined faculty of the species ... (Hoekstra & Kootj, 1988, p. 33).

Their claim that counting is inaccessible to apes is dubious as well — see page 122.

10It is interesting in itself that he appears to understand that the voice in the telephone is actually a human elsewhere, and not just a talking box.
7.2 Dolphins

The dolphin experiments of Herman et al. (1984) are quite different from the various ape experiments, in that they focus entirely on comprehension rather than production. Two dolphins, Phoenix and Akeakamai, were taught to respond to sentences in two different artificial languages, one sign language (with a human doing the signing at the edge of the pool) and one whistle language (with a computer doing the whistling, through loudspeakers in the pool), each with its own well-defined grammar. The grammar was quite simple, but did possess features like word-order dependence, in order to test whether the dolphins were sensitive to syntax as well as semantics. All sentences were in the imperative mood — instructions for the dolphins to perform various actions — so they lacked explicit grammatical subjects, but could have both direct and indirect objects. Phoenix was trained with the basic word order DirectObject + Action + IndirectObject, and Akeakamai with IndirectObject + DirectObject + Action. Modifiers to both objects and actions were used as well. Some actions had only a direct object, others had both types of objects. The same items could be used in either object position, so that both sentences in syntactic minimal pairs like \textit{HOOP FETCH PIPE} (‘Fetch the hoop to the pipe.’) and \textit{PIPE FETCH HOOP} (‘Fetch the pipe to the hoop.’) were used and correctly acted upon.

The experiments were successful, in the sense that the dolphins learned to respond correctly to sentences in their respective languages. Testing the dolphins was done with a protocol containing reasonable precautions against the Clever Hans effect and other problems that have cast doubt on many ape results. Their performance is high above chance level, also in the case of novel sentences that have not appeared in training. Typical success levels are in the 80%+ range for both dolphins. Even in semantically reversible sentences, where a correct interpretation of the syntax is vital, the dolphins performed well, with Phoenix achieving 77% entirely correct actions and Akeakamai 59% entirely correct. The incorrect responses were rarely due to sentence reversal — the direct object was correctly identified in 90%+ of the reversible sentences for both dolphins, clearly demonstrating that the dolphins are sensitive to word order.

With the description given in Herman et al. (1984), it is difficult to interpret their results without postulating both semantic-symbolic and syntactic abilities in the dolphins. The success with both dolphins, using different modalities and different grammars, indicates that dolphins are capable of learning arbitrary rules and symbols. The sign language consisted of largely iconic signs, but the whistle language resembles the modality of natural dolphin

---

11 There are two caveats to be raised here, due to the very different sensory world of dolphins. It is not self-evident that gestures by humans that appear iconic to humans also do so to dolphins — dolphins do use gestures, e.g., pointing with their snout (Xitco et al., 2001), but human arm waving may appear quite alien to somebody who lacks limbs. On the other hand, given the major use that dolphins make of sonar, it is not self-evident that whistles that do not appear iconic to humans, may not be iconic for
communication, but the sign language does not. Despite these contrasts, both were roughly equally learnable.

The performance of these dolphins is quite competitive with that of Kanzi, or for that matter that of human two-year-olds, on similar tests. However, as the experiments focused entirely on comprehension, with no production and no dialogue, only abstract language-like skills were tested, and not any social communication. The choice of modalities is perhaps unfortunate, as two-way communication with humans is impossible in both cases — the sign language is anatomically impossible for dolphins, for obvious reasons, and the whistle language is partially outside the range of human hearing, and far outside the range of human vocalization. In principle, the two dolphins could have communicated with each other using the whistle language — but only one of them was taught it. An experiment where dolphins were immersed in a social and communicative language-using environment, in a paradigm similar to those that have been successful with apes and parrots, would be very interesting.

Further experiments with dolphins would clearly appear warranted, but are even more cumbersome and expensive than ape studies; to the best of my knowledge this dolphin study has not been replicated.

7.3 Parrots

Irene Pepperberg and associates (1998; 1999; 2001) have taught a number of Grey parrots, *Psittacus erithacus*, to communicate with humans, using English words. Their star student is Alex, whose achievements rival those of many of the mammals described above. Among his apparent abilities are:

- Correctly labeling a fair number of objects
- Appropriately using hierarchical concepts, e.g., ‘Color’ $\mapsto \{\text{‘Blue’}, \text{‘Green’}, \text{‘Red’}, \text{etc }\}$.
- Appropriately using simple expression frames like ‘Want X’ or ‘Wanna go Y’, substituting labels for X and Y that appear to correspond to his desires.

The training regime used is emphasized by Pepperberg (2001) as crucially important to the success of her experiments. Unlike earlier and less successful parrot experiments, where standard conditioning techniques were used, Pepperberg and associates are using a social interaction paradigm, ‘Model/Rival’, that mimics typical social interactions of these parrots in the wild. There are interesting parallels to be drawn with the ape studies described above, where more naturalistic and spontaneous learning in a social-interaction context, as with Kanzi and his friends,
has been clearly more successful than conditioning and systematic training, in developing usable language-like communication skills.

### 7.4 Patterns of non-human language acquisition

In language teaching experiments, impressive progress has been made in establishing the capacity for at least rudimentary forms of ‘uniquely human’ cognitive and communicative skills in both apes and dolphins and possibly parrots. The dolphins and parrots, however, are phylogenetically distant from us, and the Mesozoic common ancestors that they share with us clearly did not have anything remotely resembling their current cognitive capacity. This implies that their communication abilities are the result of parallel evolution, and thus tells us little directly about the origins of human language. However, it is interesting to note that all animals that have had any kind of success in this type of experiments belong to highly social species with complex group interactions, something which may hint at social processes as a driving force behind the evolution of language, an idea further pursued in Section 10.5.

Of more direct relevance for the origins of human language are the experiments performed with non-human apes. Whether the skills acquired by these apes should be regarded as language-related is still hotly contested, but the achievements of Kanzi and his friends leave less and less room for reasonable doubt that at least some aspects of language are within reach of non-humans. A strong case can be made for the existence of symbolic thought in apes, and the use of symbols that are referential in the same sense of conventional shared meaning as human words. The case for syntax acquisition by non-humans used to be weaker, but the achievements of both Kanzi and the dolphins are hard to explain without them understanding elementary syntax.

Pinker (1995) discusses the key issue of whether the abilities displayed by chimps are homologous to human language. He concludes that:

> Though artificial chimp signaling systems have some analogies to human language (...), it seems unlikely that they are homologous. Chimpanzees require massive regimented teaching ... This contrasts sharply with human children... (p. 3, online edition).

The early ape experiments might appear to support Pinker’s point — but when the experience with Kanzi and his successors are taken into account, the opposite pattern emerges. ‘[M]assive regimented teaching’ is less successful with apes than is a more spontaneous communicative learning regime — exactly as with human children. This is seen already in the title of the paper presenting Kanzi to the scientific world: *Spontaneous symbol acquisition and communicative use by pygmy chimpanzees (Pan paniscus)* (Savage-Rumbaugh et al., 1986).

So what would the presence or absence of rudimentary language and syntax abilities in apes tell us?
Its absence would actually tell us very little (Pinker & Bloom, 1990), other than constraining the time frame for the origin of language to the time since our last common ancestor with chimps. It would be a corroboration (sensu Popper (1963)) of the hypothesis of a unique dedicated language system, but a very weak one. It would not really help us understand its evolution.

Its presence, on the other hand, would be quite informative. Any language acquisition by apes would falsify the notion that language is uniquely human, and bridge the last apparent gulf between us and other animals. And detailed information on what the apes can and cannot do with language would strongly constrain theories about the evolutionary path of our language abilities.

The patterns observed in direct comparison between human and non-human language acquisition are intriguing:

1. Systematic training of apes within a conditioning paradigm has met with very limited success
   but
   language acquisition in a social communicative context works much better.
2. Adult apes who were not exposed to language at a young age, are poor language learners
   but
   apes immersed in language from a tender age are much more successful.
3. Non-human acquisition is often a slow and tedious process in the beginning
   but
   after some threshold is passed, learning can be much more rapid and human-like (Pepperberg, 2001; Lyn & Savage-Rumbaugh, 2000).
4. When an ape (Kanzi) was directly compared with a human child with similar communicative and general cognitive skills, the human child out-performed Kanzi on list-type sentences (‘Get the apple and the orange and the banana’) but
   the ape out-performed the human on word-order sensitive sentences, where syntax would appear to be the key.

The first two points above indicate similarities between the acquisition process and optimal acquisition environment between apes and humans. It is very unlikely that a human child would acquire normal language skills in a conditioning paradigm, whereas the social immersion paradigm that is routinely and successfully used among humans is successful also with apes. Notably, the most prominent failure among chimp students, Nim Chimpsky (see page 131), was taught within a non-social conditioning paradigm, according to Fouts (1997).

12 Though it may be noted that the dolphins of Herman et al. (1984) acquired apparent Kanzi-like comprehension skills under what was essentially a conditioning paradigm.
Likewise, humans such as Genie (Goldin-Meadow, 1982) who are not exposed to language in early childhood also fail to acquire normal human language. Neither of these points indicates any qualitative differences in the acquisition process between apes and humans. It is rather the similarities that are striking.

The third point concerns the considerable quantitative differences in early acquisition rates between humans and non-humans. The occasionally observed transition to a higher learning rate indicates that a qualitative difference in the learning process may be involved — but that this qualitative difference may be bridged also by some non-humans, after some linguistic threshold has been passed. Pepperberg (2001) indicates that a similar transition occurs also in humans, though at a much earlier stage in ontogeny. Again, there is no clear evidence of any qualitative differences between human and non-human acquisition, merely a quantitative difference in the timing of and possibly requirements for the apparent transition.

The fourth point directly contradicts the commonly held views that syntax is the core of what makes human language unique, and that syntax acquisition is impossible without an innate grammar. If syntax were uniquely human and innate, a human and an ape with similar general communicative skills could be distinguished by the human excelling at syntax-based tasks. Instead, the opposite pattern is found. Again, there is little evidence of any qualitative differences between human and non-human.

Further experiments are needed for firm conclusions. But, even though skeptics remain, the preponderance of the evidence is clearly on the side of presence rather than absence of basic language skills in some non-humans. When a young child displays the same apparent abilities as Kanzi, we do not hesitate to say that the child is acquiring language (Savage-Rumbaugh & Lewin, 1994). Is a double standard sometimes employed?
7.5 Summary

Many attempts have been made to teach language to non-humans, mainly with apes, but also with parrots and dolphins. Projects using appropriate modalities have met with some degree of success.

- **Apes**, mainly chimps and bonobos but also a few orangs and gorillas.
  - **Stage 1: Inappropriate modalities.** Early attempts to teach apes to talk failed for the simple reason that apes are not anatomically and neurologically equipped to produce human-like speech.
  - **Stage 2: Insufficient controls.** The first experiments with appropriate modalities were criticized due to insufficient safeguards against the ‘Clever Hans’ effect and other forms of experimenter cueing.
  - **Stage 3: Better controls and better learning environments.** The next generation of experiments, still on-going, are performed with better methodology, in response to earlier criticism. A significant finding in these experiments, that ought not to have been surprising considering how humans acquire language (Locke, 2002), is that language acquisition in a social context is much more successful than attempts to train animals to use language by traditional conditioning methods. Results from Stage 3:
    - Apes can learn a large number of signs, and use these signs appropriately in a social communicative situation. Established in blind tests that it’s not just a matter of human cueing.
    - Apes can communicate with each other using signs, apparently achieving ‘joint reference’, and getting close to establishing symbolic capacities.
    - One ape can understand English sentences, even with non-trivial syntax, on a level comparable to that of a two-year-old human.

- **Parrots.** One successful experiment, comparable to the Stage 3 experiments with apes.

- **Dolphins.** One successful experiment, where dolphins appear to have acquired the ability to interpret syntax in a controlled setting.

It is still a matter of some controversy how to interpret the achievements of these animals. It is well established that they have some capability to handle language-like communication systems. But is this language?

Of the two major components of human language, words and rules (Pinker, 1999), it would be hard to deny that apes have shown a fair grasp of words. Evidence of ape understanding of syntactical rules is less clearcut. There is little syntax in the spontaneous utterances of apes, at best approaching that of a child at the two-word stage. Some apes, however, do quite well in comprehension tests (as do dolphins), demonstrating their understanding of syntactically non-trivial requests. Their performance on such tests compares favorably with that of two-year-old humans.

It can be concluded that at least some components of language, at least on a primitive level, are not strictly limited to *Homo sapiens*. 
Further reading

CHAPTER 8

LANGUAGE, MIND, AND SELF

What is the mind, and what does it mean to be conscious and self-aware? These are questions with an ancient and distinguished philosophical pedigree, going back at least to Plato and Aristotle, but without universally accepted answers. These concepts are relevant to the issue of language origins and animal language, because it has been argued that language and mind are intimately connected (e.g. Maturana et al. (1995), Jonker (1987)). A necessary connection between mind and language has been proposed in both directions, either with language a necessary prerequisite for conscious thought\(^1\) (Bickerton, 1995; Spangle & Menzel, 1991), or with self-awareness and intentionality a necessary prerequisite for symbolic thought and true language (Sinha, 2001; Zlatev, 2001b).

Our everyday experience of ‘inner speech’ as a major part of our thought processes, our Joycean ‘stream of consciousness’, does support some kind of role for language in the human mind, an argument that can be found in Plato’s *Theaetetus* (Glidden, 1994). But there are many possible variations on the thesis that language plays a role in human thought and consciousness. Carruthers (2002a) offers this classification:

- **Strong forms:**
  - Language is conceptually necessary for thought
  - Language is de facto used for all conceptual thought
- **Weak forms:**
  - Language is necessary for concept acquisition
  - Language serves as a scaffold for thought

Carruthers (2002a) then goes on to propose his own model, in which language is needed for all non-domain-specific thinking, all conceptual connections between modules (cf. Section 5.3.3) in the mind.

\(^1\)The idea that language is required for conscious thought has the remarkable corollary that there must have existed, at least briefly, people with language but without consciousness. Julian Jaynes (1976), reviewed in Mooneyham (1993), takes this idea to its logical extreme, postulating non-conscious people as late as the Bronze Age. While Jaynes’ specific proposal is ludicrous, the reasoning behind it is nevertheless a valid deduction from language as a prerequisite for consciousness.
With an intimate connection between language and thought, it is difficult to avoid a chicken-and-egg problem in both ontogeny and phylogeny — did thought come first, with language as the externalization of thought, or did language come first, with thought as the internalization of language? In both cases it is difficult to account for the origin of both thought and language (Harris, 2001), a problem that played a central role in early discussions of language origins — cf. Section 9.1.

To what extent our thoughts are shaped by language remains an open question. It is certainly true that language is routinely used as a ‘mental tool’ in our thinking — but does that mean that we can only think thoughts for which we have words? Obviously not, as we are demonstrably capable of coming up with new concepts and coining new words for them. The strong over-simplified Sapir-Whorf hypothesis is effectively dead — but more subtle forms of Whorfian influence are pervasive, inserting linguistic relativity into our conceptual structures and thoughts (Pütz & Verspoor, 2000; Bowerman & Levinson, 2001), and even in our perception (Davidoff, 2004). Language provides new tools for conceptualizing and thinking about the world, and while the shape of the tools might not prevent us from thinking certain thoughts, it does affect what’s easier or more difficult to think about. As Bowerman & Levinson (2001) put it: “Language invades our thinking because languages are good to think with.” (p. 584)

Fellows (2000) states that he “believe[s] that non-language-using animals lack minds because they lack beliefs.” (p 598), arguing further, mainly on philosophical grounds, that language is necessary for having beliefs. Pinker & Bloom (1990), Bloom & Keil (2001), and Bogen (1997), on the other hand, argue strongly against normal human language being the basis for thought. Zlatev (2001b) similarly argues against language as the basis for consciousness:

..., language can not be the major cause of (self-)consciousness as is claimed by numerous contemporary theorists (e.g. [...]), since its acquisition presupposes (a degree of) intersubjectivity, which presupposes consciousness. (2001b, p. 6, emphasis in original)

Furthermore, a strong connection from language to mind and thought implies that young children, as well as aphasia patients, are mindless and thoughtless, an implication that does not agree with our experiences, nor with empirical data from pre-verbal children (Hespos & Spelke, 2004; Clark, 2004a; Bloom, 2004b; Mandler, 2004b), from patients with total agrammatic aphasia (Bloom, 2000; Bloom & Keil, 2001; Siegal et al., 2001; Carruthers, 2002a), and even from effectively language-less left-hemispherectomized patients (Bogen, 1997). For that matter, looking introspectively I can state with confidence that not all my own thoughts are language-based — images and other pseudo-sensory patterns form a conspicuous part. And I am in good company here:

Words and language, whether written or spoken, do not seem to play any part in my thought processes. The psychological entities that serve as building blocks for my thoughts are certain signs or images, more or less clear, that I can reproduce at will. (Albert Einstein, quoted in Dehaene et al. (1999, p 970)).
Still, there are reasonable grounds for suspecting some kind of connection between language and mind, at least in the direction from mind to language — symbolic language makes little sense without intentionality and self-awareness. The idea that thought is required for language goes back at least to the Stoics (Everson, 1994a).

Humans are arguably better than other animals both at communicating and at understanding intentionality and mind in others. And even without one being a strict prerequisite for the other, it is pertinent to ask whether our edge in communication is caused by our edge in ‘mind-reading’, or vice versa? Or are both abilities emergent from some underlying general feature of the human mind? Are the differences between human minds and the mental processes of other animals qualitative or quantitative?

8.1 What is Mind — the ‘hard problem’

The mind, as used in this context, corresponds roughly to the non-theological aspects of the soul, the grammatical subject of the “cogito, ergo sum” of Descartes (1637; 1642, and see also Almog (2002)), but it is not easy to define the mind in any stringent way, much less operationalize the concept. See e.g. the reviews of Hofstadter & Dennett (1982) and Ran (1999) for a variety of perspectives. The debate surrounding the brain and mind has close parallels with the classical body-soul duality. Is our perception of being a self-aware mind (or soul) merely an epiphenomenon growing out of various brain activities, or does the mind have an existence beyond mere neuronal patterns in the brain? And if it does, what is its substance, and what is its connection with the material world? Why do conscious experiences — qualia — exist at all? There is something it is like to be a being with a mind (Nagel, 1974) — but where does that subjective feeling come from? This mystery is the essence of the ‘hard problem’ of consciousness, as opposed to the ‘easy problem’ of connecting psychology with physiology, discussed in the next section (Chalmers, 1995; Shear, 1999; Papineau, 2003). The hard problem is indeed very hard, and a solution to it is definitely beyond the scope of this book, which will just briefly touch upon a few aspects of it that may be relevant to the phylogenetic origins of the mind.

Lindahl (1997) discusses the evolutionary implications of different views of the relation between the mind (or ‘mental events’) and the brain (‘neural events’). ‘Mental events’ are essentially our conscious thoughts and decisions, as perceived by ourselves in introspection, and ‘neural events’ are whatever is going on in the brain when a mental event occurs — for example, when I consciously decide to raise my right arm (a mental event), the corresponding neural event is the brain

---

2It is even argued that it is unsolvable (Krakow, 2002).
activity that ends up in motor commands going out from the brain to the arm muscles. Mental events would belong to the *res cogitans* of Descartes (1642), and neural events to his *res extensa*.

Lindahl (1997) distinguishes between three³ possible relations between mental and neural:

- **Epiphenomenal.** Neural events are real, and cause e.g. bodily actions. Our perception that mental events cause actions is an illusion. To the extent that mental events exist, they are caused by the neural events as a mere accidental byproduct. Parallelism is a version of epiphenomenalism where there is a one-to-one correspondence between neural and mental (Krakow, 2002).

- **Interactional.** Mental events are real, and interact with, and can cause, neural events and subsequent physical actions. The interactional view can be further subdivided⁴ (Vanderwolf, 1998):
  - The Aristotelian view, in which the mind is primary. All functions of the body (and brain) are directly due to the mind.
  - The Cartesian dualism, in which the mind and body form a symbiotic system, with bodily functions handled mechanically but higher functions (notably language) handled by the mind (Dilley, 2004). In the Cartesian view, only humans have minds — animals are mindless automata, purely mechanistic.

- **Identity.** Mental events *are* neural events. Our introspective perception of consciousness is a neural pattern, and nothing else (Dennett, 2001; Place, 2000).⁵ Unfortunately, it is very difficult to get any empirical handles on these alternatives. Both Lindahl (1997) and Århem & Liljenström (1997) attempt to apply evolutionary reasoning, with the following basic steps:
  - Humans have minds (whatever they may be).
  - Not all living things have minds, so minds must have evolved somewhere along the human family tree.
  - Minds are complex features, the kind of features that don’t just turn up by accident in evolution, but must confer a selective advantage.
  - Features that don’t *do* anything can’t give a selective advantage.
  - The only alternative in which the mind does do something that may confer a selective advantage is the interactional perspective, in which the mind is real, and materially affects the brain and body. In the other alternatives, the mind does not affect anything else, and so cannot confer an advantage.

³Sampson (1999b) presents a fourth alternative, that mind is a social construction, “distributed among individuals, the texts they produce, the artifacts they create, and the institutions they develop.” (p. 1). However, in my judgement her alternative does not add anything useful to the present discussion.

⁴There are also other possibilities, like the event dualism of Paul Pietroski, where our actions have dual causes, one neural and one mental (Menzies, 2003). But taken at face value, dual causality appears to me incoherent.

⁵Place (2000) argues in favour of identity for mental/neural *processes*, but in favour of epiphenomenalism for mental *states*.
Both authors conclude that this argument makes a strong but not compelling case for the interactional view. What makes the argument inconclusive is that we cannot rule out the possibility that what is really selected for is a particular type of complex brain activity, of which our perception of having (being?) a mind is a mere byproduct, a spandrel (Edelman, 2003).

An argument in the other direction, against the causal effectiveness of the mind, comes from the fact that there are clear causal links from neural events to mental events. At a basic level, it is uncontroversial that there is a causal link from light absorbed in the retina (a neural event) to our perception of light (a mental event). But on a higher level, the fact that our minds can be affected by brain lesions, and by mind-altering drugs (through well-defined neurophysiological mechanisms), is good evidence that the mind is causally dependent on the brain, and not vice versa.

Also the timing of events is evidence against a causal link from mental to neural. It is reported by Georgieff & Jeannerod (1998), Deumert (2003), Libet (2004), and Wegner (2002) that mental events significantly lag behind the corresponding neural events — e.g., our conscious decision to do something takes place a few tenths of a second after the initiation of the neural premotor processes that prepare for the act. Together, the mind-altering and the timing rule out the Aristotelian view, and weakens the case for dualism.

8.2 What is mind — the ‘easy problem’ — and do animals have it?

We now leave the ultimate nature of Mind aside, and proceed with more tractable aspects of consciousness, that are empirically accessible also in non-humans. These aspects include both the neural events associated with consciousness, and the inferences about mental events and self-awareness that may be drawn from observed behavior.

The neural events that indubitably take place while we introspectively perceive mental events, are considerably easier to operationalize and study experimentally, than the mental events themselves. But even here there are vexing difficulties in correlating neural and mental events — neural events can be measured, but how do we know that a mental event has taken place in somebody else’s head? Vanderwolf (1998) discusses this problem at some length, concluding simply that we can’t know. We can only judge by external behavior (including verbal reports of purported mental events), but this may not be sufficient — see the Chinese Room parable of Searle (1980) for an extended argument against inferring mental events from external behavior.

6See Lubinski & Thompson (1993) for a review of the communication of mental (or ‘private’, as they call them) events, in both humans and other species.
8.2.1 Theory of mind

In everyday life we do not follow the logic of Searle’s (1980) argument — instead we routinely infer mental events from the external behavior of others, despite the epistemological problems involved. We assume that other humans have minds, and have mental processes that resemble our own, and we use our understanding of mental processes to predict their behavior. This ability to infer mental events is important enough to have a name of its own — we are said to possess a ‘theory of mind’ (Premack & Woodruff, 1978). Without a theory of mind — an understanding that other humans have minds like our own — we would be unable to function socially. An innate theory-of-mind module has been postulated, critically reviewed in Gerrans (2002). While the evidence is not conclusive the results of recent neuroimaging studies are suggestive — specific locations in the brain are activated in theory-of-mind tasks but not in ‘mindless’ tasks that are otherwise similar (Zimmer, 2003; Gallagher & Frith, 2003). Theory-of-mind defects have been invoked as the main cause of autism (Baron-Cohen et al., 2000), supported by the fact that autistic children who fail theory-of-mind tests do perfectly fine on ‘mindless’ tests of comparable difficulty (Garfield et al., 2001). Autism has a clear genetic component; interestingly enough, one putative ‘autism gene’ maps to the same stretch of DNA as one putative ‘language gene’ (Stokstad, 2001) — but not too much should be made of this, as the genetic links remain weak, and the common stretch of DNA is long enough to contain multiple genes. Further research has not shown any evidence of defects in the ‘language gene’ FOXP2 among autists (Wassink et al., 2002; Newbury et al., 2002).

There is, however, a severe problem both for the hypothesis that we have an innate theory-of-mind module, and for the hypothesis that a defective such module is the primary cause of autism, and that is the observation that initially healthy children who for some reason are deprived of early social interaction and linguistic input, display theory-of-mind delays reminiscent of autism. Deaf children of hearing (and non-signing) parents is a case in point (Garfield et al., 2001; Peterson & Siegal, 2000). This would seem to indicate that a theory of mind grows out of early social and verbal interactions, something which children sharing no modality with their parents might be deprived of, and which autists don’t participate normally in (Peterson & Siegal, 2000).

Children with Williams syndrome (see page 105), on the other hand, do just fine on theory-of-mind tests, better than would be predicted from their general retardation (Garfield et al., 2001). It can be noted in this context that Williams children typically do just fine in social interactions in general, supporting the point of the previous paragraph.

---

7The Turing Test (Hofstadter, 1982) is an interesting generalization of this inference, to non-human and even non-living minds.

8There are several other hypotheses concerning the causes of autism; see e.g., Frith (2003), Parisse (1999), Baron-Cohen (2003), and Siegal & Blades (2003).
The theory of mind is relevant for language origins, not least because language acquisition would be highly problematic without an understanding of the communicative intent of others, which requires a theory of mind (Tomasello, 1999a). This argument should not, however, be pushed too far, as there exists a subset of autistic people, arguably lacking a full theory of mind, who nevertheless acquire language (Glüer & Pagin, 2003). Hurford (2003b) includes some aspects of theory of mind in his list of exaptations (or pre-adaptations as he calls them — cf. p. 18) needed for language readiness, under the heading ‘pre-pragmatic capacities’.

Whether somebody possesses a theory of mind is an experimentally tractable question. And it is reasonable to assume (though not a stringent certainty) that a being with a theory of mind also possesses a mind of its own — how else can it infer the presence of minds in others, if it doesn’t have one of its own to compare with? Never mind the philosophical issues for the moment — this argument needs only a mind in the sense that we perceive ourselves as having, whatever its ontological status.

Experiments testing for the possession of a theory of mind have mainly been conducted on human children, in whom the existence of a mind is taken for granted (Feinfeld et al., 1999). It has been found that children develop a theory of mind through several regular stages at roughly predictable ages, much like language acquisition (Lee et al., 1999), starting with eye contact detection already in the first week of life (Farroni et al., 2002), and then going on to an important communicative breakthrough towards the end of the first year of life (Zlatev, 2001b; Tomasello, 2003). Of course, the acquisition process is beset by the same epistemological issues that were discussed above (Montgomery, 1997) but normal children nevertheless manage to acquire a theory of mind in about the same time it takes for them to acquire language. This may be taken as a sign that there is a connection between language and theory of mind — but a patient with agrammatic aphasia, totally lacking syntax, nevertheless had a full theory of mind (Bloom, 2000).

More interesting in this context are the experiments that have been conducted on beings that do not normally acquire language, and whose possession of a mind is in some doubt. After the seminal paper of Premack & Woodruff (1978), a number of groups have attempted to determine whether non-human primates have a theory of mind. The experiments themselves are very similar to the experiments done with pre-verbal children, testing for behavior based on inferences about the mental state (knowledge and intentions and motivations) of other beings, either conspecifics or experimenters. Heyes (1998) reviews the field, as it was twenty years after the work of Premack & Woodruff (1978), and concludes that it is very difficult to design experiments which can firmly exclude non-mind interpretations, and that the

---

9But Garfield et al. (2001) find that the pattern of development is not universal across cultures.

10Cf. the change in child-directed communication around the same time (Rivero, 2004), discussed on page 182.
data available are inconclusive. In a later review, Pennisi (1999) is more positive, unlike Tomasello (2000c) and Hauser (2001). Tomasello and colleagues, however, have recently changed their opinion due to new evidence, formally recanting their former stance in Tomasello et al. (2003a): “...we are now convinced that at least some non-human primates [...] do understand at least some psychological states in others.” (p. 153). For a more detailed discussion of what apes can and cannot do in this context, see Tomasello et al. (in press).

As research with human children has shown, a theory of mind is not a monolithic module that you either have or not (Tomasello et al., 2003b; Gärdenfors, 2001) — instead it is acquired gradually, with more and more abilities to reason about the mental states of others being added. Apes (and also dolphins) have passed tests for the abilities that children acquire first, and outperform autistic humans on these tests (Dunbar, 2001), but just as regularly fail tests for the ‘higher’ abilities, such as an understanding of false beliefs11 (though the results of Hare et al. (2001) are intriguing), that children don’t acquire until age five or so. But both apes and monkeys have considerable ‘social intelligence’ (Anderson, 1998), and a thorough understanding of the politics of a primate tribe (de Waal, 1998; Strum, 1989), which is difficult to explain without assuming an understanding that the others in the tribe are also active participants. But political astuteness is not easily amenable to experiments in laboratory settings, unlike lower-level functions that are more experimentally tractable.

One such lower-level function is gaze-following — does an ape (or a child) notice in which direction somebody else is looking, and infer that something interesting must be in that direction? This is one of the first mind-related abilities acquired by humans, before age 18 months, and several experiments have been performed with apes and monkeys, as well as with other animals.

Simple gaze-following, for which non-mentalistic explanations appear most plausible, can be found even in some birds, who avoid entering their nest when a predator is looking in that direction (Watve et al., 2002). Ordinary domestic dogs are quite adept at reading human signals, including following the gaze of humans.12 These abilities have apparently evolved in dogs during the domestication process, as they are lacking in wolves — presumably dogs good at ‘reading’ humans were more successful pets (Agnetta et al., 2000; Hare et al., 2002; Miklósi et al., 2004; Cooper et al., 2003). But here as well, non-mentalistic explanations cannot be excluded (Byrne, 2003).

---

11Bloom & German (2000) argue that the false belief task is not a relevant test for the possession of a theory of mind, since passing the test entails other abilities as well. The argument of Bloom & German (2000) is supported by the discovery of Abu-Akel & Bailey (2001) that success on false-belief tasks is strongly dependent on the presentation of the task, notably in the degree of symbolic thinking required (cf. the ontogeny of symbolic thought, discussed on page 228).

12They are also able to learn rapidly to understand large numbers of words for objects (Kaminski et al., 2004; Bloom, 2004a).
The results from primates are mixed, in that some experiments, e.g. Anderson et al. (1996), failed, whereas others, e.g., Tomasello and associates (1998; 1999), met with better success. It is conceivable that the difference can be attributed to the fact that Anderson et al. (1996) worked with rhesus monkeys, whereas the strongest successes of Tomasello et al. (1999) have been with chimpanzees. But Ferrari et al. (2000) found clear evidence of gaze-following with adult (but not juvenile) pig-tailed macaques, close relatives of the rhesus monkeys that Anderson et al. (1996) failed with, and Tomasello et al. (2001) report observations of gaze-following in both rhesus monkeys and chimps in the same experimental setup. On the other hand, Povinelli et al. (2000) also find evidence of chimpanzee gaze-following, but do not attribute it to a theory of mind — the pattern of behavior is such that non-mentalistic explanations cannot be excluded. Tomasello et al. (1999) conclude “The degree to which chimpanzees have a mentalistic interpretation of the gaze [...] of others is still an open question.” (p. 769, but see also Tomasello et al. (2003a)).

Tomasello et al. (in press) cite evidence that chimpanzees have some understanding of the intent of others, recognizing what they are attempting to do even if they fail. The chimps reacted differently to someone who tried but failed to give them food, compared with someone who could have given them food but refused to do so. They also appeared to have some understanding of accidental as opposed to intentional actions.

Closely related is the question of whether apes are aware of what others know and do not know, and what others can and cannot see. Such awareness has been difficult to demonstrate in artificial setups or human-style social settings, which may be due to the tasks being very different from typical chimpanzee social interactions. There is a fair amount of anecdotal evidence of primates employing tactical deception in the wild (Whiten & Byrne, 1988), which clearly requires a theory of mind (Gärdenfors, 1996), but this has been difficult to reproduce under controlled conditions. In more chimp-style competitive tasks Hare and associates (2000; 2001) find evidence that chimps do understand what other chimps do and do not see and know, and that they use this understanding tactically in social interactions.13 Chimps do better at cognitive tasks in general, when the tasks are presented in a competitive context (Hare & Tomasello, 2004).

The success with the apes in competitive experiments is interpreted by Tomasello and Hare and associates as strong evidence that chimps have at least some notion of each other’s minds — “Chimpanzees understand psychological states...” begins the title of Tomasello et al. (2003a) — but Povinelli and associates, the other major research group in this field, strongly disagree. The ensuing exchange in Povinelli & Vonk (2003), Tomasello et al. (2003b), and Povinelli & Vonk (2004),

13Hare et al. (2003) have also repeated exactly the same experiment with monkeys, who failed in the crucial tasks where chimps succeeded.
is illuminating mainly in the philosophical assumptions that are revealed. Some of
the relevant epistemological issues are reviewed in Schilhab (2002).

But the fact that apes succeed in competitive tasks, but fail in simple cooperative
tasks that appear obvious and natural to humans, is interpreted by Tomasello et al.
(in press) as a sign of another type of difference between us and the other apes. It
is argued that, while all apes have the basics of a theory of mind, humans alone are
motivated to cooperate and share emotions and activities, and use their theory of
mind for collaborative rather than competitive purposes.

8.2.2 Self-awareness

Another aspect of the mind is self-awareness, the insight that cogito, ergo sum.
This can be regarded as having included in my inner representation of the world a
representation of my own inner representation. According to Gärdenfors (1996),
this presupposes a theory of mind (a theory of other minds, that is) — ‘you-
awareness’ comes before ‘me-awareness’.

Self-awareness is another concept the presence of which is experimentally dif-
ficult to assess, but Gallup (1985) proposed the following test:

1. The subject is placed in a room with a mirror, and is given time to get ac-
quainted with how a mirror works.

2. The subject is rendered unconscious

3. A marker that cannot be felt is placed on the body of the subject, in a place
where it cannot be seen directly. A drop of paint on the forehead is commonly
used.

4. When the subject wakes up, there is only one way for the subject to discover
the mark — by using the mirror. Operationally, this is measured by observing
whether the subject tries to scratch or rub away the mark, before and after
having seen it in the mirror, and if the rate of mirror use increases when the
mark is detected.

5. Scratching a mark on your own body, after seeing it in the mirror, implies that
you understand that the image in the mirror is an image of yourself. And this in
turn implies that you understand that you have a self — which is pretty much
the definition of self-awareness.

The mirror test for self-awareness has been used extensively in the years since
Gallup proposed it, both by Gallup himself, reviewed in (1998), and many others.
Humans pass the test from the age of 2 or so (Hauser et al., 1995). Concerning
chimpanzees, there is near-unanimity that they succeed in the mirror test (Lin et
al., 1992; Kitchen et al., 1996; Povinelli et al., 1997, among others), with about
half of a total of 163 tested chimpanzees apparently recognizing their mirror-image
(van den Bos & de Veer, 2000). The success of other apes is more mixed, with con-
siderable doubt remaining about the abilities of gorillas in particular (Shillito et al.,
but it appears that at least some gorillas and orangutans pass the test (Taylor-Parker et al., 1999; Hauser et al., 1995; Tobach et al., 1997). It is interesting to note that to date, the only gorillas to pass the test have been human-reared participants in language-learning experiments, such as Koko (see page 135), whereas gorillas growing up without intimate human contact consistently fail (van den Bos & de Veer, 2000).

Outside the hominoids, successful mirror tests have been reported for cotton-top tamarins (Hauser et al., 1995) and recently also for dolphins (Reiss & Marino, 2001) and killer whales (Delfour & Marten, 2001). In humans, brain scans have located several brain areas that are activated specifically by self images, but not by viewing images of others (Kircher et al., 2001) — replicating this experiment with apes would be interesting.

Heyes (1996) remains skeptical also of the mirror test, as she is of theory-of-mind experiments (Heyes, 1998), mainly invoking methodological concerns. Both Mitchell (1995), Povinelli et al. (1997) and van den Bos (1999) address her concerns, and the latter two present new data which firmly exclude the alternative interpretation proposed by Heyes (1996). In conclusion, it appears well established that at least some non-humans can recognize themselves in a mirror, from which some level of self-awareness can reasonably be inferred — but see also Schilhab (2004) for a deeper discussion of the methodological and philosophical issues involved.

The issue of self-awareness raises the issue of what the self is. Disregarding philosophical and theological complications, introspection tells me that a large part of what makes me me is my memories, and my self-awareness is to a considerable extent built from my awareness of my memories. Lou et al. (2004) have identified a brain location in the parietal cortex associated with such self-memory retrieval.

Animals of all kinds certainly have memories — but are they consciously aware of what they remember? In an ingenious sequence of experiments, Hampton (2001) has determined that monkeys are. Rhesus monkeys were faced with the task of remembering which of a set of test pictures they had seen before — but they were given the opportunity to refrain from taking the test, with suitable incentives so that the wise course was to take the test when they remembered, and decline otherwise. Rational choice here is possible only for beings who are aware of their memories, and able to ‘look at’ their memories and evaluate if they’re good enough for the test. The monkeys did choose rationally, which can reasonably be interpreted as their being consciously aware of their memories (Hampton, 2001; Griffin, 2001), which opens the possibility of their building a sense of self in the

---

14 The tamarin is a small South American monkey, only distantly related to us hominoids. Given that African monkeys, much more closely related to us, consistently fail the mirror test (van den Bos, 1999), the success of the tamarin is rather odd, indicating parallel evolution.

15 Note that ‘memory’ is not a homogeneous concepts — there are several types of memories, not all of which are connected with awareness (Schachter, 1998; Clark & Squire, 1998).
same way as we do. There is some evidence of episodic memory also in gorillas (Schwartz et al., 2004).

Gusnard et al. (2001) have found (in humans) a specific brain area, in the medial prefrontal cortex, that is activated in the kind of introspective tasks that the monkeys above were engaged in — to the extent that the self is a matter of brain activity, this may be its location in the head. Similarly, some progress has been made in identifying what it means, neurally, to pay conscious attention to something (Stryker, 2001; Pascual-Leone & Walsh, 2001; Thier et al., 2002; Taylor, 2002b; Koch, 2004), and Lau et al. (2004) have also found apparent neural correlates of intention, as in consciously deciding to make a movement. Interesting in this context is that none of these neural correlates of various conscious activities have been found to be limited to humans (Griffin & Speck, 2004).¹⁶

In conclusion, it does appear as if apes, and possibly some monkeys, do show the external signs of possessing at least a subset of the features that we humans have, and that we label as our minds. The features detected in apes correspond to those that develop early in human ontogeny — this may be interpreted as a case of ontogeny recapitulating phylogeny (Povinelli, 1993). The mind is not a monolithic entity that one either has or not – there are degrees of mindedness and intentionality, something to keep in mind if we are to “characterize the practical abilities and perspective of nonhuman animals correctly, as neither too rich nor too impoverished.” (Hurley, 2003, p. 252).

But there is far from total consensus on the issue of ape minds. Part of the lack of consensus concerns the standards of proof (see e.g. Povinelli & Vonk (2003) versus Tomasello et al. (2003b)) — evidence that would be sufficient to conclude that a child has a theory of mind (or has a mind), is not regarded as sufficient in the case of a chimpanzee (Griffin, 2001). This may be reasonable, as each one of us personally has first-hand evidence of a human child growing up to a being with a theory of mind, but lacks similar first-hand evidence for non-humans, but it places an unreasonably heavy burden of proof on the proponents of ape minds. Should we accept or reject ape minds based on the preponderance of the evidence – or should we reject all such notions as long as the slightest possibility of a non-mental explanation remains?

One may also consider whether lack of mind is really the appropriate null hypothesis in ape studies:

Isn’t it far more economical to assume that if two closely related species act in a similar way, the underlying mental process is also similar? If wolves and coyotes were being compared, there would be immediate agreement about that. Why should we adopt another logic when comparing chimpanzees and humans? (Boesch, 2001, p. 526, but see also Povinelli (2004)).

The situation resembles the case of ape language acquisition, where the burden of proof is similarly lopsided (see page 140).

¹⁶Griffin & Speck (2004) is also a good recent review of the evidence of animal consciousness in general.
The presence of some aspects of mind in apes does not support the view that language is required for thought and consciousness. Instead, the evidence indicates a gradual evolution of the mind, possibly along the lines sketched by Povinelli (1993), with metacognition, reviewed in Nelson & Rey (2000), as the central theme — cognition about cognition, the ability to think about one’s own thoughts, is regarded as the core of the self-aware mind (Vaneechoutte, 2000). A key concept here is attention, directing the metacognition at some particular underlying cognitive event (Posner, 1994; Koch, 2004). Churchland (2002) argues for self-representation as the relevant complex brain activity, with the ability to distinguish between self and non-self as the evolutionary roots of consciousness, whereas Edelman (2003) instead focuses on reentrant interactions enabling the brain to perform higher-order discriminations. A related conception of the mind is the ‘higher order thought’ hypothesis of Rosenthal (2000), reviewed in Gennaro (2004), but Dennett (2001) paints a different picture, where consciousness is the result of a ‘political’ battle for influence in the brain between competing neural processes.

8.3 Summary

— *The hard problem* remains hard. But clear evidence of causal links from neural to mental, and the temporal ordering of events with neural coming before mental, makes it difficult to hold on to theories involving mental primacy.

— *Theory of mind.*
  - Theory of mind is not a monolithic capacity, but is built from several subcomponents.
  - Humans acquire a theory of mind gradually from birth to age 5 or so. Social interactions play a prominent role in acquisition.
  - Insufficient evidence concerning innate theory-of-mind modules.
  - Theory of mind in non-humans:
    - Apes pass some theory-of-mind tests, but not all.
    - No consensus whether this is due to their having a *bona fide* theory of mind, or if non-mentalistic explanations remain tenable.
    - No consensus whether the appropriate null hypothesis is that apes resemble their human cousins, or that they resemble mindless animals.
    - Dogs are remarkably adept at reading human signals, often better than apes. Have dogs adapted to their peculiar ecological niche as our pets?

— *Self-awareness.*
  - Gallup’s mirror test is commonly used to assess self-awareness.
  - Humans pass it from age two or so.
  - Many apes and a few other mammals pass it as well.
  - The situation similar as for theory of mind above — the performance of apes equals that of young children, in whom we infer the presence of a mind. But
no consensus on whether the same inference from the same evidence is valid for non-humans.

- **Memories.**
  - Memories important component of self.
  - Some non-humans appear to be introspectively aware of their memories.

- **The mind in the brain.**
  - Brain scans have found neural correlates of various consciousness-related mental events.
  - Similar brain activity takes place in non-human primates in similar tasks.

- **Language and mind.**
  1. Humans without language are not mindless, nor unthinking.
  2. Some evidence of mind in language-less non-humans.
  3. Not all human thought is language-based.
  4. 1 & 2 & 3 ⇒ Language cannot be a prerequisite for mind and thought.

- **Mind and language.**
  1. Understanding communicative intent important in language acquisition.
  2. Social interactions important in language acquisition.
  3. Theory of mind would be very helpful for 1 & 2.
  5. 3 & 4 ⇒ Theory of mind useful but not strictly necessary for language.

**Further reading**


The previous chapters have all dealt with various background material needed in order to understand the constraints on language evolution hypotheses. In this and the following chapters, the focus will be on the main issue itself — why and how and when did the human language capacity evolve among our ancestors? There are two main issues in explaining the evolution of any feature (Byrne, 2000):

- Historical: at what time, and at what point in the family tree, did different aspects of language appear?
- Causative: what were the selective advantages that drove the evolution of language, and what evolutionary precursors did it evolve from?

The causative issue is the main focus of Chapter 10 and Chapter 11, with historical data used mainly to constrain causative hypotheses. Possible selective advantages are discussed in Chapter 10 and possible evolutionary precursors in Chapter 11. The main thrust of the current chapter is to clarify the structure of the problem.

It is clear from the previous chapters that there is much that we simply do not know about the human capacity for language, certainly concerning its history, but also concerning the details of its implementation in modern humans. It is far from well established exactly how and where the human brain processes language, and the links between linguistic theory and neurological observables are tenuous at best. This means that firm conclusions will be difficult to achieve.

A reasonable starting point in the analysis of the evolution of language, is the last common ancestor of us and the chimpanzees. Presumably this ancestor had roughly the same capabilities and exaptations that modern chimpanzees do, so what needs to be explained here is how we went from chimpanzee-like to human-like linguistic abilities, in less than ten millions years. The principal questions to be answered here are the two that Bickerton (2001) succinctly express as “How did meaningful units (words or signs) evolve?” and “How did syntax evolve?”(p. 583). All else is ancillary.

---

1The linguistic abilities of chimpanzees are not negligible, as shown in Chapter 7, but we are concerned here only with the capabilities that humans have but chimps lack, notably the universal acquisition of and habitual use of a rich language with complex syntax.
This removes quite a few areas from consideration, notably the entire sensory system — as shown in Chapter 5, the senses of an ape are perfectly adequate for language perception already. Likewise, the apparent capacity for at least protosymbolic thought (see page 131) and rudimentary self-awareness (see page 154) among apes show that these two areas also can be dissociated from the origin of language.

At the opposite extreme, those unique human features that are exquisitely adapted for language, notably our vocal tract, cannot be invoked as explanations for the evolution of language either — language must have been in use before natural selection had any reason to adapt the vocal tract for it. So the vocal tract can be disregarded as well, at least in the early stages of language evolution. The vocal tract adaptations can, however, be used to constrain the time frame of speech origins.

9.1 Historical background

The issue of language origins has been discussed by philosophers for as far back as we can trace the history of philosophy. The contrast between speaking humans and dumb beasts has played a central role in this discussion, and possible scenarios for a transition from one state to the other were proposed well before Darwin firmly anchored our origins with apes rather than angels. There are many still-current hypotheses that have precursors in previous centuries.

Aristotle saw language as a major distinction between humans and animals, and regarded our ability to communicate thoughts as the key difference between our language and animal cries, but he never put together a coherent proposal for the origins of language (Everson, 1994a). Epicurus, writing around 300 B.C., may have been the first to do so, with a stimulus-response scenario in which external events caused people to utter certain sounds, a specific noise in response to each type of stimulus. Different people had the same innate pairing of stimuli and response-noises, setting the stage for joint reference to emerge (Everson, 1994b).

More serious proposals concerning language origins came forth from several Enlightenment philosophers in the 18th century, when both the essentialism of classical times, and the Christian doctrine of divine creation, began to lose their hold. Leibniz (1710) was one early pioneer, with his lingua adami with onomatopoetic roots. Condillac (1746) worked out a more detailed and influential scenario of language invention (Wells, 1987; Coski, 2003; Taylor, 1997). In Condillac’s scenario, language starts with gestures (cf. Section 9.6 below). The first gestures were actually not intended for communication, just natural reactions to a situation, but were sufficiently self-explanatory (iconic or otherwise) for viewers to understand. Condillac’s prime example involves somebody reaching out for an object barely out of reach, and somebody else noticing this and giving it to him.
Over time, the sender first notices the receiver’s reaction, and then starts exploiting it. The gesture ceases to be an actual attempt to reach out and becomes stylized, eventually conventionalized. Pretty much the same process can be observed in the ontogeny of many children, where an attempt to reach up to Daddy to be picked up eventually becomes conventionalized as a ‘pick-me-up’ gesture with minimally raised arms.

Syntax, in Condillac’s scenario, emerges from the natural iconicity of the gestural-mimetic depiction of sequences of events. Reid (1765) further emphasizes the role of mimesis and pantomime (cf. page 175), drawing a connection between language and art.

The relationship between language and thought, assumed at the time to be intimately connected, is a problem in this scenario, as it appears to require non-negligible powers of thought before language gets off the ground. Condillac (1746) is criticized on this point by both Rousseau (1755) and Müller (1866), even though Condillac (1746) circumvents the problem quite elegantly. In Condillac’s solution associative learning, together with what amounts to memetic evolution of word-meaning associations, is sufficient for joint reference to emerge (Wells, 1987).

Rousseau’s own scenario for language origins (1781) posits a first stage with both gestures and cries, gestures driven by need and cries by passions. Rituals and songs had a key role in the transition from cries to more language-like systems (Gans, 1999a).

Monboddo (1774) is the first to explicitly emphasize the social aspect of language origins, though it is implicit in both Condillac (1746) and Rousseau (1781). In Monboddo’s scenario, there are four original types of communicative self-explanatory signs:
- Facial expressions
- Painting
- Emotional cries
- Imitative iconic sounds

Imitation is central for Monboddo in the further development of language, though he is not able to explain the transition from iconic noises to arbitrary conventionalized words. Monboddo actually got some of his inspiration from first-hand observations of orangutans, which he regarded as prelinguistic humans in the natural state (Limber, 1982).

Herder (1772) is the next key figure in the history of language origin hypotheses, with his seminal *Abhandlung über den Ursprung der Sprache*. In it, Herder is quite critical of Condillac and other Enlightenment thinkers on this topic, perhaps unfairly so (Wells, 1987).

Herder starts with a discussion of ‘natural language’, the emotional cries that humans and animals have in common, but concludes that these are irrelevant for the origin of ‘real’ human language. Instead, he tries to identify a key difference between humans and animals, that can explain why we speak but they don’t. The
key that he proposes is that animals are specialists, guided by instinct, with skills limited to a narrow context, the animal’s ‘Kreis’, to use Herder’s German word, whereas humans are not so limited with our generalist minds and lack of instincts. Not all animals have equally narrow skills, there is a continuum in Kreis size, with humans at one extreme of the distribution with an effectively infinite Kreis. The special faculty that we use to handle our Kreis is our powers of reason. It is our reason which, in the view of Herder (1772), makes language possible, or even inevitable.

Our reason, together with our lack of instincts, enables us to regard objects with a kind of detached curiosity. We can see a lamb (Herder’s example) and neither pounce on it like a wolf nor mount it like a ram, just watch it curiously — which leads to our giving it a name! It appears that Herder assumes that purportedly instinctless humans nevertheless have some sort of ‘naming instinct’, an innate desire to label objects with sounds.

Despite the apparent continuity between animals and humans in the scenario of Herder (1772), Herder (1784) explicitly denies the possibility of evolutionary transitions between species.²

Herder (1784) is also troubled by the chicken-and-egg problem of language and thought mentioned above, but fails to find an acceptable solution. It is also unclear in his scenario where our ‘naming instinct’ comes from, and how to proceed from the one-word stage it leads to, on to full human language. A further problem is that Herder’s scenario is basically asocial — a single man invents ‘words’ on his own (Gans, 1999b).

During the 19th century, linguistics gradually developed as a science. In some respects it was consciously modeled on more established sciences, with focus on the discovery of general laws. In such a context, there was little room for speculation on language origins, and the subject gradually fell out of fashion among linguists. Müller (1866) did discuss possible ideas for language origins, notably the ‘bow-wow’ and ‘pooh-pooh’ theories, basing language on onomatopoeia and emotional cries, respectively — but even so he explicitly rejected the possibility of a transition from animal to human:

Language is our Rubicon, and no brute will dare cross it. [...] no process of natural selection will ever distill significant words out of the notes of birds or the cries of beasts. (Müller, 1866, p. 354, quoted in Limber (1982)).

In 1866, the topic was even banned by the Linguistic Society of Paris, a ban that was for all practical purposes in force for more than a century, with a handful of exceptions (Geiger, 1868; Wundt, 1921). Mario Pei (1965), in his popular overview of linguistics, reports the state of the art of language-origins studies at the time on just three pages (pp. 24-26), with a style and choice of examples clearly

²This was of course well before Charles Darwin was born, but evolutionary ideas were nevertheless discussed at the time — see e.g., Erasmus Darwin (1795) — and Monboddo (1774) may have considered the possibility in connection with language origins (Wells, 1987).
indicating that he does not take the issue seriously. In addition to the ‘bow-wow’ and ‘pooh-pooh’ theories mentioned above, he brings up and dismisses the ‘ding-dong’, ‘yo-he-ho’, ‘sing-song’, and ‘ta-ta’ theories, and even manages to find a 20th-century defender of the Tower-of-Babel story (Italian linguist Trombetti, no reference given).

There has been a remarkable amount of resistance to ideas about language evolution also among prominent present-day linguists, notably Noam Chomsky (1988; 1990), but he is far from the only one; see e.g., Piattelli-Palmarini (1989; 2000), Bickerton (1995; 1998), Fodor (1998), not to mention the nit-picking in nine papers by Botha (1997a; 1997b; 1998a; 1998b; 2000; 2001a; 2001b; 2002a; 2002b). A central part of Chomsky’s theory of language is a universal grammar, innate in all humans (1965; 1986), further discussed in Section 9.7 below. This in itself does not exclude the possibility that language evolved — on the contrary, both Pinker (1994) and Dennett (1995) see a natural connection between the innateness of and the evolutionary origins of language — but Chomsky emphasizes the gulf between his system of universal grammar, and any other system of communication, and is openly skeptical about the power of Darwinian evolution to bridge the gap, with statements like:

Evolutionary theory is informative of many things, but it has little to say, as of now, of questions of this nature [such as the origin of language]. [...] In the case of such systems as language or wings it is not easy even to imagine a course of selection that might have given rise to them (Chomsky, 1988, p. 167).

Several more Chomsky quotes in the same vein can be found in Pinker & Bloom (1990).

There are several possible explanations for the persistence of the resistance to evolutionary explanations within linguistics. The Chomskian attitude to language is basically Cartesian (Chomsky, 1966), with the human language faculty being human-specific, monolithic, and part of our innate essence (Chomsky, 1990; Mueller, 1996). But Chomsky’s essentialism has roots also in Plato (Chomsky, 1988; Bates, 2003). Some of the roots of the resistance to evolutionary thinking among linguists may also be traced to the structuralist tradition of Saussure (1916), the essentialist character of which effectively excludes evolutionary explanations (Bichakjian, 2002). And even though Chomsky is often regarded as reacting against structuralism, its essentialism is retained in the Chomskian paradigm (Croft, 2002).

But other philosophical traditions may also pose obstacles, notably the one stretching from the *tabula rasa* of Locke (1689) to 20th-century behaviorism (Skinner, 1957), in which all behaviour is learned and nothing whatsoever is innate,

---

3In a more recent publication (2002), Piattelli-Palmarini is more ambivalent, not to say contradictory, in his attitude towards evolution, saying both “...natural languages have been shaped by the haphazard biological evolution of the human brain.” and “Human natural grammars are trimmed to the barest essentials, ..., rather than along the whimsical contours of evolution.” in the same article (both quotes from Piattelli-Palmarini (2002, p. 129)).
leaving precious little room for biological evolution of behaviour, including lan-
guage.

Other possible explanations for why linguists have avoided the topic of lan-
guage evolution are reviewed in Newmeyer (2003b).

In recent years, however, some linguists have started to take the evolution of
language more seriously. Pinker & Bloom (1990) is one seminal paper, followed
up by Pinker’s (1994) popular book on the subject. But work on the evolution of
language within a Chomskian paradigm remains problematic (Uemlianin, 1999).
Nevertheless, Pinker (2000) has a point in that “[t]he study of the evolution of
language, ..., has returned to respectability.” (p. 441), an observation also made
by Carstairs-McCarthy (1996). And some former opponents of Darwinian ex-
planations appear to be switching sides, notably Derek Bickerton, co-authoring
a book with William Calvin (see section 5.1.3) with the significant title Reconc-
ciling Darwin and Chomsky with the human brain (Calvin & Bickerton, 2000).
Even Chomsky himself has softened his stance recently, most notably in Hauser &
Chomsky & Fitch (2002a), a highly significant article both for its combination of
authors, for its message of interdisciplinary cooperation, and for the “firm support
of the adaptationist program” (p. 1574) that permeates the paper, contra practi-
cally everything Chomsky has previously written on language origins. Now, in
the 21st century, the “scientific study of the evolution of language has apparently
come of age.” (Fitch, 2002b, p. 278).

9.2 Dimensions of language evolution hypotheses

There are several dimensions along which to classify hypotheses about how we
acquired our language capacity. Among the more important ones are:5

- Adaptation vs. spandrel
- Early vs. late
- Gradual vs. sudden
- Speech first vs. gestures first

---

4 Hauser is an animal communications expert, and Fitch an expert on the comparative anatomy of the
vocal tract. The works of both, e.g., Hauser (1997) and Fitch (2000b), are extensively cited elsewhere
in this book.

5 Hauser & Chomsky & Fitch (2002a) propose a related hypothesis space, but with three dimensions:

- “Evolved as a unique adaptation for communication vs. some other computational problem”, corre-
sponding to my ‘Adaptation vs. spandrel’.
- “Gradual vs. saltational evolution”, corresponding to my ‘Gradual vs. sudden’.
- “Uniquely human vs. shared with other species”, which I have chosen not to include as an in-
dependent dimension. Some aspects are covered in my ‘Innate vs. learned’ dimension. (Hauser,
Chomsky & Fitch, 2002a, Fig. 3, p. 1571)
Hypotheses of language origins

163

9.3 Adaptation vs. spandrel

Evolution is a strong force for shaping our bodies and minds. But this does not mean that every single feature has been shaped by natural selection to perfection. Many aspects of our bodies may have evolved for some other use than their current function (exaptations), or may simply be accidental byproducts (spandrels) or leftovers (vestigial), with no particular adaptive function in themselves (Gould & Lewontin, 1979; Gould, 1997, and see also page 18 above). Male nipples are a case in point — female nipples are obviously adaptive, but it is likely that males have nipples, not because they are of any use, but simply because both male and female embryos follow the same developmental program, and it’s embryologically simpler to give nipples to both of them than to just one (Gould, 1992).

From the point of view of language, the difference between exaptation, spandrel, or vestigial feature, really doesn’t matter — neither of them evolved for language. Whether they have, or had, some non-linguistic use is beside the point.

And given that it is established beyond reasonable doubt that we, with all our advanced cognitive and linguistic abilities, have evolved from ape-like creatures lacking those features, it is not a matter of whether the features that we use for language are the product of evolution — they must be. The question is whether they were shaped by natural selection for linguistic purposes, or not. Botha, in a series of papers (2001a; 2002b; 2002a), attempts to show that there is insufficient evidence to establish either of these possibilities, mainly due to what he regards as various definitional and epistemological shortcomings in the literature that he reviews. But Botha, apart from spending too much effort on unhelpful word games, appears to have missed the point that either one or the other (or some combination) must be true, unless one wishes to postulate some model of language origins totally at odds with evolutionary biology.

Here, as elsewhere in this chapter on language evolution hypotheses, it must be kept in mind that these issues are not black-and-white dichotomies. Some language-related features may be adaptations, and others may be spandrels. And even a single feature may have a mixed origin, starting out as a spandrel and then being fine-tuned — adapted — for language.
But do adaptations or spandrels predominate among the features that we use for language? To begin with, there is a chicken-and-egg problem at the very beginning of language evolution — with no language at all there will be no selection pressure towards adapting our bodies and minds for language use, and without such selection pressure we won’t be adapted for language use — implying that the first steps towards language had to be based on pre-existing features that had originally evolved for some other purpose. The co-opting of exaptations is thus a necessary first step in language evolution, or for that matter in the origin of any evolutionary novelty.

But what about language in its modern form? Pinker & Bloom (1990) argue strongly in favor of language as an adaptation, based on both its complexity and its obvious usefulness:

Evolutionary theory offers clear criteria for when a trait should be attributed to natural selection: complex design for some function, and the absence of alternative processes capable of explaining such complexity. Human language meets this criterion:... (p. 707).

The argument is further elaborated by Pinker (1994; 1998a) and others, to the point where Győri (2001c) can write about “the general recognition of language as a complex adaptive trait...” (p. 124).

Carstairs-McCarthy (1999) and Bierwisch (2001) on the other hand apparently consider language to be a pure exaptation, and Gould (1997) and Bickerton (1995) seriously consider the possibility of language being a spandrel. Chomsky (1988) can also be interpreted this way — he certainly argues that our mathematical ability is a spandrel (pp. 168f), but he is less explicit about language; the closest he comes is the quote on page 161 above.

Ragir (2001) also regards language as a spandrel, but from a different perspective; both language and other species-specific behaviors are in her view byproducts of our encephalization. None of them, nor Botha (2002b), however, offers any strong counters to the complexity argument of Pinker & Bloom (1990) above.

According to Christer Johansson (2001), language emerges through cultural evolution on a pre-existing biological substrate. In this case, there would be no adaptive biological evolution involved, only exaptations — language would be culturally adapted to us, we would not be biologically adapted to language, we would just be the selective environment for language (cf. section 3.5.1).

The issue of whether language is an adaptation also hinges on whether language is functional. Only functional features, that are actually useful, can provide the kind of selective advantage needed for a process of Darwinian adaptation. But in the Chomskian paradigm, the essence of language is structural, formal, and autonomous, disconnected from any considerations of meaning or function. Such an abstract autonomous language faculty can hardly convey any fitness advantage except by pure accident, and it is thus difficult to argue that it is an adaptation (Dor & Jablonka, 2001). But at the same time, the notion of an innate Universal Gram-
Hypotheses of language origins

Language as an adaptation comes much more naturally within a functionalist paradigm like Cognitive Grammar (Győri, 2001c). E.g. Langacker (1987) and Givón (1995) provide theories where language is explained as a functional consequence of general cognition. If language is functional, then Darwinian evolution for better functionality ought to be straightforward.

Lightfoot (2000) presents a rather peculiar argument against language being an adaptation. To begin with, he brings up an ultra-adaptationist strawman that he calls a ‘singularist’ (p 235), arguing that singularists believe that every single feature of every organism is adaptive in itself, and that nothing but natural selection ever affected any feature. Then he goes on to argue that a specific grammatical rule, applied in a particular subcase, appears to be dysfunctional for that subcase, therefore that rule for that subcase cannot be an adaptation and must be a spandrel. Thus the strawman is defeated — but there hardly exist any real adaptationists (as opposed to strawmen) who would deny that features exist that have side effects that are not necessarily adaptive; the main feature can still be an adaptation, shaped by natural selection, if its benefits outweigh the side effects. Furthermore, it is far from obvious that the feature invoked by Lightfoot is actually dysfunctional — both Bickerton (2003) and Deacon (2003b) propose functional explanations for it. Nevertheless, Lightfoot seems to believe that he has ruled out adaptation as an explanation for this grammatical rule — and then in a total non sequitur he goes on with the argument:

...of course, precisely the same could be true of UG as a whole: UG may have evolved as an accidental side effect of some other adaptive mutation. [...] Natural selection may have played no direct role in the evolution of UG specifically. (Lightfoot, 2000, p. 245).

Evolution is a complex process, with many subprocesses. Natural selection is one of them, but nobody is claiming it is the only one — the question is how important it is, how much of the present state of, in this case, our biological language endowments, have been shaped by natural selection for linguistic purposes. Arguing like Lightfoot (2000) does not move that debate forward.

Andrews et al. (2002) is a more serious discussion of how to disentangle natural selection from other evolutionary processes. It is not specifically about language, but more concerned with general principles of evolutionary inference. But its examples are largely picked from human cognition, so much of it may be adaptable to the case of language. In the article, Andrews et al. analyze a number of related criteria that may be used to distinguish adaptation from non-adaptation:

- Comparative evidence

---

6This particular strawman is not unique to Lightfoot (2000), it can be found also in e.g., Wuketits (in press) and in various other places in the anti-adaptationist literature.

7This is common enough in biology to have a technical term of its own, pleiotropy — see e.g., Futuyma (1998).
Origins of language

Fitness maximization
Beneficial effects
Optimal design
Tight fit between feature and function
Special design

None of these criteria is sufficient on its own — all are susceptible to both type I and type II errors, both failing to find adaptations when they are real, and finding adaptations where there aren’t any.

The ‘optimal design’ criterion is closely related to ‘Chomsky’s paradox’ (Li, 1997), the seeming contradiction between the apparently highly optimized mathematical elegance of Universal Grammar, and the generally non-optimized ‘bricolage’ (Duboule & Wilkins, 1999) character of evolved systems (Botha, 1999). But, as Li (1997) shows, this contradiction is only apparent, and not a serious argument against language being a product of evolution. Likewise, Newmeyer (1992) argues that autonomy of grammar does not exclude functional explanations, from which it follows that evolutionary ones are not excluded either. Jackendoff (cited by Botha (1999)) instead resolves the paradox by arguing that language isn’t perfect, that it does have the patchwork character typical of evolved systems. Marcus (2004c) takes this argument one step further, identifying patchwork candidates in our language capacity, ‘fossils’ of its evolutionary history. The presence of such fossils gets around ‘Chomsky’s paradox’, demonstrating that language does have an evolutionary history, but also demonstrating that some aspects of it are exaptations, not optimized for language.

The main conclusion of Andrews et al. (2002) is that it is far from easy to demonstrate conclusively either that any particular individual feature is an adaptation, or that it isn’t, but that the burden of proof must be balanced between adaptationists and ‘exaptionists’. In an attached commentary Haig & Durrant (2002) add the important point that we should be less concerned with proof for or against adaptation, and more concerned with inference to the best explanation.

On a genetic level, it is possible to distinguish genes that have been subject to recent natural selection from genes that have changed merely due to random unselected mutations — the statistical distribution of gene variants in the population is different. It is interesting to note that the only known ‘language gene’, FOXP2 (see page 104), shows a distribution in modern humans indicating strong natural selection (Enard et al., 2002; Pinker, 2003), which strengthens the case for language being an adaptation.

The case for language as an adaptation, at least in its full modern form, is compelling. Both the complexity criterion of Pinker & Bloom (1990) and the majority of the criteria from Andrews et al. (2002) listed above are amply fulfilled. This by no means excludes the possibility that language co-opted numerous other systems, either spandrels or exaptations — on the contrary, Elizabeth Bates is very likely
right in that “[l]anguage is a new machine that Nature built out of old parts.” (2003, p. 263). Nevertheless, the final assembly and refinement of the human language capacity into the exquisitely fine-tuned complex system we have today, must have been an adaptive process. It is indeed largely built from old parts, but the old parts are assembled and tuned adaptively.

This conclusion does not, however, tell us to what extent this adaptation is a matter of biological evolution, and how much of it is cultural or memetic evolution, to what extent we are adapted to use language, and to what extent language is adapted to be used by us. That issue, already discussed in Section 3.5.2, will be further addressed in Section 9.7 below.

9.4 Early vs. late

Did our language capacity evolve long ago, in the early stages of hominid evolution, or was language evolution a late development, taking place in anatomically modern Homo sapiens? ‘Early’ would mean at least several hundred thousands of years ago, and possibly one or two million years ago (Wildgen, 2004),\(^8\) or even longer (King, 2003), whereas ‘late’ would be within the past 100,000 years or so (Li & Hombert, 2002). As noted earlier, the time frame of language evolution is not strongly constrained by either fossils or anatomy alone. Our biological language adaptations cannot be younger than 60,000 years or so, and are very unlikely to be younger than 100,000 years (see page 74), but they can in principle be much older. Exactly how much older depends on the language capacity of apes — but even without ape language, human language could have evolved at any time after our common ancestor with chimpanzees, 5 million years ago or more (page 51). Neither ‘early’ nor ‘late’ hypotheses can be firmly excluded on paleontological grounds alone, though ‘late’ hypotheses with biologically based language faculties are severely constrained. Hypotheses in which language emerges through cultural evolution are less constrained.

The constraints get quite a bit firmer when the evolution of our speech organs and hearing is taken into account. As discussed in sections 5.1 and 5.2, there are signs of speech adaptations in Neanderthals, implying that the last common ancestor of us and the Neanderthals had some form of speech, pushing back the lower limit on the origin of speech to half a million years or so, effectively ruling out ‘late’ hypotheses. It should be noted, however, that this does not mean that full human syntactical language has to be that old — some simpler form of spoken proto-language may be enough to drive the evolution of speech adaptations.

\(^8\)But Wildgen (2004) is not entirely consistent on this point — he also talks of “protolanguage in the time of the Neolithic revolution...” (p. 184).
9.4.1 Art and technology as proxies for language?

The archeological record has frequently been invoked as support for the late, sudden appearance of language, due to the perception of a technological and creative revolution around 40,000 years ago (Li & Hombert, 2002).

Language use, of course, does not fossilize, at least not before the invention of writing, but other forms of symbol use may, and may be used as indicators that some level of symbolic abilities has been reached. The use of ancient art, including pigments and personal ornaments, as indicators that the artists were capable of symbolic thought, or even as an indicator that language had evolved (Mellars, 1998), is fairly common:

The pieces of ochre, ... were clearly intended for decorative or ritual use. This proves that the people who made them must have been capable of subtle thought, and probably indicates that they spoke a language of syntax and tenses, Professor Henshilwood said. (Henderson, 2002, p. 1, online version; see also Henshilwood et al. (2004)).

The connection between the decorative use of ochre, and grammatical details is, however, not overwhelmingly supported.

The supposedly sudden appearance of advanced art and advanced tools in the caves of Europe about 40,000 years ago is taken as evidence of a cognitive leap. However, the appearance of a sudden dramatic ‘cultural revolution’ around 40,000 years ago, has turned out to be largely an illusion caused by the predominance of European sites in the documented archeological record, and possibly some Eurocentrism among archeologists (Henshilwood & Marean, 2003). *Homo sapiens* did indeed invade Europe rather suddenly about 40,000 years ago, bringing along an advanced toolkit — but that toolkit had been developed gradually in Africa9 over the course of more than 200,000 years (McBrearty & Brooks, 2000; Van Peer et al., 2003). Kuhn et al. (2001) remain skeptical of the interpretation of McBrearty & Brooks (2000), but later discoveries of less ambiguous works of abstract art (Henderson, 2002; Henshilwood et al., 2002; Balter, 2002a; Recer, 2002; Harms & Yellen, 2002), pigment use (Barham, 2002), and personal ornaments (Henshilwood et al., 2004; Holden, 2004a) add further support to the long timescale of McBrearty & Brooks (2000). The debate over the supposed revolution is reviewed by Balter (2002c), Bar-Yosef (2002), and Henshilwood & Marean (2003).

Art is reasonably regarded as indicative of abstract and symbolic thought, and it is commonly argued, though not self-evident, that “[a]bstract and symbolic behaviors imply language, ...” (McBrearty & Brooks, 2000, p. 486), but McBrearty & Brooks (2000) certainly have a point also in the less commonly realized continuation of the sentence “..., but it is doubtful that the point at which they first can be detected coincides with the birth of language.” (p. 486). If we can observe signs of

---

9 According to d’Errico (2003), there are precursors also at Neanderthal sites in Europe, and Hovers et al. (2003) present ochre finds from modern humans in the Middle East from around 100,000 years ago.
art or other symbolic behavior, we might infer, following the logic of McBrearty & Brooks (2000) and others, that the artists had language, but the converse does not apply — the absence of fossilizable art does not imply absence of language.\footnote{For that matter, the absence of fossilizable art does not even imply the absence of art — most art among modern humans isn’t fossilizable, and it is not difficult to imagine a long period with only perishable art (body painting, wood carving, etc.) before anybody got around to making stone statuets or painting in deep caves.}

This inference from art to language, or at least from art to symbolic capacities adequate for language, is interesting in view of the additional evidence that has been uncovered recently that appears to show that simple art actually predated the appearance of anatomically modern *Homo sapiens* (Keys, 2000; Bahn & Vertut, 1997; Bednarik, 2003), in the context of *Homo heidelbergensis* or possibly even *Homo erectus*. Objects that can reasonably be interpreted as art have been found associated also with Neanderthals (Appenzeller, 1998; Wynn & Coolidge, 2004; d’Errico et al., 2003), though much simpler than the figurative art of later *Homo sapiens* (Conard, 2003), which would push back the origin of the biological capacities needed for art at least to the common ancestor of Neanderthals and us, some 500,000 years ago. And given that the symbolic capacities needed for art are also needed for language, and are interpreted by some as indicative of the presence of language, this adds support to the possibility of an early appearance of language, in agreement with the limits inferred from anatomy on page 167. As for art itself, as a cultural phenomenon, either independent invention in both lineages, or horizontal memetic transfer between us and Neanderthals, are conceivable, but a common origin of art may still be the simpler hypothesis. So far, we have insufficient data for any firm conclusions on that point.

One serious problem with the inference from art to human language, is that traces of both proto-symbolic thought and artistic activities have been observed in apes. Both chimpanzees and gorillas happily produce paintings when supplied with canvas, brushes, and paint. And at least one language-trained ape has even been reported to describe what her works of art represent (Patterson, 1981), which would seem to indicate that the apes themselves regard their art as representational. One could, rather optimistically, argue that chimps have the capacity for both art and language, saving the inference, but making it useless for elucidating the history of human language. The alternative appears to be to exercise caution in drawing conclusions from art to language.

A related argument is that of Barnes (1997), who postulates language as a requirement for religion, for much the same reasons as for art — religion requires the ability to reason symbolically about abstract categories. Müller (1866) proposed instead a more direct role for religion in the origin of language, with religious awe as the root of the need for speech (Gans, 1999c).

Archaeological data on the origins of religion are unfortunately sparse and controversial — much Paleolithic art, from statuetts to cave paintings, has been interpreted in religious terms, but other interpretations cannot be excluded (Bahn &
Vertut, 1997). The same is true for prehistoric ceremonial burials (Gargett, 1999). This uncertainty means that the religious argument adds no constraints to the possible origins of language.

9.5 Gradual vs. sudden

Did we acquire our language capacity in one single step, without intermediate forms, or did we go through a long sequence of successive proto-language stages?

To begin with, it should be noted that there is perennial confusion over the word ‘sudden’ as used in deep historical and geological contexts. A process that took, say, 10,000 years would appear very gradual to the participants — but would appear instantaneous in the fossil record to paleontologists working a million years later, and would be labeled as a ‘sudden’ event by them. Many evolutionary transitions belong in this category of events that are paleontologically sudden but on human timescales gradual, and this is the root of the debate surrounding ‘punctuated equilibrium’ (Eldredge & Gould, 1972) — the hypothesis of punctuated equilibrium proposes that evolutionary transitions are geologically sudden, not necessarily sudden on human timescales:

... the punctuations of punctuated equilibrium do not represent de Vriesian saltations, but rather denote the proper scaling of ordinary speciation into geological time. (Gould, 2002, p. 768).

However, most proponents of gradual evolution of language intend the process to be geologically slow, and most proponents of sudden evolution are saltationists, talking about a single jump from ape-like to human-like language abilities, so this problem is not severe when it comes to differentiating between hypotheses in this context. What is a problem, however, is that Gould’s point in the quote above is commonly forgotten, and Gould is often cited in support of saltationism (cf. footnote 2 on page 13).

Another problem with this dichotomy between gradual and sudden language evolution, is that both sides are primarily discussing the biological evolution of the human language capacity. But biology is only one aspect of language evolution and, as discussed in Section 3.6, the aspect slowest to evolve. Cultural and memetic evolution is relevant as well, and can be orders of magnitude faster.

But regardless of whether we are discussing biology or memetics, the sudden single-step evolution of something as complex as the human language capacity is highly problematic. If we have an innate dedicated ‘language organ’ and a universal grammar that is genetically specified at the level of detail assumed in e.g. Lightfoot (2000), with genes for individual grammatical rules, this requires a large number of highly specific genes working together in a coordinated pattern. And the simultaneous de novo evolution of many coordinated genes is so utterly unlikely that ‘sudden’ hypotheses in that case become totally untenable without
divine intervention, *contra* Chomsky (1988) and Bickerton (1990).\footnote{It should be noted that Bickerton himself, to his credit, has now acknowledged that his earlier position was biologically ridiculous — see footnote 2 on p. 80 in Bickerton (2003).} The only context in which ‘sudden’ single-step hypotheses are not totally ridiculous is if most of the bodily and cognitive features that we use for language evolved for some other purpose, and were available as exaptations, with only some minor additional change needed to put all the pieces together as a workable language organ, and even then some intermediate stages of proto-language would appear necessary to render the hypothesis evolutionarily plausible. Carroll (2003) definitely has a point in that “the temptation to invoke macromutational models for ‘rapid change’ [...] must be resisted in the absence of genetic evidence.” (p 852).

The existence of master regulatory genes is sometimes invoked as an explanation for sudden evolutionary saltations (Schwartz, 1999), a dubious notion occasionally seen also in the context of language origins:

> The mechanism underlying the sudden origin of phenotypic characteristics whether anatomical, physiological or behavioral is the duplication of the master regulatory genes, the so-called Homeotic genes. (Li & Hombert, 2002, p. 185)

As an explanation for the sudden origin of our language capacity,\footnote{It should be noted that Li & Hombert (2002) state that by ‘sudden origin of phenotypic characteristics’ they do not mean sudden origins of language.} or even our large brains, this is nonsense. The homeotic genes exist, and do act as ‘master switches’, turning on or off developmental programs — but only if those programs already exist. Mutations in the homeotic genes can cause body parts to move around or duplicate or disappear, or cause new copies of old parts to sprout in odd places, but genuinely new features require changes in the developmental programs themselves, not just in the master switches. There is in any case no evidence that any homeotic genes have been duplicated in the human lineage, as all mammals appear to have the same set, in the same number of copies. The canonical homeotic genes of the Hox family aren’t expressed at all in the relevant parts of the vertebrate brain (Rancourt, 1998), though other homeotic genes of the Emx and Otx families are (Deacon, 1997). Mutations — not duplications\footnote{Duplications as such are common enough in the human genome (Abdellah et al., 2004), but no recent homeotic duplications have been identified. In any case, the effects of duplications are more a matter of long-term evolvability than of sudden saltations.} — in regulatory genes\footnote{More likely in the DNA sequences controlling the timing and pattern of homeotic gene expression, rather than in the homeotic genes themselves.} are perfectly plausible as explanations for the massive brain growth in *Homo sapiens* — but not in sudden jumps\footnote{Which in any case is not what the fossil record indicates; cf. Section 5.3.4.} and definitely not for filling our new brains with specific capacities.

When discussing language evolution, the prerequisites for evolutionary processes (listed in Section 3.1) must be kept in mind. An important point here is that heritable variation in language abilities is necessary, otherwise there is nothing for
natural selection to select. To the extent that language is innate, this heritable variation must be genetic. For gradual language evolution to be tenable, the variations ought to be of rather modest magnitude, whereas hypotheses of sudden language origins ought to predict all-or-nothing variation, either full language capacity or nothing at all.

As was shown in Section 5.3.3 in the context of SLI, the evidence supporting the existence of genes that affect language is quite compelling, at least in the case of FOXP2. But FOXP2 defects (and SLI in general) only cause partial loss of language, not the total loss that would be expected if language were the result of a single macromutation. Furthermore, Stromswold (2001) finds strong evidence of a heritable component in the existing variation in language abilities, even between people with no evident language abnormalities. The existence of such small-scale genetic variability is consistent with expectations from gradual, but not sudden, hypotheses.

One might invoke also the non-negligible heritability of verbal IQ, but it is unclear both to what extent verbal IQ is independent of other cognitive abilities (Alarcón et al., 1999), and to what extent verbal IQ actually measures language abilities in the sense relevant here.

Pinker & Bloom (1990) add some more data and anecdotal evidence supporting variability in our syntactic abilities, but they also point out that, while feeding on variation, natural selection also eliminates variation — if only the most able individuals breed, and their offspring inherit their abilities, the spread in ability will decrease with each subsequent generation, unless new variation is added in the form of mutations. Early hominids may well have varied in linguistic abilities, even if little such variability had remained today.

Also to be considered in this context is the argument, usually based in the Chomskian paradigm, that our language capacity is a monolithic universal grammar module (Chomsky, 1982), a unified whole in which variation is logically impossible. But there are several ways around this argument:

- Even if grammar, as an abstract entity, may be monolithic, its implementation in our brain may be more or less efficient — even if all people use the same universal grammar, it is possible that some can acquire and process language faster and easier than others. That shows us a conceivable evolutionary path from an initial state where the same grammar was handled in a slow and muddled way by whatever cognitive and heuristic abilities were available, through more and more efficient neural circuits, towards the modern human brain with which we effortlessly acquire language at an early age.

- It is not self-evident that grammar actually is monolithic, with no imaginable partial proto-grammar. We’ll return to this point in Section 11.4 below.

- The existence of SLI and aphasia patients with partial language deficits demonstrates that blocks can be taken out of the ‘monolith’ without the total collapse of language.
The gradual evolution of tightly coupled apparently monolithic systems was discussed on page 18, and there is no reason to believe that the conclusion there isn’t applicable to language. The fact that for Chomsky “...it is not easy even to imagine a course of selection [towards language] ...” (1988, p. 167) is not a strong counterargument. Pinker (2000) has a better case when he states that “the game theorists have demonstrated the evolvability of the most striking features of language...” (p. 442, emphasis added).

In conclusion, the gradual evolvability of our apparently monolithic grammar is far from excluded (Pinker, 1994; Jackendoff, 1999b). And given the near-impossible odds against the single-step appearance of something as complex as language, we can conclude that the evolution of language is overwhelmingly more likely to have been gradual, in the sense of entailing many small evolutionary steps, rather than a single leap. If biological evolution dominated the process, as it would have to if language is innate in any strong sense, then the process can be expected to be geologically slow. On the other hand, if language is largely the product of memetic evolution, then even a gradual process may appear geologically sudden.

9.6 Speech first vs. gestures first

Did language first evolve in the spoken modality dominant today, or was another modality, presumably gestures, used in the early stages? Darwin (1872) felt quite certain about the origin of language:

I cannot doubt that language owes its origin to the imitation and modification, aided by signs and gestures, of various natural sounds, the voices of other animals, and man’s own distinctive cries. (Darwin, 1872, p. 56)

Unfortunately, this is one of the rare cases where Darwin’s intuition led him partially astray — there is good reason to doubt the homology of animal calls and human speech. This means that it is not self-evident that language started with sounds, precursors to the speech modality. The ‘signs and gestures’ that Darwin invokes as aids may conceivably have been the main modality of early language instead.

Language per se is basically modality-independent, as long as the modality used supports a sufficiently rich structure. In modern society, a large fraction of all language use is written rather than spoken. If anything, the written modality supports more complex language than the spoken. Other alternative modalities can easily be imagined, and quite a few have been used, both in ape language experiments (see Chapter 7) and in the teaching of severely retarded non-speaking children (Savage-Rumbaugh & Lewin, 1994).

16But see Cowley (2002).
17An interesting case is when the two chimps Sherman and Austin (p. 131) apparently invented a new modality on their own, spontaneously, when deprived of their usual computerized system (Savage-Rumbaugh & Lewin, 1994).
Written language is of course derived from spoken in evolutionarily recent times, and so it is not highly relevant to the origin of our language capacity. But another alternative modality, sign language, is more interesting in this context. Sign language, just like spoken or written language, is a bona fide language (Sandler, 1993; van der Hulst & Mills, 1996), with all the functionality of any other modality.


Sign language displays the same features as spoken language, not only in its mature form, but also in its development and in its neurological organization. Children of deaf signing parents ‘babble’ in sign language during their early development (Petitto & Marentette, 1991; Berent, 1996; Petitto et al., 2001b), start signing at the same age and with the same basic vocabulary as the first words of children of hearing and speaking parents (Cheek et al., 2001), and their further development goes through basically the same stages as hearing children (Locke, 1997). In the case of bimodal bilingual children, simultaneously acquiring both a signed and a spoken language, the parallels are very clear, with the same child attaining various linguistic milestones simultaneously in sign and speech (Petitto et al., 2001a). The formation of pidgins and creoles have been observed among deaf people (Goldin-Meadow & Mylander, 1998; Goldin-Meadow, 1999; Helmuth, 2001a), with fundamental language properties like combinatorial discreteness emerging spontaneously (Senghas et al., 2004). Brain lesion studies, as reviewed by Hickok et al (1996; 1998a; 2001), show a pattern of sign language aphasia among the congenitally deaf that resembles speech aphasia among hearing patients in the correlations between deficit patterns and affected brain areas. Likewise neuroimaging experiments (Neville et al., 1998; Hickok et al., 1998b) see similarities between speaking and signing. There are also minor differences between speech and signing in the brain, but it is unclear how much of this is simply attributable to the different sensory and motor areas involved.

\[\text{Written language is of course derived from spoken in evolutionarily recent times, and so it is not highly relevant to the origin of our language capacity.}\]

\[\text{But another alternative modality, sign language, is more interesting in this context. Sign language, just like spoken or written language, is a bona fide language (Sandler, 1993; van der Hulst & Mills, 1996), with all the functionality of any other modality.}\]


\[\text{Sign language displays the same features as spoken language, not only in its mature form, but also in its development and in its neurological organization. Children of deaf signing parents ‘babble’ in sign language during their early development (Petitto & Marentette, 1991; Berent, 1996; Petitto et al., 2001b), start signing at the same age and with the same basic vocabulary as the first words of children of hearing and speaking parents (Cheek et al., 2001), and their further development goes through basically the same stages as hearing children (Locke, 1997). In the case of bimodal bilingual children, simultaneously acquiring both a signed and a spoken language, the parallels are very clear, with the same child attaining various linguistic milestones simultaneously in sign and speech (Petitto et al., 2001a). The formation of pidgins and creoles have been observed among deaf people (Goldin-Meadow & Mylander, 1998; Goldin-Meadow, 1999; Helmuth, 2001a), with fundamental language properties like combinatorial discreteness emerging spontaneously (Senghas et al., 2004). Brain lesion studies, as reviewed by Hickok et al (1996; 1998a; 2001), show a pattern of sign language aphasia among the congenitally deaf that resembles speech aphasia among hearing patients in the correlations between deficit patterns and affected brain areas. Likewise neuroimaging experiments (Neville et al., 1998; Hickok et al., 1998b) see similarities between speaking and signing. There are also minor differences between speech and signing in the brain, but it is unclear how much of this is simply attributable to the different sensory and motor areas involved.}\]
Even among people using spoken language, gesturing is firmly wedded to language use (Goldin-Meadow, 1999; Bates & Dick, 2002) — your hands are likely to be moving even when you are talking on the telephone, and even congenitally blind people (who can hardly have acquired the habit by observing others) gesture while speaking, also when addressing a blind listener (Iverson & Goldin-Meadow, 1998). Normal hearing children acquire the use of communicative gestures in parallel with speech acquisition (Acredolo & Goodwyn, 1988; Bates & Dick, 2002), and there is some evidence that gesturing actually precedes speech in acquisition (Goodwyn & Acredolo, 1993; Goldin-Meadow, 1999). At the very least, gesture is as important as speech in early child communication, before the advent of rudimentary syntax in the two-word stage (Iverson et al., 1999).

And the manual dexterity required for gesturing is present in many primates, including our closest relatives (see Chapter 7), so it is reasonable to assume that it has been present for a long time among our ancestors, tens of millions of years at least. Apes also have the cortical control of their hands needed for sign language (Corballis, 1999), while lacking the corresponding vocal control, as discussed on page 81 (though a complication is that they also appear to lack voluntary control of facial expressions, ubiquitously used in human gestural communication (Premack, 2004)). Accordingly, wild chimpanzees can communicate voluntarily and flexibly with gestures, whereas their vocalizations are mainly involuntary (Tomasello, 2003). Interesting in this context is that in chimpanzees fine motor movements of the hands are frequently accompanied by sympathetic mouth movements, hinting at a possible path from gestures to speech (Waters & Fouts, 2002).

The transition to bipedalism may nevertheless have been an important exaptation in this context (Corballis, 2002), as it in the short run freed the hands for gesturing even while moving around, and in the long run freed the hands from selection pressure for locomotor efficiency. As already mentioned on page 82, our bipedal posture may be important for speech as well, as it decouples breathing from stride.

9.6.1 Mimesis

Mimesis (or mimetics21 — not to be confused with memetics; see page 25) concerns the art of miming or, as Donald (1997) puts it, “*us*[ing] the whole body as a represen*ta*tional device” (p. 4, online edition) or “...as a communication device...” (p. 6), which both Donald (1997) and MacNeilage (1994) regard as a vital first step in the evolution of language. It is not an unreasonable suggestion that miming, imitating, and pretending can be regarded as proto-symbolic activities that may be related to the origin of language, particularly if language started in a gesturing modality, for which miming abilities are plausible exaptations — modern sign lan-

---

21 Clark (2004b) makes a distinction between mimesis — telling-by-showing in general — and mimetics — mimesis with sounds only, onomatopoeia.
guages still have considerable mimetic components (Newport, 1982). Mimesis has many properties that may be useful bridges between animal communication and human language (Zlatev, 2003b):

- Partial generativity
- Intentionality
- Public representation
- Parity
- Iconicity

A possible mimetic origin for syntax was discussed already by Condillac (1746) and Reid (1765); cf. page 159.

Zlatev (2001a) identifies a mimetic stage in human ontogeny, at which pre-verbal children acquire awareness of self and others, and take the first steps on the road towards social communicative competence, using miming and gestures for communication. This stage, and the self-consciousness and social interactions that it entails, is a key stage in the acquisition of true meaningful language, according to Zlatev (2001a). The role of mimesis in language ontogeny is further discussed in Vihman & Depaolis (2000).

There is little clear evidence of mimesis in apes, but it is not unknown in dolphins (Bauer & Harley, 2001).

9.6.2 Mirror neurons

The hypothesis of Rizzolatti & Arbib (1998), that the roots of language can be traced to the so-called ‘mirror neurons’, reviewed in Stamenov & Gallese (2002), has some parallels with mimesis, but has a neural rather than a behavioral basis. ‘Mirror neurons’ make up a neural system that is activated both by performing a certain action, and by observing — either seeing or hearing (Théoret & Pascual-Leone, 2002; Buccino et al., 2003) — the same action performed by others. This is very likely part of the neural basis of imitative learning, with the mirror neurons performing a high-level synthesis role in the network of neural connections reviewed by Schaal (1999). Rizzolatti & Arbib (1998) hypothesize that the mirror neurons also led to a system of gestural communication, with iconic gestures mimicking the action that’s the topic of communication. The gestural system would have included both manual and oro-facial gestures, with speech growing out of the oro-facial gesturing system. It is interesting to note that the mirror system in monkeys is located in their equivalent of Broca’s area (Schaal, 1999; Théoret & Pascual-Leone, 2002).

The mirror neuron system was first identified in monkeys (Gallese et al., 1996; Fogassi & Gallese, 2002). It was subsequently discovered that humans share the mirror system with monkeys and apes (Rizzolatti et al., 2002), apparently located in our Broca’s area (Iacoboni et al., 1999; Buccino et al., 2003).
Despite the common neural mechanisms in the mirror system, there are qualitative differences between the imitative learning of humans and other primates (Call & Tomasello, 1995; Tomasello et al., 1993; Nagell et al., 1993, but see also Voelkl & Huber (2000) and Whiten et al. (2004)) — but interestingly enough, enculturated apes who have grown up with humans show human-like imitative learning (Tomasello et al., 1993; Bjorklund et al., 2002, but see also Bering (2004)).

The mirror neurons may be a useful exaptation in the emergence of a gestural language, but cannot by themselves explain its emergence — after all, lots of monkeys have mirror neurons but no trace of language (Deacon, 2004a), nor any theory of mind or human-style imitative or communicative capacity (Stamenov, 2002).

9.6.3 Why switch to speech?

But if language did first evolve in a gesturing modality, why did we switch to speech? This question can only be answered speculatively, but there are obvious advantages of speech over gestures:

- Speech is more efficient, using less time and energy (Knight, 2000).
- There is no need to see each other, an advantage in the dark, or in heavy vegetation (Rousseau, 1755).
- Speech calls attention to the speaker in a way that gestures do not (Rousseau, 1755).
- The hands are not needed for communication, making it possible to work or carry things while communicating (Carstairs-McCarthy, 1996).

Sign language has corresponding advantages in very noisy environments, or when stealth is an issue, situations in which people even today communicate with gestures. One can well imagine a gradual transition from gesturing to speech, with intermediate stages similar to those depicted in Auel (1980), in which sign language is augmented by a few sounds.

Bradbury & Vehrencamp (2000) review the economic viability of communication systems, setting a lower limit for the accuracy of signal coding, below which it is not worthwhile for receivers of signals to pay any attention to their content. In this model, it makes sense for communication systems to start out by using as signals such behavior that potential receivers have already evolved to pay attention to for other reasons. Much animal communication can readily be interpreted within such a framework. It is unclear, but would be relevant to investigate, whether hominoid vocalizations or gestures are better from this perspective.

An alternative possibility is that gestures and speech were used in parallel in the beginning, while the production and reception of both modalities were still in their infancy (Bickerton, 2003). According to Rowe (1999), such multicomponent signaling improves detectability and discriminability beyond that possible with either component alone.
If gestures came first, then this implies early language, since anatomical speech adaptations turn up in fossils well before the postulated time frame for late language.

Alternatively, if speech came first, then we have two possibilities:

- Early speech, gradually evolving in articulation, starting with the sounds that apes can produce, with selection pressure from speech driving the anatomical reconstruction of the speech organs. This kind of coevolution of speech organs and language is evolutionarily plausible.

- Late speech — but this is problematic for the same reasons as late signing; language must be in place before obvious anatomical language adaptations.

In either case, language evolution must be well underway before the anatomical speech adaptations can be selected for. And since some of these adaptations go all the way back to the last common ancestor of *Homo sapiens* and Neanderthals, more than half a million years ago (see page 80), this effectively rules out late language.

### 9.7 Innate and genetically determined vs. learned and culturally determined

Reviewing the full debate on whether language is innate in humans, and if so, what this means, is beyond the scope of this book. On one level, innateness is hardly controversial in the limited sense that Chomsky alluded to when he said:

> I have no idea what the phrase [innateness hypothesis] is supposed to mean and correspondingly have never advocated any such hypothesis — beyond the truism that there is some language-relevant distinction, ..., between my granddaughter and her pet kitten [...]. (Chomsky, quoted in Stemmer (1999))

It is self-evident that humans have innate, genetically determined language-relevant abilities that kittens don’t. It is also uncontentious that language is learned, in the limited sense that the particulars of individual languages aren’t innate. What is controversial, however, is to what extent the innate abilities that we unquestionably do have are specifically linguistic, and to what extent they constitute a genetically hardwired ‘universal grammar’ incorporating what Pullum & Scholz (2002) call “specific contingent facts about natural languages” (p. 10). Chomsky, e.g., (1965), as well as other linguists working within the Chomskian paradigm, e.g. Pinker (1995), commonly make much stronger claims about innateness than Chomsky’s kitten quote above. But the debate is often unnecessarily polarized — it is not a matter of total genetic determinism on one side, and total *tabula rasa* conditioning on the other (Seidenberg & MacDonald, 1999; Jackendoff, 1999a).

---

22 Innateness is a somewhat problematic concept, lacking a clear and coherent definition — see e.g., Scholz (2002) and Griffiths (2002) for brief reviews of the complexities involved.
Hypotheses of language origins

despite the rhetoric of both sides. Innateness is more complex than that (Khalidi, 2002).

The claims of ‘strong innateness’ rest on three main pillars:

1. The universality of certain language features (Chomsky, 1988). Particularly compelling is the emergence of the same universal features in the independent origin of creoles (Bickerton, 1995, but see also Owens (1990) and Mufwene (2002)) and sign languages (Siegal, 2004).

2. The poverty of the stimulus (Chomsky, 1980; Laurence & Margolis, 2001; Thomas, 2002) — the impossibility of language acquisition without having the central concepts of language available a priori, in an innate language acquisition device, LAD (Chomsky, 1965; Wanner & Gleitman, 1982; Fodor & Crowther, 2002; Lasnik & Uriagereka, 2002; Legate & Yang, 2002; Lidz et al., 2003). This is not just a practical matter of learners receiving insufficient and too noisy input — it is argued that, as the space of all possible grammars is infinite, it is impossible in principle to identify the target grammar of acquisition without innate constraints on the search space. There is no learning algorithm that can learn an arbitrary language from finite input, without constraints (Gold, 1967; Chomsky, 1975).

3. Patterns of language acquisition. Not just the fact that children can acquire language at all, but also the patterns seen in their early efforts, are invoked as evidence of an innate grammar. Particularly interesting here are the errors that children never make, errors that would have been easy to make if learning were inductive, but that violate principles of Universal Grammar (Crain & Thornton, 1998; Jackendoff, 2002). The errors that the kids do make are instead such that they conform to possible human languages other than the target one (Crain & Pietroski, 2002).

Arguments against innateness take several different forms. Laurence & Margolis (2001) review and dismiss a variety of philosophical objections; the focus here will be more on empirical issues, also from outside linguistics proper:

1. Language universals may have other causes than an innate grammar. And how universal are they really? Linguists who search for universals in language will

23DeGraff (2003), a linguist who happens to be a native Creole speaker, objects quite strongly to how Bickerton and others portray Creole languages, arguing basically that Creoles are no different from any other languages, and should not be treated as primitive ‘linguistic fossils’.

24Lidz et al. (2003) present experimental evidence that they interpret as favoring an innate LAD — but remarkably enough, they also say that up until their work, almost 40 years after the innateness hypothesis was proposed, it hadn’t been experimentally tested (cf. Pullum & Scholz (2002)): Generative linguistic theory stands on the hypothesis that grammar cannot be acquired solely on the basis of an analysis of the input, but depends, in addition, on innate structure within the learner to guide the process of acquisition. This hypothesis derives from a logical argument, however, and its consequences have never been examined experimentally with infant learners. (Lidz et al., 2003, p. B65, emphasis added)
generally find what they are looking for. But what conclusions can be drawn from this? Are the type of universals observed across languages actually evidence of UG (Haspelmath, 2004)? Here is a list of some conceivable ways of explaining language universals without innate grammar:

- When complex sets of data are studied and modeled, spurious structures and correlations are often found even when in reality there are none whatever. Is this problem excluded in the search for Universal Grammar? Tomasello (2003) and Sampson (1999a) question the reality of putative grammatical universals, apart from those deriving from general cognitive considerations, arguing that they are based on a too-narrow sample of languages, or on forcing ‘odd’ languages into a prescribed form, making the argument for universals circular. Some are simply empirically false even in English (Sampson, 2002).

  I have surveyed this literature [on language universals] as exhaustively as I could manage (...), and I concluded that no candidate for the status of contingent linguistic universal survives scrutiny. (Sampson, 2002, p. 100)

- Many similarities between languages may be adequately explained by their having a common origin. It appears quite likely that all human languages have a common origin, if one goes far enough back in time — otherwise one would have to assume that language was independently developed by several distinct groups of proto-humans. This is certainly possible, but the evolution of a single innate universal grammar, common for all mankind, actually requires that all languages have a common origin, spoken by the first people to evolve UG, in order to be compatible with standard Darwinism. And if a common origin has to be postulated anyway, why not let this common origin explain the universal features — to introduce innateness at this point would seem to go against Occam’s razor.

- All extant languages have been acquired by human children. Biases in the acquisition system — which need not be a matter of innate grammatical principles — can shape what form of language is acquired. The observed universals may reflect more general acquisition biases, rather than specifically an innate grammar (Kirby & Christiansen, 2003; Kirby et al., 2004). Languages, as memetic species, will adapt to be acquirable by whatever cognitive equipment children have — are universals the result of natural selection among languages?

- In order to be a useful instrument for communication, a language has to meet certain basic criteria. Is it possible that some principles of Universal Grammar can be explained by their being, logically or pragmatically, necessary features of a language? In this case, language would be the result of neither nature nor nurture, neither genes nor learning, in any simple sense. Deacon (2003b) develops the idea of logical necessity further, invoking semiotic constraints — symbols have to connect with their referents in a way that can be
Hypotheses of language origins

parsed — to explain the universal features of grammar. Language emerging from the pragmatic constraints involved in mapping between meanings and speech is reviewed in Bates (2003), with examples of alternative explanations for putative innate universals.

2. It is not empirically well established that stimulus is actually poor, in the sense of the ‘poverty of the stimulus’ argument (Pullum & Scholz, 2002); cf. the quote from Lidz et al. (2003) in footnote 24 above. The debate introduced by Ritter (2002) is a good review of this topic, with both pro (Fodor & Crowther, 2002; Lasnik & Uriagereka, 2002; Legate & Yang, 2002; Crain & Pietroski, 2002) and con (Pullum & Scholz, 2002; Sampson, 2002; Scholz & Pullum, 2002) arguments represented.

The lack of negative evidence in the learner’s input is frequently cited as evidence against learnability (Marcus, 1993; Marcus, 1999b; Pinker, 1995; Fodor & Crowther, 2002), but Saxton (1997) and Strapp (1999) provide examples of negative input that children may use. Sampson (2002) argues instead that the shortage of negative evidence is not unique to language acquisition — it applies equally well to e.g., scientific discovery, where nobody would argue that the results are innate. Furthermore, comprehension comes before production in language acquisition (Bates, 1993; Burling, 2000; Newmeyer, 2003a) — and there is no shortage of negative feedback for miscomprehension (Savage-Rumbaugh, 1990; Savage-Rumbaugh et al., 1993). Regier & Gahl (2004) also discuss under what circumstances the absence of evidence for a construction can be taken as evidence of its absence in the target grammar.

One may also consider that children from the same stimulus manage to acquire not only general syntax but also all the idiosyncratic peculiarities of the target language, including but not limited to a huge lexicon — but nobody argues that all these particulars are innate. Children clearly need powerful learning mechanisms — but do they need an innate grammar any more than they need an innate lexicon?

And the speech that language learners hear is rather different from normal adult discourse. Surprisingly young children can exploit linguistic and non-linguistic cues as an aid in speech perception and language acquisition (Shady & Gerken, 1999). As is well known, those adults who interact with language acquirers enrich their speech in such cues, sometimes to the point of ungrammaticality (Chafetz et al., 1992), creating what is known as ‘motherese’ (Elliot, 1981; Pinker, 1995), ‘parentese’ (Chafetz et al., 1992) or ‘child directed speech (CDS)’ (Cameron-Faulkner et al., 2003; Rivero, 2004) when directed towards

---

25Paradoxically, this would make Universal Grammar more universal than the innatist ‘Universal Grammar’ — the semiotic constraints apply not only to human language users, but to any symbolic system of communication. UG would be truly universal in the same way, and for the same reasons, as mathematics.

26A usage which I, being a father, consider sexist.
children, and ‘teacher talk’\textsuperscript{27} (Håkansson, 1987) when directed towards adult learners. Even phonetic contrasts are enhanced (Kuhl et al., 1997), making phonemes easier to distinguish, and the segmentation of speech into words is facilitated by many parents commonly using isolated words rather than full sentences (Wagner, 2001b; Cameron-Faulkner et al., 2003). Full canonical sentences are actually quite rare in child directed speech — but the sentence fragments and non-canonical sentences that are more common, may be at least as informative. Sentence fragments commonly consist of a grammatical unit, an isolated NP, VP or PP, which may help children identify them as coherent constituents of language (Cameron-Faulkner et al., 2003), facilitating the recognition that grammatical rules are structural, not linear. The gestures that accompany speech are likewise modified into a ‘gestural motherese’ that may function to reinforce or disambiguate speech (Iverson et al., 1999). Child directed speech is pragmatically adapted to the communicative competence of the child, with interactions kept very simple for the first nine months, and then increasing rapidly in complexity (Rivero, 2004). There are, however, examples of cultures where speech to children does not appear to be adapted like this, without obvious ill effects on language acquisition (Pinker, 1995).

3. The errors that children supposedly never make, the absence of which is taken as evidence of an innate grammar, do occur occasionally. Drozd (2004) presents several types of errors in child speech and child grammaticality judgements, that appear to violate UG constraints. Some are found even in adult speech — Sampson (1999a) quotes from a real-life conversation a canonical example of a UG-violating error: “Am what I doing is worthwhile?” (2002, p. 86).

4. The impossibility of language acquisition without an innate language acquisition device is not self-evident.\textsuperscript{28} To begin with, this argument is based on particular assumptions about what is actually acquired in language acquisition. “The notion of what constitutes important evidence for learning a particular structure is not theory-neutral.” argue Seidenberg & MacDonald (1999, p.

\textsuperscript{27}Why not ‘teacherese’?

\textsuperscript{28}In an interesting twist of logic, Bever (1982) reverses the logic of the impossibility argument. The standard syllogism of the impossibility argument can be stated as follows (Adapted from Bever (1982), p. 432):

\begin{enumerate}
  \item \textbf{Language has property }$P_1$
  \item $P_1$ cannot be learned by any known theory of learning
  \item Therefore $P_1$ is innate
\end{enumerate}

But what about this syllogism:

\begin{enumerate}
  \item \textbf{Language has property }$P_1$
  \item $P_1$ cannot be transmitted by any known genetic mechanism
  \item Therefore $P_1$ is learned
\end{enumerate}

It is not self-evident that one syllogism is more valid than the other. Bever (1982) proceeds from this point into an odd Platonic essentialist view of language; this is better regarded as a challenge to premise (b) of both syllogisms. Premise (b) of the first syllogism is related to the classical ‘Poverty of the stimulus’ argument, but contains also more general learnability arguments, whereas premise (b) of the second syllogism is similarly related to the ‘Poverty of the genes’ argument below. Whether either poverty argument is valid, is an empirical issue that remains to be settled.
Hypotheses of language origins

... many of the classic arguments rest on the assumption that the child’s task is grammar identification, and these arguments simply no longer apply if the task is instead acquiring the performance system underlying comprehension and production. (Seidenberg & MacDonald, 1999, p. 574).

Similarly, Clark (2001), using a statistical instead of a symbolic grammar in his computer model, “conclude[s] that the Argument from the Poverty of the Stimulus is unsupported by evidence.” (p. 1).

But even within the Chomskian paradigm, the issue is not entirely clearcut. The theoretical impossibility of grammar identification (Gold, 1967) applies only in a totally unrestricted search space. Even with very general restrictions, such as an upper limit on the number of rules in the grammar, the task is no longer impossible (Petersson et al., 2004). Grammar identification from positive evidence only is also possible with very loose restrictions on the class of grammars considered (Shinohara, 1994; Scholz & Pullum, 2002).

5. The timing of language acquisition, and particularly the relative timing of monolingual and bilingual acquisition, does not support the existence of an innate Universal Grammar (UG).

In a simplistic form, the argument here is that, if children do have UG innate, then no time is needed to acquire it. All the time a child uses for language acquisition is then spent on acquiring the particulars (lexicon, parameter settings, etc.) of whatever language(s) the child acquires. Acquiring two languages doubles the amount of particulars to learn, which ought to double the acquisition time. Thus, the innateness hypothesis predicts bilingual acquisition to be much slower than monolingual, contrary to observations.

More formally, the argument can be expressed as follows:

\[ t_{a1} = t_{UG} + t_\ell \]  
\[ t_{a2} = t_{UG} + t_{\ell i} + t_{\ell j} = t_{UG} + 2t_\ell \]

using the symbols defined below:

- \( t_a \): The acquisition time needed for a child to acquire its native language(s).
- \( t_{a1} \): \( t_a \) for a monolingual child.
- \( t_{a2} \): \( t_a \) for a bilingual child.
- \( t_{UG} \): The part of \( t_a \) spent in acquiring Universal Grammar, the core common to all languages.
\( t_\ell \) : The time it takes to acquire language-specific features (lexicon, language-specific rules and parameter-settings) of a single language. It is probably a fair approximation to assume that \( t_\ell \) is the same for all languages. \( t_{\ell i} \) and \( t_{\ell j} \) represent the acquisition times for the two languages of a bilingual.

If the innateness hypothesis is true, then \( t_{UG} = 0 \) (as UG is then innate, no acquisition of it is needed). Thus:

\[
\begin{align*}
t_{UG} &= 0 \\
t_{a1} &= 0 + t_\ell = t_\ell \\
t_{a2} &= 0 + t_{\ell i} + t_{\ell j} = 2t_\ell \\
t_{a2} &= 2t_{a1}
\end{align*}
\] (9.3) (9.4) (9.5) (9.6)

If the innateness hypothesis is false, then it can be assumed that \( t_{UG} \) is large compared to \( t_\ell \), a reasonable assumption considering the view of innatists that UG cannot possibly be acquired in the time available to a child. Thus:

\[
\begin{align*}
t_{UG} &\gg t_\ell \\
t_{a1} &= t_{UG} + t_\ell \approx t_{UG} \\
t_{a2} &= t_{UG} + t_{\ell i} + t_{\ell j} = t_{UG} + 2t_\ell \approx t_{UG} \\
t_{a2} &\approx t_{a1} \approx t_{UG}
\end{align*}
\] (9.7) (9.8) (9.9) (9.10)

There exists a fair number of studies of rates of language acquisition in bilingual children (see e.g., Romaine (1989), Harding & Riley (1986), Petitto et al. (2001a), and references therein). The variations between individual children are very large (as is also the case for monolingual language acquisition), but the consensus that can be extracted is that \( t_{a2} \) is possibly somewhat larger than \( t_{a1} \), but not significantly so, and by no means twice as large:

In very general statistical terms, bilingual infants and children start speaking slightly later than monolinguals, but they still remain well within the degrees of variation for monolingual children. (Harding & Riley, 1986)

Even when the onset of acquisition is delayed in the bilingual, children apparently make up for the time lost,. . . (Romaine, 1989, p. 195)

The results provided strong evidence that bilingual acquisition caused no language delays. (Wagner, 2001a, p. 509).

Romaine (1989) and Petitto et al. (2001a) also discuss another aspect of bilingual acquisition, namely the pattern of acquisition:

. . . bilingual children seem to pass through the same developmental milestones in much the same order and the same way in both their languages as monolinguals do in their respective languages,. . . (Romaine, 1989, p. 195).
Romaine (1989) takes this as evidence in favour of the innateness hypothesis, but the reasons for this are not evident — it just implies that the methods children use for acquiring languages do not depend on whether one or two languages are acquired, which in itself tells us very little about what those methods may be; cf. the next point below. The comparative acquisition times clearly do not support the innateness hypothesis.

6. There is no shortage of alternative theories of language acquisition that do not postulate an innate language acquisition device with a genetically specified grammar. They generally do postulate other innate capacities, but less detailed and less language-specific. A rough classification of language acquisition theories:

- **Empiricist theories**
  - Connectionism, reviewed by Rispoli (1999), with attached discussion and comments, pro and con: (Chater & Redington, 1999; Ellis, 1999; Feldman, 1999; MacWhinney, 1999; Maratsos, 1999; Gobet, 1999; Hahn, 1999; Valian, 1999; Plunkett et al., 1999).
  - Probabilistic and distributional approaches (Saffran et al., 1996; Redington & Chater, 1997; Plunkett, 1997; Seidenberg et al., 2002; Clark, 2001; Gerken, 2004; Mintz et al., 2002).

- **Cognitivist theories**
  - Schemas, of several types (Arbib & Hill, 1988; Mandler, 1994).
  - Functionalist approaches (Bates & MacWhinney, 1982).
  - Language emergent from cognition (Gomez & Manning, 1997; Bates, 2003).

- **Social-cognitive interactionist theories**
  - Socio-perceptual language emergence (Zukow, 1990).
  - Cultural acquisition of language (Harkness, 1990).
  - Ecological language acquisition (Dent, 1990a; Dent, 1990b).
  - Context-based language acquisition (Walczak, 2002).
  - Usage-based language acquisition (Tomasello, 2000b).

- **Neo-nativist theories**
  - Chomsky (1965)
  - Optimality (Prince & Smolensky, 1997; Tesar, 1998; Archangeli, 1999, but see Fodor (1997)).

Optimality and connectionism have the attractive feature that they are amenable to direct computer simulations of language acquisition, and appear to work, at least for the ‘toy languages’ that are computationally tractable (Prince & Smolensky, 1997; Parisi, 1997), with some modest achievements also with
natural language (Palmer-Brown et al., 2002). Morris et al. (2000) find grammatical relations emerging in their connectionist simulations. Interestingly enough, simple recurrent neural networks show the same type of bias in language learning as human learners (Kirby & Christiansen, 2003), and also similar patterns of over- and under-generalization (Morris et al., 2000). According to Womble & Wermter (2002), the addition of a simulated mirror neuron system (see Section 9.6.2) to a connectionist network can improve grammar learning considerably. Furthermore, both optimality and connectionism are eminently compatible with Chomsky’s (1982) Government & Binding grammar (Uriagereka, 1999), and only take issue with Chomsky’s innatist language acquisition model. Smolensky (1999) presents a formal grammar within a connectionist framework.

Probabilistic and distributional approaches can also be simulated with computers, but have been studied with real children as well, learning artificial ‘languages’. It turns out that small children are equipped with quite powerful statistical-learning capacities with language-like input (Saffran et al., 1996), even extracting syntactical and other patterns believed to require algebraic processing (Marcus et al., 1999; Altmann, 2002). Yang (2004), however, argues that this type of statistical learning isn’t powerful enough to handle realistic natural-language input without considerable innate scaffolding. But Mintz et al. (2002) manage to identify words and nouns in simulations of distributional learning with real CDS corpora as input. Furthermore, the existence of statistical learning in children, including apparent innate knowledge of what features are relevant for statistics-gathering, is sometimes in itself regarded as part of an innate language acquisition device. Possibly relevant in this context is that the kids manage this statistical learning not only with speech-like input, but also with tone sequences, and that the rule learning experiments of Marcus et al. (1999) have been replicated with monkeys (Hauser et al., 2002c). The presence of statistical learning outside linguistic contexts, and even in species that do not have language, implies that whatever cognitive machinery is used isn’t language-specific (Gomez & Gerken, 2000), and cannot be a language adaptation.

It should be kept in mind also that language acquisition is an iterated process — the output of language acquisition in one generation, becomes input for the next — and that both the human language acquisition equipment, and language itself, can evolve over time. The evolutionary iterated learning of Kirby & Christiansen (2003) is an attractive structure taking all these processes into account.

7. The poverty of the genes. We simply don’t have enough genes to specify in detail all the complex neural connections in a putative language organ (Mueller, 1996). After the Human Genome Project, we know that no more than 20,000 – 25,000 genes have to account for the entire human body and brain (Abdellah et
Hypotheses of language origins

Even when the number was still believed to be three times as large, this was regarded as a severe problem for any hypothesis proposing detailed genetic specification of our cognitive capacities (Buller & Hardcastle, 2000).\(^\text{29}\) The problem is exacerbated by the fact that we share the vast majority of those genes with the other apes, and even with other mammals. Changes in gene expression and regulation can explain quantitative differences easily enough — but complex and truly unique human features place a heavy burden on the tiny number of non-shared genes, if they are to be genetically specified. Humans build a brain three orders of magnitude larger than the brain of a mouse, with about the same amount of genetic information (Deacon, 2000), and largely using the very same genes. Worden (1995), discussed in Section 3.6, calculates a very low limit, a few kilobytes, on the amount of new genetic information that can have accrued in our genome since our last common ancestor with chimps — if his limit is accurate, this severely limits any innate differences between us and chimps, leaving barely enough room even for the obvious bodily differences, much less any innate universal grammar. Lorenzo & Longa (2003) argue that Chomsky’s Minimalist program requires just a small number of genes for the specification of its postulated innate components, but their argument is not compelling — see page 39. Jackendoff (2002) has a better case when he argues for a much weaker form of universal grammar, with only fragments of rules serving as scaffolding for language acquisition.

Brain development is highly plastic (see page 111), and dependent upon the sensory impressions received at an early age (Wong, 1995; Mueller, 1996). Those systems that are understood in more detail (like vision; see page 111) are not genetically hardwired in the detailed sense that an innate universal grammar would need to be.\(^\text{31}\) Instead, only the rough outlines are laid down genetically — the optical nerve is led to the occipital lobe of the brain under genetic control — and the detailed neural connections are then gradually formed and pruned, in response to the sensory data received during a critical period. That language

\(^{29}\)This estimate has been gradually going down, as our knowledge of the human genome has improved. The current number is from the near-final version published in 2004. According to Abdellah et al. (2004), the upper limit of 25,000 is quite firm. Estimates from the draft version of the genome were on the order of 30,000 (Pääbo, 2001; Claverie, 2001). The reasons for this discrepancy are discussed in Abdellah et al. (2004). Earlier estimates were several times larger, up to 100,000.

\(^{30}\)Two caveats are in order here:
(a) Processes such as alternative splicing can produce more than one protein from one gene.
(b) New patterns of gene expression can re-use the same gene in new contexts. But neither of them changes the effective number of genes by the order of magnitude needed to affect the substance of the argument. The first effect is estimated to given an increase on the order of 50%, with Abdellah et al. (2004) reporting a total of 34,000 transcripts coming from 22,000 genes. Marcus (2004a; 2004b) attempts to counter the ‘poverty of the genes’ argument, invoking these arguments as well as general developmental biology, but he is successful only in showing that the brain can have innate structures at the gross level where everybody agrees that it has them. His arguments are insufficient to address innateness at the detailed level required for a genetically hardwired Universal Grammar.

\(^{31}\)Smell is an apparent exception (Barinaga, 2001), but smell is evolutionarily ancient, and does use up a very large number of genes, about 5% of the total number of genes in our entire genome.
acquisition is not handled by a hard-wired device in a specific location in the ‘language areas’ of the brain is demonstrated by the fact that language acquisition can follow the normal pattern even if the entire left hemisphere of the brain is absent (Stowe & Haverkort, 2003).

9. It is well established that children from different backgrounds are all equally able to acquire any human language — if placed in the right environment at a tender age, children of Chinese descent or Basque descent or whatever are equally good at acquiring either the language of their own ancestors or some other language. There are no signs of innate predispositions to acquire any specific language. But the speakers of some language families have been isolated from the bulk of humanity for millennia, time enough for some genetic fine-tuning of bodily features to take place — why, then, hasn’t their innate LAD been fine-tuned for their language family, with pre-set parameters or whatever? This argument is due to Lieberman (2003b), invoking the example of Tibetans, who are genetically adapted to high altitudes (Beall et al., 2004), but not genetically adapted to learning Tibetan. I am not entirely convinced of its validity, but it is at least worth contemplating, even though its force is blunted considerably by Lieberman (2003b) on the next page in the same paper arguing that there has been insufficient time for UG to evolve, hardly consistent with the argument above that’s based on there having been too much time for UG not to be fine-tuned for individual languages.

10. The ape language experiments reviewed in Chapter 7, to the extent that their results are accepted, argue against the necessity of innate language-specific abilities (cf. page 133).

The innateness issue is far from settled. There is a disturbing tendency for the debate to be split along disciplinary lines (Yang, 2004), with mainly linguists on the innateness side, and mainly psychologists and cognitive scientists on the other, which indicates a lack of adequate communication between the fields.32 There is some merit in the arguments of both sides. On one hand, the arguments for underlying universals in the structure of human languages are compelling — but on the other hand the successes, however modest, of computational and statistical models for language acquisition indicate that the supposed impossibility of language acquisition without a Chomskian innate Universal Grammar may not be as absolute as claimed. Innate capacities may well be needed for language acquisition, but it is very far from established that these capacities need to be language-specific in any strict sense, much less that a hardwired Universal Grammar is needed.

Furthermore, the ‘poverty of the genes’ argument together with the compelling evidence of brain plasticity in ontogeny show that strong claims of a complex and fully genetically determined innate grammar are untenable. But that doesn’t make the ‘poverty of the stimulus’ argument go away, leading to what Dor & Jablonka (2001) call ‘The paradox of domain-specificity’.

32But the growth of cognitive linguistics may signal a bridging of this gap.
The impact of the innateness issue on language evolution is actually rather modest, if subtle. Nobody doubts innateness in Chomsky’s kitten sense (see page 178), which implies that some language-relevant, if not necessarily language-specific, genetic changes must have taken place along the human lineage, since the last common ancestor we shared with kittens, which was a primitive mammal sometime in the Cretaceous, perhaps 100 million years ago (Murphy et al., 2001). Many, but not all, of those changes can be located to the last five million years, after we and the other apes parted company in the family tree — there is certainly a difference in language abilities, not only between young Miss Chomsky and her kitten, but also between the gorilla Koko and her kitten.\(^33\)

Innateness does have an impact on the issue of gradual vs. sudden language evolution, as mentioned above. Innate complex features cannot evolve suddenly with any reasonable probability — intermediate steps are necessary. Even with the rather modest degree of innateness that is empirically well-established, a gradual transition is more biologically plausible, contra Chomsky (1988).

Nevertheless, even though an innate grammar may not be necessary for language acquisition, this does not prove that humans don’t have one anyway, since an innate grammar would certainly facilitate language acquisition, even without being strictly necessary. The Baldwin effect, described on page 30, implies that if language has been a central part of human behavior for a sufficiently long time, an innate predisposition to acquire language is a possible result. But an innate predisposition is a far cry from an innate Universal Grammar — it may be nothing more than a disposition to attend to human voices, and some biases in what aspects and types of patterns to pay attention to. There are also question marks concerning the circumstances under which the Baldwin effect would be effective, and whether they actually apply in this case (Deacon, 2004b); cf. page 30.

Furthermore, Deacon (2004b) argues that it may well be the case that innate mechanisms aren’t just unlikely to evolve — they would be evolutionarily unstable if they did, subject to ‘devolution’. Even if we were somehow miraculously endowed with a full innate genetically specified Universal Grammar, if at the same time we had efficient learning mechanisms, which we do, this would relax the selection pressure needed to weed out random mutations in the ‘grammar genes’, which would eventually ruin them, much the same way that our unnecessary genes for vitamin C production have been ruined (Nishikimi et al., 1994; Ohta & Nishikimi, 1999). It is far from clear whether Baldwin-like processes, building up innate structures, or devolutionary processes, tearing them down, predominated in our evolutionary history.

On the non-genetic side of the issue, we need to consider the different levels of language-related evolution, discussed on page 31. Very little empirical data is available concerning the cultural or memetic evolution of language in the relevant timeframe, but it would nevertheless be an error to discount such processes and

\(^33\)Yes, the gorilla Koko (see page 135) also had a pet kitten (Patterson, 1981).
focus exclusively on the biological evolution of a hypothetical innate language acquisition device. Even though little is known, it would be highly surprising if language, regarded as a memetic-type entity, did not change over evolutionary time. As discussed in Section 3.5.1, selection for both improved learnability and communication can be expected to occur. In parallel with this memetic evolution, there may be biological evolution towards an innate language acquisition device — but memetic evolution is a much faster process, so the result is likely to be biased towards languages that are easy to learn, rather than towards learners who are innately good at learning languages (Bull et al., 2000).

In conclusion, it may well be that the final resolution of the innateness debate will be a compromise, with coevolution of language memes and acquisition genes (Kirby, 1996). It is certainly not a simple black-or-white dichotomy. Quite possibly innate features and biases will be identified in human children on many different levels, some very general and possibly some more focused on language acquisition — but today we are still far from pinpointing what these features may be. The possibility that language is to some extent shaped by neither nature nor nurture, but instead by more general principles such as semiotic constraints (Deacon, 2003b), is also interesting.

9.8 Summary

- Adaptation vs. spandrel
  - Not either-or, has to be some of both.
  - Spandrel/exaptation:
    - Cannot adapt for language until language already present ⇒ First step towards language must be based on spandrels/exaptations.
    - Many features that we use for language already present in other apes ⇒ Exaptations.
  - Adaptation:
    - Obvious selective value today — fitness of a language-less human near zero.
    - Complex package appearing to be designed for its current function.
    - Some features fine-tuned for language use.
  - We have biologically adapted to language use, and language has culturally adapted to be used by us — but which process is more important?
- Early vs. late
  - Speech adaptations detectable in fossils:
    - Hearing fine-tuned.
    - Breathing control enhanced.
    - Hyoid bone in modern form.

All of the above present in Neanderthals, and by implication in the common ancestor of us and Neanderthals, 500,000 years ago.
Symbolic behaviour:

- Archeological signs of early symbolic behaviour:
  - Engravings.
  - Ornaments.
  - Pigment use.
  - Burials.

- These signs do not suddenly appear 40,000 years ago, as commonly believed.
  - Early gradual appearance of more and more signs of symbol use, across at least 100,000 years, mainly in Africa.
  - Possible hints of symbolic behaviour outside *Homo sapiens* as well, in Neanderthals and possibly *H. heidelbergensis*.

- Early appearance of speech adaptations and symbolic behaviour rules out a late appearance of language. Our ancestors 500,000 years ago had some form of speech, if not necessarily full human language, and our ancestors at least 100,000 years ago, and possibly 500,000 years ago had some symbolic capacity.

**Gradual vs. sudden**

- Two-pronged argument for gradual appearance:
  - Fossil and archeological signs of language do not appear suddenly all at once — see the previous point.
  - Language is a complex adaptation. To the extent that language has a biological basis, it must be a matter of many genes. Lots of coadapted genes do not suddenly appear together in a coordinated package, but have to coevolve gradually. Furthermore, some of our features are fine-tuned for language.

- How gradual is gradual?
  - What is strictly ruled out is single-step saltationism.
  - Gradual evolution need not be geologically slow — a process that takes 10,000 years will still look instantaneous in the fossil record.
  - The actual time needed for language to evolve depends on many factors, including:
    - How much of our language ability is based on pre-existing exaptations, and how much new features are needed?
    - How much biological evolution, and how much cultural evolution?

**Speech first vs. gestures first**

- Speculations that the first language may have been a sign language have a long history, from Condillac (1746), and are still popular.

- Arguments for gestures first:
  - Apes have both the dexterity and the cortical control needed for gesturing, but not for speaking.
  - Easier to imagine the very earliest stages of proto-language, with mimesis and iconic proto-words, in a gestural system rather than a spoken one.
  - Mirror neurons provide a possible path into iconic gestures — but monkeys have these neurons as well, so why don’t they gesture?

- Arguments for speech first:
  - Speech is universal among human cultures today.
- If gestures were first, an additional evolutionary step, the switch from gestures to speech, is required. More parsimonious to postulate that speech was first.
- Either gestures first or speech first remains tenable. Insufficient evidence to exclude either possibility.
- It need not be one or the other — the earliest forms of language may well have used both.

Innate vs. learned grammar

- Arguments for innate and genetically specified:
  - Universals in language.
  - Poverty of the stimulus, and related learnability issues.
  - Patterns in language acquisition
- Arguments for learned and culturally emergent:
  - Poverty of the genes.
  - Brain plasticity in ontogeny.
  - Alternative views of language acquisition.
  - Some language abilities present in non-humans.
  - Language can memetically adapt to our brains faster than we can genetically adapt to language.
- None of the possibilities has overwhelming support. Particularly, the case for a genetically specified grammar and an innate language acquisition device is not nearly as strong as is commonly believed.

Further reading

CHAPTER 10

WHY DID LANGUAGE EVOLVE?

In order to understand the evolution of language, we must advance our understanding of the purpose of language, and ‘purpose’ in an evolutionary context is synonymous with selective advantage (Ganger & Stromswold, 1998). Different hypotheses concerning the original selective advantage that drove the evolution of language will be discussed in this chapter.

Evolution of complex and specialized features does not occur without being driven by some selective pressure — some evolutionary advantage accruing to those who possess the feature. As discussed in Section 9.3, it is well established that language is such a complex and specialized feature (Pinker & Bloom, 1990). But what was the crucial advantage conferred by language, that drove its evolution? One may think that the advantages of having a language are obvious — as Lieberman (2003a) points out “it is difficult to identify any aspect of human behavior (...) that would not profit from [...] language, ...” (p. 670) — but that would instead raise the question of why only humans have acquired it, why not a lot of other animals as well, if it is so useful?

Szathmáry (2001) identifies two possible explanations for the uniqueness of an apparently useful feature, such as language:

1. **Variation-limited:** The requisite combination of mutations has a very low probability of occurring.
2. **Selection-limited:** The feature will bring a selective advantage only under very rare circumstances.

The first possibility is in principle conceivable in the case of language, but it would mean that the appearance of language in the human lineage, rather than that of chimps or cockroaches, is a matter of pure chance. In that case, the problem of language origins is not amenable to analysis, and not very interesting in any case (though Szathmáry (2001) calls the possibility ‘amusing’). The miraculous saltationist models of language origins formerly favored by e.g., Chomsky (1988) and Bickerton (1990), discussed in Section 9.5, belong in this category.

The second possibility is more interesting, and does appear more plausible, as humans have many other unusual features, that may provide the unique circum-
stances that would favor language emergence. In that case, the problem at hand is to identify whatever it was in human history, that made language particularly advantageous for our ancestors, but not for the other apes. This has close parallels with hypotheses concerning why our brains are so much larger, discussed in Section 5.3.1. The different levels of evolution, as well as the features and limitations of evolutionary processes, reviewed in Chapter 3, should also be kept in mind.

A particularly common pitfall here is the teleological scenario — that our ancestors evolved language because they needed it for some purpose or other. Future utility of a trait is never valid as an explanation for its evolution, as this entails backwards causation. Tenable hypotheses must be structured around selection scenarios, instead — why did people with incipient stages of language have a reproductive advantage over people without?

The issues raised in Section 6.3 must also be considered — the selective advantage of language must actually benefit the spread of the speaker’s genes, not just the general welfare of the group. Hypotheses of language evolution cannot be based exclusively on how information recipients benefit from language, without explaining how that benefit spreads back to the speaker.

Another important aspect, related to the previous points, is the context in which language evolved. Before the advent of agriculture (which certainly postdates language), humans lived as hunter-gatherers, in modest-sized tribes, presumably with a lifestyle and social structure not vastly different from that of the few remaining present-day hunter-gatherers. The chimpanzee lifestyle can reasonably be called hunter-gatherer as well, though it is quite different from that of human hunter-gatherers. Thus it appears not unreasonable to assume that language evolved in a hunter-gatherer context, in tribes with a social structure somewhere in between that of chimps and modern hunter-gatherers. This would be the ‘Environment of Evolutionary Adaptedness’ (EEA), the Stone Age environment for which evolution has shaped us — the time since we abandoned the hunter-gatherer lifestyle is much too short for any substantial evolutionary changes (Daly & Wilson, 1999; Nesse & Williams, 1994; Byrne, 2000, but see also Irons (1998)). Reasonable hypotheses of language evolution must postulate that language carried some crucial advantage for people in such an EEA society — advantages that language confers only in modern industrialized (or even farming) societies are irrelevant.

Numerous hypotheses have been proposed as answers to these questions about human language evolution, a selection of which will be evaluated here. In order to provide some structure and overview, the hypotheses will be classified under several categories, though the boundaries between the categories are sometimes fluid.
10.1 Hunting

An obvious starting point is to postulate that language evolved for the purpose of communication, and that the main selective advantage gained by improved communication was enhanced coordination of group activities. Prominent among the group activities discussed in this context is hunting, which has played a central role in many scenarios for human evolution (Landau, 1991; Sagan, 1977), but more general resource acquisition activities can be included here as well (Cziko, 1995). Apart from communication during hunts, the hunting argument has also been connected with the teaching argument of Section 10.4 below:

The ‘quality education’ needed to become an expert Pleistocene hunter could not do without a complex form of information transmittal interaction in which the transcendence of the *here and now*; [...] played a key role. (Roebroeks, 2001, p. 451, emphasis in original).

Hunting-related communication could be a significant force in our evolution, only if hunting was actually of major importance for our subsistence. Contrary to this, it is commonly asserted that most hunter-gatherers, at least in the tropical areas relevant to early human evolution (Ragir, 2000), don’t get nearly as much food from hunting as from gathering. Kaplan et al. (2000), however, contest this claim, and review a number of studies of actual calories hunted and gathered by members of various tribes — on average, an adult hunter produces twice as many calories per day as an adult gatherer. Kaplan et al. (2000) furthermore trace the ‘common knowledge’ that hunting is unimportant back to a single study of questionable generality (Lee, 1979, cited in Kaplan et al. (2000)).

The Inuit, who live in an Arctic environment with little plant food to gather, incontrovertibly get most of their food from hunting, and very likely so did the Neanderthals of Ice Age Europe, for similar reasons. Isotopic evidence from Neanderthal fossils indicates that meat from large herbivores was a major part of their diet, which implies that hunting was an economically vital activity (Richards et al., 2000; Bocherens et al., 2001). But living in a glacial environment is a very late development in human history, well after *H sapiens* and Neanderthals had gone their separate ways, so an Arctic origin of language does not appear plausible.

Nevertheless, with the evidence presented by Kaplan et al. (2000) it would appear that hunting is highly significant for modern human hunter-gatherers, and by implication has been important for at least the later part of our evolution, so that a role for hunting in the evolution of language is not excluded on these grounds.

There are, however, a few problems with the notion that hunter coordination was a major driving force in language evolution. To begin with, modern humans do not use all that much language during a hunt — it is a rather silent activity (Dunbar, 2003b).

Furthermore non-human social carnivores manage to coordinate their collaborative hunts without using language (Brinck & Gärdenfors, 2003). This of course
includes wolves and lions — but also chimpanzees (Mitani & Watts, 1999; Plummer & Stanford, 2000; Mitani & Watts, 2001; Stanford, 1998) and orangutans, who do hunt on occasion (Utami & van Hoof, 1997). The observation that chimpanzees hunt caused considerable surprise when first reported by Goodall (1963) — until then, our cousins had been regarded as peaceful frugivores in stark contrast with the bloody history of Man (Dart, 1953; Ardrey, 1961). Hunting by other carnivores may not be so relevant in the present context, but hunting by our nearest relatives certainly is.

Orangutan hunting is rare and mostly opportunistic, but chimpanzee hunting appears more organized and purposeful. Chimpanzee hunts take place in modest-sized groups (almost exclusively male),¹ and give the impression of being collectively planned and coordinated in advance. The communication systems that chimps have in the wild are to all appearances perfectly adequate for hunting purposes, as evidenced by a hunting success rate exceeding 50% (Mitani & Watts, 2001), which compares favorably with human hunting prowess.

There is some evidence that chimpanzees in more open savannah habitats hunt more often than their rainforest relatives, as do savannah-dwelling baboons (Domínguez Rodrigo, 2002). This is of some interest in the context of human history, as our early ancestors also appear to have lived in fairly open habitats, woodland or savannah rather than rainforest.

Gathering may well be as plausible as hunting as an explanation for language, since gathering in modern human hunter-gatherers relies on an encyclopedic knowledge of thousands of edible plants and other items,² the communication and discussion of which may well be highly advantageous (Pinker & Bloom, 1990). And gathering, particularly in a savannah environment, involves much more than just picking fruits off branches — a major part of a typical human gatherer diet consists of tubers and other underground plant parts, difficult to find and extract, and commonly laced with noxious chemicals, requiring extensive preparation to render them edible (Ragir, 2000), multiplying the demands on learning and thus communication.

As noted earlier, a plausible hypothesis of language evolution should explain not only why our ancestors did evolve language, but also why the chimp ancestors

¹Chimpanzees do not hunt out of economic necessity — hunting is less common when fruit is scarce (Watts & Mitani, 2002). Instead, the evidence indicates that the main function of chimpanzee hunting is male bonding and coalition building (Mitani & Watts, 2001; Stanford, 1998). It is not entirely obvious how this differs from hunting among humans in modern Western countries. O’Connell et al. (2002) argue that early hominid meat consumption followed a similar pattern.

²Note, however, the counterargument of Alvard (2003):

Language could not have evolved initially to facilitate the passage of a complex database of knowledge because, in the absence of language to produce it, the database did not yet exist. (p. 143).

This is a valid point, or would be if no such database existed in our non-speaking relatives — but the corresponding database that chimpanzees possess is not entirely negligible (Huffman, 1997), and the extension of the database that language makes possible, might well contribute to the selective value of language once it has evolved.
didn’t. And in the case of gathering it is not obvious that there should be any notable difference between the two, though the human reliance on underground food mentioned above is a possibility. Concerning hunting, chimpanzees do hunt, as mentioned earlier, but they don’t do it very often. Meat is regarded by them as a rare delicacy, but it does not contribute significantly to their subsistence, unlike the situation for human hunter-gatherers.

It is a matter of some controversy when hunting became economically important for our hominid ancestors. There is evidence for meat-eating and meat-processing quite early in the fossil record — isotopic hints of a dietary shift are found in 3-million-year-old Australopithecus africanus (Sponheimer & Lee-Thorp, 1999), and bones with cut marks from stone tools first occur around 2.5 million years ago (Semaw, 2000) and become common from about 2 million years ago (Brantingham, 1998; Domínguez Rodrigo, 1997; Capaldo, 1997; Fernández-Jalvo et al., 1999b), in the context of Homo habilis or possibly Homo ergaster (O’Connell et al., 2002).

But evidence of meat-eating need not be evidence of active hunting. It is not obvious to what extent animals were actively hunted by early hominids, and to what extent meat was scavenged from carcasses killed by carnivores. Scavenging is argued on archaeological (O’Connell et al., 2002) and ecological grounds (Brantingham, 1998), whereas cut marks are invoked both to support hunting (Domínguez Rodrigo, 1997) and scavenging (Capaldo, 1997). One should perhaps also make a distinction here between ‘active’ scavenging — driving away a predator from a fresh kill — and ‘passive’ scavenging — picking over carcasses after the primary predator has left them. Active scavenging is archeologically difficult to distinguish from hunting — Domínguez Rodrigo & Pickering (2003) conclude that hominids around 1.5 million years ago got their meat either through hunting or active (but not passive) scavenging. But Domínguez Rodrigo (2002) was skeptical about early hominid active scavenging, given the risks involved in confronting a lion or saber-toothed cat over its fresh kill.

It may be relevant to note here that chimpanzees get all or almost all of their meat from active hunting — according to Plummer & Stanford (2000) they never scavenge, but Mitani & Watts (1999) and Domínguez Rodrigo (2002) say that they do scavenge occasionally. According to O’Connell et al. (2002) they engage in active but not passive scavenging. Studies of the remains of chimpanzee meals known to be the result of active hunting, like Tappen & Wrangham (2000), may yield clues to the interpretation of fossil bone assemblages.

An argument against scavenging, at least of the passive variety, is that unlike e.g., jackals we get sick if we eat rotting meat. The digestion of humans, and probably primates in general, is not adapted to handle the bacteria and toxins in decomposing carrion (Ragir, 2001), which it ought to have been if passive scavenging had been important in our history.
Even in the case of Neanderthals and early modern humans, there have been some arguments about whether they hunted or scavenged, but more recent evidence clearly supports hunting in these cases (Marean, 1998; Milo, 1998; Richards et al., 2000; Bocherens et al., 2001). As a matter of fact, it appears that Neanderthals hunted *more* than modern humans, who had a broader diet (Richards et al., 2001). Under the hunting hypothesis this would imply that Neanderthals, who according to Balter et al. (2001) obtained 97% of their sustenance from meat, would have had even more reason to evolve language than we did.

For a recent review of the issue of hunting versus scavenging among early humans, see Domínguez Rodrigo (2002).

Regardless of whether the meat was hunted or scavenged, there is evidence for a major dietary shift in early hominids, somewhere between 3 and 2 million years ago, from a diet similar to that of chimpanzees, to one including significant amounts of meat from larger animals. It is intriguing that the timeframe coincides with the emergence of the genus *Homo* and the first stone tools, as well as the start of human brain growth. New communicative needs associated with this new lifestyle are not ruled out as explanations for language. But communication about scavenging or gathering is at least as likely as communication about hunting.

A variation on the hunting theme is the proposal of Stanford (1999, cited in Heesy (2000)), that the triggering factor was the social machinations involved in meat-sharing in the group, rather than in the demands of hunting *per se*. Closely related is the dual economy, the division of labor between hunting men and gathering women, with the organized cooperation and exchange of food that it entails, invoked by Quiatt (2001) as an explanation for the emergence of language. But both of these proposals belong rather in Section 10.5 below, as special cases of the social hypotheses discussed there.

Another variation is that of Deacon (1997), where male cooperative hunting in conjunction with our mating system drove the evolution of language — see Section 10.3 below.

### 10.2 Tool making

Tool use and tool making is not entirely limited to humans. The *making* of tools is reported only for chimpanzees (Savage & Wyman, 1844, cited in Whiten & McGrew (2001)), bonobos\(^3\) (Boesch & Boesch, 1990), elephants (Hart et al., 2001), and crows (Stone, 2002; Hunt & Gray, 2004), but numerous species have been found to *use* tools, including all the great apes (Sugiyama, 1994; Nakamichi, 1998; Peters, 2001; O’Malley & McGrew, 2000; Fox & bin’Muhammad, 2002; van Schaik et al., 2003b) and some monkeys (van Schaik et al., 1999; Westergaard

---

\(^3\)Bonobos can even be taught to make stone tools (McNeil, 1996; Savage-Rumbaugh & Lewin, 1994).
et al., 1998; Worch, 2001; Hauser et al., 2002b), as well as numerous others, from spiders to naked mole rats (Shuster & Sherman, 1998). Chimpanzees even use tool-composites, several tools used in conjunction to achieve a goal (Sugiyama, 1997), and adapt their tool making to the task at hand (Humle & Matsuzawa, 2002). That chimpanzees use tools has been known at least since the 19th century (Savage & Wyman, 1844), but the uniqueness of human tool use was nevertheless still argued a century later (Oakley, 1956, cited in Ambrose (2001)).

Still, tools do appear to have played a major role in human evolution, as reviewed by e.g., Schick & Toth (1993) or Ambrose (2001), and are frequently invoked as one of the factors that drove the evolution of hominid brains and intelligence, a hypothesis gaining some support from the complexity found in the brain activity of modern humans during stone-tool manufacture (Stout et al., 2000). This argument is commonly connected with the belief that the early tools were mainly hunting weapons, a belief not borne out by observations of chimpanzees, who use tools regularly for gathering, but rarely for hunting (Tappen & Wrangham, 2000). As for the driving forces behind tool-use evolution itself, see the review of van Schaik et al. (1999).

A few authors, such as Gibson (1990, cited in King (1996)), have invoked tool making as a driving force also behind the evolution of language. But it is not entirely obvious just why language would be of such selective advantage for tool making. In general, flint knapping is a solitary activity in which language plays little role even among those modern humans who still make stone tools. Teaching it is typically done by demonstration rather than verbal instruction (Dunbar, 2003b). It appears implausible to have improved social communication evolving due to pressures that have little to do with social communication.

The oldest recognizable stone tools are about 2.6 million years old (Semaw et al., 2003), and bone tools of comparable age have been found as well (Backwell & d’Errico, 2001; Shipman, 2001). It is not known whether less durable tools, like wooden sticks, were also used by early hominids, as they are by chimpanzees, since such tools are rarely preserved; the diversity of tools in the fossil record is likely to be underestimated. Likewise, the use of unmodified natural stones as tools is difficult to recognize.

Pseudo-archeological excavation studies of known sites of chimpanzee tool use, like that of Mercader et al. (2002), may well provide the means to recognize the remains of more primitive tool use among our earliest ancestors (Vogel, 2002; Bower, 2002). Mercader et al. (2002) studied the remains of chimpanzee nut cracking, their main use of stone tools. Similar remains of nut cracking do oc-

---

4But see Bridgeman (2002) and Simão (2002) who caution against assuming that improved tool manufacture necessarily has a biological basis — cultural evolution is a distinct possibility as well.

5The oldest known wooden tools are less than 400,000 years old (Klein, 2000), in the probable context of Homo heidelbergensis.
cur in the human archeological record, but do not predate the oldest regular stone axes (Goren-Inbar et al., 2002).

Hand anatomy may provide another clue to when tool use first became important. Chimpanzee hands are a compromise between the demands of quadrupedal walking, tree climbing, and object manipulation, and are not optimal for any of them — for example, the chimpanzees do not have the forceful human precision grip with the thumb and index finger. But already *Australopithecus afarensis* may have had it (Alba et al., 2003; Tocheri et al., 2003), well before the appearance of stone tools in the archaeological record, and later australopithecines almost certainly did (Aiello, 1994). It is, however, unclear whether the hand modifications are due to selective pressure for tool making, or simply due to relaxed selection for tree climbing and knuckle walking.

For the first million years or so, a simple stone tool kit called Oldowan was used; a typical Oldowan tool is shown in Figure 10.1. Interpretations differ on whether there was any significant change through time, with de la Torre et al. (2003) and Kimura (2002) supporting change and Semaw (2000) stasis. The Oldowan industry is mainly associated with *Homo habilis* (Kimbel et al., 1996), though tool-using robust australopithecines remains a possibility, according to Susman (1994; 1998), and both Parker (2002) and Semaw et al. (2003) connect the earliest tools with *Australopithecus garhi*; cf. page 63. Some early *H. ergaster/erectus* may have retained Oldowan tools as well.

After the Oldowan, and coincident with the arrival of *Homo erectus*, the Acheulean tool kit, with more advanced and consistently shaped hand-axes, also shown in Figure 10.1, became widespread instead. The significance of this change is, however, controversial (McPherron, 2000). Aiello (1998) attributes it to the same cognitive breakthrough that he postulates as the first step in the evolution of language, whereas Wynn (2002) proposes a different cognitive breakthrough related to symmetry and the ability to impose a preconceived shape on a lump of rock.

For another million years, throughout the lifespan of *Homo erectus* and possibly a bit beyond, the Acheulean tools underwent little change. Klein (2000) states that “[l]ater Acheulean bifaces tend to be more extensively flaked and more carefully shaped... (2000, p. 23), and Wynn (2002) finds more complex symmetries in late Acheulean tools, but the basic design remained unchanged. Even more remarkable, the rate of technological diffusion was so slow that Oldowan tools were still used in Europe as late as less than a million years ago, even though they had been obsolete in Africa for almost a million years by then (Roebroeks, 2001).

The Mousterian tools that are usually associated with Neanderthals (Churchill & Smith, 2000) were likewise fairly uniform in time, from half a million years ago onwards, though some technological progress has been observed through time.

---

6 Sometimes spelled Acheulian.
7 There is not a simple one-to-one correspondence between species and stone-tool industries (Davidson, 2003).
Why did language evolve?

Figure 10.1. Examples of tools from two early cultures. Left: Oldowan, commonly associated with habilines. Right: Acheulean, commonly associated with Homo erectus.

at Neanderthal sites that were occupied during many millennia (Simek & Smith, 1997; Otte et al., 1998), as well as adoption of some inventions by late Neanderthals (Churchill & Smith, 2000), presumed to be copied from Homo sapiens (Hublin et al., 1996), though Wynn & Coolidge (2004) argue for emulation rather than copying, and d’Errico et al. (2003) for independent invention.

And then Homo sapiens came along, and with us a clear acceleration of the pace of technological and cultural innovation, an acceleration that is still with us today. Here we have the main role of tools in the debate about language evolution — the sudden increase in creativity is interpreted by many, e.g. Mellars (1998), Diamond (1991), and Donald (1997), as evidence that we had suddenly acquired language, or at least speech (Corballis, 2002), though a large number of other hypotheses have also been proposed (Gabora, 2003; Carruthers, 2002b). However, as discussed in Section 9.4.1, the suddenness is largely an illusion. There was no sudden revolution, and thus no need for a sudden cognitive leap (as in sudden language acquisition) to explain the revolution (d’Errico et al., 2003).

Ambrose (2001), citing earlier work by among others Greenfield (1991), discusses another possible role for tool making in the origins of language. The mak-
ing of composite tools is a sequential and combinatorial activity — but combining multiple elements into a structured sequence is characteristic also of language. In this model, tool making coevolves with syntactical language, but one may also consider the possibility that composite tool making evolved first, and provided mental tools that were perfect exaptations for language (Wildgen, 2004).

10.3 Sexual selection

The core of natural selection is reproductive success — no matter how successful you are by other measures, if you do not reproduce you’re an evolutionary failure. Sexual selection, the selection of certain features because they directly influence the mating success of their bearers, therefore plays a prominent role in evolutionary theory (Darwin, 1871), accounting for innumerable features from peacock tails to birdsong to horns (Emlen, 2001). If one observes (or participates in) the mating rituals of *Homo sapiens* today, it is obvious that language plays a non-negligible role. The possibility that this may have been true also for early hominids is considered by many authors, e.g. Cziko (1995) and Wildgen (2004). In this scenario, the selective advantage that drove language evolution may simply have been that the better speakers were preferred as mates, and so got more offspring (Miller, 1999, cited in Dunbar (2003b)). A variation on the same theme is the hypothesis of sexual conflict as a driving force (Aiello, 1998), as is the observation that verbal skills can be translated into political power, which in turn enhances reproductive success (Pinker & Bloom, 1990) — as Henry Kissinger reputedly expressed it: “*Power is the strongest aphrodisiac*.” This ‘political hypothesis’ is further discussed on page 211 below. Another related idea is that human vocal capabilities, and thus speech capacity, may have been shaped by selection for the ability to produce pleasant, modulated, musical — sexy — sounds (Darwin, 1871).

But it is not sufficient that it is plausible that the fitness of language is based on sexual selection — we also need to explain why chimps and other apes did *not* evolve language if it’s so great for your sex life. Human sex life is certainly very complex, with our officially monogamous but actually semi-polygamous system8 (Diamond, 1991), with multiple mating strategies available for both men and women (Gangestad & Simpson, 2000), with the strategies at least partially having a biological basis, under hormonal control (Thornhill & Gangestad, 1999). Such a complex system would enhance the adaptive value of language. But it is not obvious that the other apes have simpler mating systems:

* Gibbons* (1998b) were long thought to have a purely monogamous system, similar to the official human one, but have turned out to have an actual system very

---

8Genetic screening of human infants, done on more-or-less random samples for other reasons than to determine paternity, regularly turn up on the order of several % children, even within superficially monogamous families, who were not fathered by the ‘official’ father (Salter, 1996; Marlowe, 2000).
close to the actual human one instead (Brockelman et al., 1998; Gibbons, 1998b; Fuentes, 2000).

**Gorillas** were likewise believed to have a fairly simple system, based on the physical dominance of a single large male, who monopolized a harem of females. But also in this case, careful observations have shown that the actual system is more complex, with ‘illicit’ matings with other males than the harem owner very common, at least in some gorilla populations (Robbins, 1999), though Parnell (2002) finds evidence of a more nearly pure harem system elsewhere.

**Bonobos** are probably the most sexually active of all the apes, including us, and use sex for a multitude of purposes apart from reproduction — sex for friendship, reconciliation, or even pure recreation, is ubiquitous, in all conceivable combinations of gender and number (Savage-Rumbaugh & Lewin, 1994; Hohmann & Fruth, 2000). They are also famous for having the same taste as humans in sexual positions, with the missionary position and other face-to-face positions being the most common, unlike most other mammals, including the other apes, where the male enters the female from behind.

**Common chimpanzees** may have the most complex system of all, with both males and females choosing between several different mating strategies (Goodall, 1971; de Waal, 1998; Gagneux et al., 1999b). A female chimp may some of the time engage in promiscuous mating within the group, at other times seek out partners in other groups for a quick discreet mating in the forest, or form temporary monogamous relationships with a preferred male. Males may either try for political power in the group, thereby gaining preferred access to females, or try to mate on the sly with willing females, either in his own group or from another.

Empirical data on the prevalence and success of different strategies are not readily obtained, and the data on patterns of paternity within chimpanzee communities are contradictory (Vigilant et al., 2001).

**Orangutans** are solitary and arboreal, and difficult to study in the wild. Until recently, little was known of their sex life. Nadler (1977) describes orangutan sex in captivity, but the context is quite unnatural, and unlikely to be informative. According to Schwartz (1987), they are mainly monogamous, but more recent studies contradict this. To begin with, their considerable sexual dimorphism argues against monogamy (see page 204 below), as does their territorial structure, with large male territories each containing several smaller female territories (Fuentes, 2000). According to Singleton & van Schaik (2002) the orangutan system is quite complex, with different males adopting different strategies. Dominant males are able to semi-monopolize a group of females, apparently with the consent (and sometimes active cooperation) of the females, but other males roam the forest and may either attempt to force matings, or find willing females when the dominant male is absent. Homosexual behavior between males has also been observed (Fox, 2001).
There is no strong reason to believe that language is more advantageous in the human mating system than in any of the others. Particularly the gibbon system is quite similar to ours, including similar levels of paternal care, something which the other apes are not known to engage in (Ross & MacLarnon, 2000). Furthermore, there is some evidence that the mating system of humans has changed during the past few million years.

One fossilizable mating-system indicator is male-female dimorphism, reviewed in Plavcan (2001) — in monogamous species, males and females are very similar, whereas in polygynous species males are commonly larger and more robust. Gorillas, with their enormous and fearsome males and comparatively tiny females, have a typical polygynous system, whereas gibbon males and females, with their quasi-monogamous system, are hard to tell apart. Humans and chimps have similar levels of dimorphism, which might have been interpreted as a sign of continuity in mating systems since the last common ancestor — but some australopithecines are commonly believed to have had a male/female size ratio comparable to gorillas (Silverman et al., 2001; Ward et al., 2001), hinting at a strongly polygynous system. Reno et al. (2003), however, contest the earlier dimorphism studies, arguing that the dimorphism of *Australopithecus* is not significantly different from that of modern humans.

In other primates, the canines are commonly used for intraspecific displays and conflicts between males, and unsurprisingly males of polygynous species typically have very large canines. In humans, canines are not used for this purpose, and are correspondingly modest in both sexes — but it is intriguing and puzzling to note that this trend towards smaller and non-dimorphic canines was begun among the australopithecines, where the opposite trend would naïvely be expected if they really were highly body-size-dimorphic (Plavcan, 2001). Their canine size makes more sense in the context of the results of Reno et al. (2003) mentioned above.

From *Homo ergaster* onwards, at least, there is consensus on a human-like modest degree of sexual dimorphism (Wrangham et al., 1999; Plavcan, 2001), which might be taken as a sign that the human mating system had been the same since then. But Dupanloup et al. (2003) present genetic results that can be interpreted as evidence that the human mating system was mainly polygynous in the very recent past, within *Homo sapiens*. The issue of prehistorical human mating systems remains open.

There is one aspect of the human mating system that is novel among apes, and that Deacon (1997) considers to be of prime importance for the origin of language, and that is the fact that we have a more-or-less monogamous system while living many pairs together in larger social groups. This is unique among apes, if not

---

9 More subtle forms of paternal care have, however, recently been observed among baboons (Sherman & Neff, 2003), and cannot be excluded among apes.

10 Such changes are unremarkable. Mating systems are not evolutionarily stable, and may differ even in very closely related species, like the baboons studied by Bergman & Beehner (2004).
in the rest of the animal world — lots of birds, for example, form monogamous pairs within larger groups — and gibbon pairs do live within earshot of each other, even if they do not really form groups. Deacon connects this aspect of the mating system with paternal provisioning and male cooperative hunting, arguing for the necessity of an efficient social communication system to prevent cheating.

Other novelties in the human mating system include concealed ovulation and the fact that females are sexually receptive and active also when fertilization is not possible (Rodriguez-Girones & Enquist, 2001), unlike e.g., chimps, where a female clearly advertises her fertile periods and neither sex cares about sex when the female isn’t fertile. This adds a layer of tactical complexity for humans, but hardly enough to make our system obviously more complex than those of the other apes.

But even if language didn’t evolve because it is directly useful in the mating game, sexual selection has the odd property that features can be selected for even though they are disadvantageous by any objective criterion, if they happen to attract the opposite sex. The tail of a peacock male is of absolutely no practical use, and is actually detrimental to his survival — but since peacock females prefer to mate with males with large tails, a large-tailed male will have higher reproductive success, despite dying young himself when his large tail makes him too slow to escape a predator. Why a certain feature attracts females may well start out completely at random — if both the feature in males and the preference in females are inheritable, runaway sexual selection can result.

It is also common for mate preferences to display a bandwagon effect — a mate whom others have found attractive is commonly regarded as more desirable, regardless of objective characteristics. This has been shown in experiments, mostly with birds, e.g. White & Galef (2000), where females are given the choice of several males, and regularly choose the one they have seen others choose. This tactic makes evolutionary sense, as a male whom others find attractive will presumably sire sons whom others find attractive, giving you many grandchildren, but can easily lead to accidental runaway selection of features that aren’t superior in any sense but simply belonged to a male who got lucky.

Quite a few features in humans may well have a similar origin in runaway sexual selection, notably the ones that we find attractive in the opposite sex. For example, there is no obvious functional reason why human females should have breasts several times larger than the breasts of chimp females — chimps manage to produce just as much milk. The evolution of large breasts in humans may instead be adequately explained by the human (but not chimp) male reaction to big breasts. The male attraction to breasts may be interpreted as males looking for

---

11 'Groupies’ may be interpreted as human examples of the same phenomenon.
12 In the interest of gender equality, it may be noted that the male penis has been enlarged during the course of human evolution about as much as the female breast. It would be imprudent to speculate about possible interpretations — though Gallup et al. (2003) present some empirical results.
signs of health and fertility in females (Manning et al., 1997), but may as well simply be a case of accidental sexual selection. Darwin’s (1871) notion of selection for a sexy voice, mentioned above, may be considered in this context.

It is conceivable that language was sexually selected in this random fashion (Okanoya, 2002), but the obvious utility of language for various purposes, both sexual and non-sexual, points towards an adaptive explanation instead — sexually selected features are commonly either useless or harmful.

10.4 Child care and teaching

Primate babies in general mature slowly, and need parental care for an extended period, often several years. During these years, the juvenile primates not only grow up physically, but also learn about their environment. Behaviors that have to be learnt include practical issues like food gathering, but the main thrust is towards learning to function in their social environment (Joffe, 1997).

This extended childhood period has been carried to an extreme in *Homo sapiens*, as reviewed by Leigh (2001), with humans taking fifteen years or more to reach physical maturity, and sometimes even longer to become socially adult. At the same time, human infants are born immature and helpless compared to other primate babies, and require intensive parental care during the first years, severely restricting the caregiver’s activities.

The gestation time in humans is normal for a mammal of our size, following closely the multivariate allometric relation calculated by Sacher & Staffeldt (1974). But the factor most affecting gestation time is brain size at birth — Sacher & Staffeldt (1974) even argue that brain growth rate is the principal limiting factor — and, as noted on page 92, human brains continue growing rapidly after birth as well, unlike most other mammals. It is commonly argued that we are systematically born premature, but comparative data does not support this (Deacon, priv. comm.). The only way to get our gestation time to look unusual is to plot it against adult brain weight — but this is hardly relevant, as adult brain size is largely the product of postnatal growth in humans.

In other words, human babies are much like other ape babies at birth — but unlike other ape babies, the brain has a lot of growing left to do, which may explain its immature state at birth. In principle, we could deviate from the normal allometric relation and prolong our gestation time until our abnormally long period of brain growth is done *in utero*, but given the rate of brain (and head) growth of the human fetus, after nine months the baby’s head size reaches the diameter of the birth canal through the pelvis, and it had better get out before it’s too late.

13 With the caveat that oxygen delivery constraints may also set a limit.
Increasing the diameter of the birth canal is not an option, since a wider pelvis would lead to inefficient bipedal running (Byrne, 2000).

An oddity in human development is the slow growth during childhood, even slower than the prolonged juvenile period would warrant, followed by a rapid adolescent growth spurt. This pattern, with a long childhood and then a sudden transition from child to adult, may be unique to humans (Bogin, 1997; Bogin, 1999).

There is little and conflicting evidence to indicate when the prolongation of childhood took place during human evolution. We have no reason to believe that australopithecines differed from other apes in this regard, but early Homo may be different. The KNM-WT 15000 skeleton of a pre-teen boy may be a key fossil here. Its interpretation is difficult, but Clegg & Aiello (1999) and Smith (2004) find its growth pattern consistent with the modern human range of variation, though faster growth remains a tenable hypothesis as well, and Zihlman et al. (2004) find a close match with chimpanzee growth rates instead. Reaney (2001) and Dean et al. (2001) report that indications of prolonged growth are first found in Neanderthals, and are absent in Homo erectus and earlier fossils. This is supported by the erectus baby found at Mojokerto, Indonesia, which had nearly an adult-sized brain already at a very tender age (Coqueugniot et al., 2004), unlike the typical human pattern, discussed on page 92, where the fetal period of rapid growth of the brain, which stops soon after birth in most primates, is prolonged for several years into early childhood.

The presence of a human growth pattern in Neanderthals would indicate that the common ancestor of Homo sapiens and Neanderthals most likely also had prolonged growth, something which is consistent with Bermudez de Castro (2002), who reports that Homo antecessor had the modern human pattern of growth, 800,000 years ago, but this is apparently contradicted by Ramirez Rozzi & Bermudez de Castro (2004), a more recent paper sharing one author with Bermudez de Castro (2002), according to which both H antecessor and Neanderthals grew more rapidly than H sapiens. There is obviously little consensus on this issue.

The extended juvenile period in humans, as in other primates, is largely spent on getting an education. But human societies, even hunter-gatherer tribes, are much more complex than ape societies (see Section 10.5 below), and so presumably require more time to learn. Active teaching of children, by parents and others, is ubiquitous among humans, across all human cultures studied (Tomasello, 1999b). Among apes, there are many anecdotal reports of deliberate teaching (King, 1996; Savage-Rumbaugh & Lewin, 1994; Byrne, 2000), but no solid experimental evidence (Tomasello, 2000c) and only one systematic study in the wild that I am aware of (Boesch, 1991, cited in Boesch (2003)). According to Premack (2004), teaching “is strictly human” (p. 318).

Teaching is a process that would obviously benefit from having a language, and King (1996) proposes this as a driving force behind language evolution. This is a

---

V. Very likely a Neanderthal ancestor, and claimed to be a modern-human ancestor as well.
somewhat plausible suggestion, but it should be seen in the wider context of our social complexity, as discussed below. Furthermore, since there is little evidence that other apes engage in teaching, we should take care to avoid the teleology pitfall discussed on page 194 above — it cannot be postulated that language evolved so that we could start teaching our kids. Instead, a selection scenario must postulate that teaching became important first, creating a selection pressure for better communication which eventually lead to language.

10.5 Social relations in groups and tribes

Humans are social animals, as are chimps and gorillas and most other primates, living in groups with complex social relationships; O’Neil (2001) provides a brief overview of primate social structures, and Kappeler & van Schaik (2002) discuss their evolution. Krause & Ruxton (2002) provide a general overview of the biology of group life.

But humans differ from the other apes in that human social groups are much larger and more complex than chimp or gorilla groups, with correspondingly heavier demands on our ability to handle social relationships. Whiten (2000) also argues that “we [humans] are more deeply social than any other species on earth in our cognitive makeup.” (p. 477, emphasis in original). These increased social demands are very likely the main cause of our increased brain size and intelligence, according to the hypothesis of ‘Machiavellian intelligence’ (see p 97).

Group size in animals is affected by many different factors:

Predation

On one hand, there is safety in numbers, both for mutual defense against predators, and simply because when a predator attacks a large group, the probability of any given animal being eaten is small. On the other hand, large groups are conspicuous and may well attract more predators than a small discreet group. The optimal group size depends on predator characteristics. In the case of hominoids, the group defense argument appears most pertinent, favoring large groups (Aiello, 1998; Dunbar, 1996; Lewin, 1993). Some data exist indicating that primates subject to heavy predator pressure live in larger and more cohesive groups on average (Heesy, 2000; Domínguez Rodrigo, 2002), and the same group may behave more or less cohesively depending on local risk level (Bergman & Beehner, 2004). Stanford (2002b) presents data indicating that both attack probability and defensive success increases with group size, as expected, but that the net effect appears to

---

15 A single example of a non-social human tribe, the Ik, is reported in the anthropological literature (Turnbull, 1978). However, the Ik are described as being able to communicate with visiting anthropologists, so presumably they had language, as well as some notion of human contact, so one may wonder how complete their supposed non-sociality was.
be that groups subject to heavy predation are smaller, contra Heesy (2000). Data from Zuberbühler & Jenny (2002) further complicate the issue, as they conclude from their study of leopard hunting that predation rate is positively correlated with group size, body size, and number of defenders. It is worth pondering cause and effect here, as the direction is far from obvious — does a monkey species form large groups as a defense against being the leopards’ favorite dish, or do their large groups attract a disproportionate number of leopards?

**Inter-group competition**

Competition between groups may be an important factor affecting optimal group size (Isbell & Young, 1996; Zemel & Lubin, 1995). In many social animals, including both humans (Malmberg, 1980) and some other apes (Williams et al., 2004), groups maintain territories and defend them against other groups of conspecifics. Large groups have obvious advantages in such conflicts, which may well translate into more territory per member in larger groups, again favoring large groups for our ancestors; cf. Williams et al. (2004), who present data on the relation between group size, territory size, and food availability, among chimpanzees.

**Intra-group competition for resources**

In a group, the members are close together during feeding, and are likely to be each other’s worst competitors for food and other resources, decreasing food availability particularly for low-status members of large groups, as well as requiring increased traveling for the whole group in order to gather enough food for all (Koenig, 2002). At an extreme, a very large group may strain the carrying capacity of the local environment (Caporael, 1996). On the other hand, if food occurs in rare but rich patches (e.g. fruit trees in the jungle, or large carcasses to scavenge on the savannah), the larger number of food seekers in a large group increases the chance of finding a food patch in which everybody can feed well (Zemel & Lubin, 1995; Lewin, 1993), and the decreased predation risk in a group can increase the time spent actually looking for food rather than looking for enemies.

Beauchamp (1998) reviews the empirical data available on the relation between group size and food intake in birds, with the conclusion that large groups are generally favored, but that it depends on the type of food as well as other factors. Chimpanzee group size (which is quite variable in the wild) is to a considerable extent affected by local food availability (Matsumoto-Oda & Hosaka, 1998).

**Mating opportunities**

In a large group, more potential mates of the opposite sex will be available — but more competitors of your own sex as well. The net effect will be strongly dependent on your status in the group. Given the generally larger variance in reproductive success for males than for females, particularly in non-monogamous species,
the effects of competition for mates can be expected to be more pronounced for males than for females. It is indeed not rare among mammals for groups to contain more females than males, with the balance of the males living alone. Whether small or large groups were advantageous for the sex life of our ancestors depends largely on the outcome of the intra-group politics discussed in the next point below.

Another mating-related aspect of group size concerns defense against sexual predation. In some primate species, females form groups that may be motivated by defense either against male sexual coercion (rape) or infanticide (Treves, 1998).

More generally, it may be worth remembering that males and females can have very different motives for joining a group, and different optimal group size. According to Kappeler (1997), a common pattern is for females to form groups based on ecological factors, and males to join groups in order to gain access to females.

**Intra-group aggression and politics**

Given the fierce competition for food and mates in large groups, conflicts between group members are likely to become more common the larger the group — but open aggression is costly, both in time and energy spent, and in risk for injuries and death. All parties would gain if all conflicts could be resolved peacefully. However, if all other group members retreat from aggression, a single aggressor can invade and dominate the group, to the detriment of everybody else. Even though all individuals would gain by peace, in a peaceful group each individual can gain by aggression — this paradox is related to the well-known ‘Prisoner’s dilemma’ game (Koeslag, 1997). This dilemma has been the subject of considerable research for some years, seeking the roots of the cooperative and altruistic behavior that we do observe in many groups, and seeking to explain the comparative rarity of open aggression, despite its apparent benefits to the aggressor.

The solution to the dilemma lies in the evolution of social intelligence, starting with the recognition of other individuals, remembering past interactions with each individual, and differentiation of behavior towards other individuals depending on their past behavior. But in a large group, this taxes the brain power of most animals, limiting either the group size or the complexity of the social system (Dunbar, 1993).

A static status hierarchy is a common solution with limited cognitive demands — but the complexity rises fast if the status hierarchy isn’t static, and if status relations aren’t transitive. And if you’re not at the top, a static hierarchy isn’t to your advantage, so it pays to keep track of the actual power of all those around you, and figure out the right moment for a bid for higher status — but this raises the cognitive stakes again.

What really leads to a cognitive arms race is when the simple hierarchy based on one-on-one dominance relationships is abandoned, and the possibility of status based on friendships and coalitions and negotiations is considered. In this case, you have to keep track of not only your own position in the hierarchy, but also
the social networks of everybody in the group — who is friendly with whom, who might consider an alliance with whom, and so on. You must be able to solve social equations like what happens if you attack individual A, when B, C, and D are also present — B might side with you, since you did him a good turn last week, but that might cause C and D to gang up on him, in which case...

The possibilities are endless, and grow exponentially with increasing group size. In the simplest case, social networks are based on blood relationships, so that kin form permanent alliances. This is fairly common among e.g., baboons (Strum, 1989) and various other mammals. Kin networks are permanent, making them easy to keep track of, limiting the cognitive demands, but also limiting the political possibilities. DNA testing on chimpanzees has shown that their politics is not limited in this way — coalitions of non-relatives are common (Goldberg & Wrangham, 1997; Mitani et al., 2000). They are not totally unknown among baboons either (Byrne, 2000), though less common.

Here we have the basis for the hypothesis of ‘Machiavellian intelligence’ mentioned earlier. Considering that both humans and chimps spend a major fraction of their time and cognitive abilities pondering intra-group relations, and considering that humans do it in much larger and more complex groups than chimpanzees, the hypothesis appears highly plausible — we hardly need such an enormous brain to outwit Nature, but we might well need it to outwit our fellow humans (Pinker, 1994), in order to win the game that we might call ‘social chess’ (Andrews, 2001).

In the context of our early ancestors, it would appear that larger groups were desirable for several reasons, notably predator and territorial defense connected with the changed diet and more open habitat, that our ancestors switched to at about the same time as our brains started growing.

But a system of highly complex intra-group political machinations had evolved already before our common ancestor with the chimpanzees, as both we and the chimps are living with such systems (de Waal, 1998), which would limit the maximum group size to something an early ape could handle mentally. A typical chimpanzee community consists of a few dozen individuals, with the largest ever recorded having 117 members, including infants (Mitani & Watts, 1999), whereas already the average human hunter-gatherer community is larger than the chimpanzee maximum, and as we all know there are humans who have managed to rise to the political top in communities that are orders of magnitude larger. Chimpanzees can apparently keep track of and be competent participants in the politics in a group with a dozens of members — but they would likely be lost and confused if they formed a typical human-sized tribe (Dunbar, 1996).

If other selective pressures forced proto-humans into larger groups than they were used to, and they already had chimp-style politics, there would be very strong selective pressure towards improved socio-political capacity. As Dunbar (2003a) notes, the direction of causation has to be from group size to cognition, not the other way round.
highly political creatures was noted already by Aristotle: “Now, that man is more of a political animal than bees or any other gregarious animal is evident.” (Aristotle: Politics 1.2, quoted in Everson (1994a, p. 7))

This political argument strongly suggests a socio-political hypothesis for the evolution of language — because politics, also in the ape version, is very much about communication. The Aristotle quote above continues: “Nature, as we often say, makes nothing in vain, and man is the only animal who has the gift of speech.”

Agreements, formation of alliances, trading of favors, coalition building, all aspects of politics place heavy demands on the communication between the parties involved. A better communicator is a better politician, and a better politician can gain higher status in the group, and reap the associated reproductive benefits for his communicator genes and memes — and spread them in the tribe. Dessalles (2000; 2003a; 2003b; 2004) discusses this further, also invoking a direct role for speech in gaining status. In the scenario of Dessalles, “talking individuals engage in a competitive display to advertise their informational abilities.” (2004, p. 1). People are valued as political coalition partners in direct proportion to the amount of valuable, relevant, and novel information that they can present.

Bickerton (2003) objects to the political hypothesis, with the counterargument that (proto-)language must have started with a tiny vocabulary, insufficient to do any social manipulation. But chimpanzees manage a fair bit of social manipulation with even less language, which removes most of the force from his argument.

There is little evidence in the archeological record concerning prehistoric group size, not until very recent times in Homo sapiens (Wynn, 2002) — but we already know from modern data that Homo sapiens lives in very large groups — so it is very difficult to tell when the transition from chimp-sized to human-sized groups took place. The hypothesis of Machiavellian intelligence would predict that group size growth caused brain size growth, in which case it can be inferred from fossil skulls that group size increased gradually during the past 2 million years, with a couple of growth spurts (Dunbar, 2003a); cf. Figure 5.5. But independent confirmation is lacking, and Jeffares (2002) cautions against drawing far-reaching behavioral conclusions from brain-size data. Inferences about cognition are even more problematic — drawing conclusions like those of Dunbar (2003a), concerning levels of intentionality in fossil humans, is imprudent.

Dunbar (1993; 1996) adds another twist to the socio-political hypothesis. Social and political relations are based on friendship, and friendship requires maintenance — in order to become and remain friends with people, you have to spend time bonding with them. Among monkeys and apes, this bonding largely takes the form of mutual grooming, reviewed in Schino (2001) — but grooming takes time (Byrne, 2000), and apes commonly spend a substantial fraction of their time grooming each other, time during which they cannot pursue either food or sex. The time needed for grooming is essentially constant per individual groomed, so the total time an ape spends grooming would grow linearly with group size. Al-
ready for the small groups of apes, grooming is eating up significant amounts of time — if that time is extrapolated to common human group sizes, the time required is around twelve hours per day, most of our waking hours, leaving precious little time for other activities. But not spending the time required for social maintenance leads to instability in the group. It follows that grooming is untenable as social lubricant in complex groups much larger than those of chimpanzees.

In order to maintain friendships in very large groups, a more time-effective method is needed. Dunbar (1993; 1996) proposes that language evolved for this purpose, as a tool for ‘verbal grooming’. He notes that a large fraction of all human speech consists of friendly gossip, which serves both to bond those who gossip together, and to exchange valuable social information about others. In other words, gossip fulfills the function of grooming, with an information bonus added. Furthermore, gossip is much more time-effective than grooming, both because other chores can be done in parallel, and because it doesn’t have to be performed one-on-one like grooming — a group of ten people can gossip together simultaneously, and get bonded in a small fraction of the time it would take them all to groom each other in all pair combinations. As we all know from experience, it is perfectly possible for a human to keep up with the gossip in a fair-sized group, without having to spend twelve hours a day at it. One might, however, wonder why a grooming tool would need to be such an efficient information transfer system — as Bickerton (2003) puts it, “a similar result could have been achieved simply by using pleasant but meaningless noises.” (p. 79). Additional selective pressures, apart from grooming, would appear to be needed to make language anything like what it is today.

There is a fair correlation, as expected under Dunbar’s hypothesis or for that matter under the general socio-political hypothesis, between neocortex size and both group size (Dunbar, 1996) and social network size within groups (Kudo & Dunbar, 2001).

Further support for the social hypothesis comes from the various experiments attempting to teach language to non-humans, reviewed in Chapter 7 above. The only species having any kind of success in these experiments, apes, dolphins and parrots, are highly social with complex group relations. Additionally, teaching experiments are significantly more successful if they take place in a social setting (Pepperberg, 2001; Savage-Rumbaugh & Lewin, 1994).

Gärdenfors (2002) and Tomasello et al. (in press) argue that a key event in the evolution of human cognition was the emergence of skills and motivations for collaborating more deeply with our peers, putting our social intelligence to cooperative use, “creating common visions” (Gärdenfors, 2002, p. 2),17 as opposed to the competitive contexts in which it originally evolved. Language is regarded by Tomasello et al. (in press) as derivative of these cooperative social and cognitive skills.

---

17My translation from the Swedish “skapa gemensamma visioner”.
Aiello (1998) extends the socio-political hypothesis in another direction, arguing not only that social complexity drove the need for language, but also that social intelligence provided the cognitive structure for language:

There are similarities in reasoning processes or procedures between primate social intelligence and the computational basis of language processing including both the semantic aspects of language and syntax... (Aiello, 1998, p. 29).

It is argued that we store social events as scripts that provide the procedural basis for syntax. We shall return to this possibility in Section 11.4.

All in all, the basic socio-political hypothesis appears highly plausible as a basis for the selective advantage of language (Dessalles, 2000). The verbal grooming hypothesis of Dunbar (1996) is worth considering as well.

10.6 Miscellaneous ideas

In this category, hypotheses are included that are far from the mainstream, and have few supporters. The list below could have been much longer than it is, but obvious crackpots have been excluded.

Children at play

Playing is proposed as the main environment of early language evolution by Foster (1991), with language not used for communicative purposes, but instead in mimicry and random repetitive playing. Foster, however, clearly illustrates the point of my reversed quote from page 3: *we cannot leave the discussion of language evolution to those linguists who have yet to understand the concerns of modern evolutionary theory*; she is one of those linguists, displaying serious misunderstandings of (and a negative attitude towards) evolutionary biology.

Knight (2000) and Ragir (2001) return to the idea of childhood play as an expectation for language, particularly its creative and combinatorial aspects, and support it with a much better understanding of evolution than Foster. In Knight’s version, it appears fairly plausible that play contributed at some level to the origin of language.

Music

Music and singing is the theme of Vaneechoutte & Skoyles (1998), according to whom “[s]ong ... underlies both the evolutionary origin of human language and its development during early childhood.” (p. 2), with the early stages analogous to birdsong, and selected for similar reasons. On a related note, Dunbar (2003a) suggests that vocal chorusing may have had a group bonding function in *Homo erectus*.
The notion that music had a role in the origin of language has a distinguished pedigree, with both Rousseau (1781), Darwin (1871; 1872) and Jespersen (1922) entertaining related ideas.

Some role for singing and rhythmic music in early language evolution cannot be excluded, but the dismissal of language evolving as an adaptive communication system is not plausible — the role of music, if any, is very unlikely to be as dominant as Vaneechoutte & Skoyles (1998) propose. Furthermore, music appears to have its own neural circuits, distinct from those used for language — people exist who are completely amusical but still have normal language skills (Balter, 2001e) — whereas considerable overlap would be expected if music lay at the roots of language, though Peretz & Hyde (2003) argue that amusia is caused not by defects in any music-specific brain module, but by defective pitch perception.

Verhaegen (1998) also links music to language origins, but in the context of the Aquatic Ape theory (Morgan, 1982), which does not add credibility (Langdon, 1997).

Storytelling

The central function of (proto-)language in the stories told by McNeil (1996) and Heeschen (2001) is storytelling. This is indeed something that is ubiquitous and important in human societies and as far as we can tell totally absent among other species, and language is obviously vital for our narrative capacity. But storytelling without language is difficult, so it is not immediately obvious how to avoid backwards causation in a scenario where storytelling provides the selective pressure for the origin of language. Possibly mimesis, discussed on page 175, could be a scaffold here?

Art

Art has been invoked in two opposite causative roles here. Davidson & Noble (1993, cited in Savage-Rumbaugh & Lewin (1994)) cast it in a positive causative role, driving language evolution. Even more odd, Bowles (1998) presents an argument that inverts both the causal direction and the connection between language and art, asking whether it was the absence of language that drove the evolution of art. The empirical basis for this hypothesis is, however, modest, apparently limited to a single autistic child (presumably lacking language), who draws images that resemble cave art.

Art in an indicative rather than causative role was discussed on page 168.
10.7 Why us and not the other apes?

Several of the ideas discussed above are plausible driving forces behind language evolution. But general plausibility is not sufficient — as has been pointed out repeatedly, a serious theory of language evolution must also account for the apparent absence of language in the other apes. Many of the factors that might have driven human language evolution, could equally plausibly have driven chimpanzee language evolution — except that in reality there must have been insufficient selective pressure towards language in chimpanzees, since they didn’t evolve human-level language.

The split between the human and chimp lineages is commonly attributed to the progressively drier climate in Africa in the relevant time frame, causing deforestation and the spread of savannah biotopes (Lewin, 1993; Isbell & Young, 1996). The general idea is that chimp and gorilla ancestors stayed in the remaining forests, retaining something close to the ancestral lifestyle, whereas human ancestors got stuck on the newly formed savannah, or in isolated patches of forest. The oldest hominid fossils are associated with woodland fauna (Haile-Selassie, 2001; Leakey et al., 2001; WoldeGabriel et al., 2001; Schoeninger et al., 2003), but later ones lived in more open terrain (Sponheimer & Lee-Thorp, 1999; Vogel, 1999b).

This new open-terrain lifestyle is the major ecological difference between our ancestors and chimp ancestors, and it appears highly likely that whatever factors drove language evolution in humans but not in chimps, are associated with this lifestyle. Larger group size, likely driven both by the more intense predation pressure on the savannah, and by the patchier food distribution, is an obvious factor to consider. It is commonly found, when comparing forest-dwelling arboreal and savannah-dwelling terrestrial relatives, that the terrestrials live in larger groups (Lewin, 1993). Given a larger group size, the socio-political complications that follow would generate a strong selective pressure towards improved social cognition and communication, as discussed above. The sexual aspects, increased teaching needs, and the dietary shift towards more hunting and/or scavenging, would all reinforce this trend, but it appears likely that the social pressures dominated. This is a reasonably plausible scenario that could lead to the evolution of language in the human lineage but not in chimps.

The ecological niche of Homo sapiens differs from chimps not only in its physical environment, but also in the way in which we exploit the environment. Our lifestyle is fundamentally based on cognition and information, to such an extent we have been called ‘informavores’ (George Miller, cited by Pinker (2003)) and our niche ‘the cognitive niche’ (Tooby & DeVore, 1987). Language may in this perspective be seen as part of a larger package of information-handling adaptations to our information-based niche (Pinker, 2003). But language as a means for transferring information still leaves open the questions ‘why inform each other?’ (cf. Section 6.3) and ‘information about what?’, leading us back to the various alter-
natives discussed above — technological information, environmental information, social information, or what?

10.8 Summary

Many factors have been proposed as the main selective pressure behind the origin of language. The major ones are:

- **Hunting.**
  - **Pro:**
    - Economically important for our ancestors, at least within the genus *Homo*.
    - New ecological niche for *Homo* with increased hunting.
  - **Con:**
    - Other apes hunt, too — why don’t they talk?
    - Other social carnivores do just fine without language.
  - Scavenging and gathering should also be considered in this context, as both may have been important in our past. It is not clear why language would be more useful for hunters than for gatherers.

- **Tool making.**
  - **Pro:**
    - Large-scale making of complex and durable tools major factor in our evolution.
    - Tool complexity increased in several stages in the archeological record. Both early and late theories of language evolution can find a matching tool-stage.
    - Human-style tool making is a structured sequential activity — much like language. Possible coevolution or exaptation?
  - **Con:**
    - Other apes make tools, too — why don’t they talk?
    - Why would a basically solitary activity like flint knapping drive the evolution of improved social communication?
  - Tool making does not appear plausible as a major factor behind language. But structured sequences may be a useful language exaptation.

- **Sex.**
  - **Pro:**
    - Language use important for mating success in modern *Homo sapiens*.
    - Human mating system more complex than that of many other animals — greater need for communication.
    - Runaway sexual selection can drive evolution of all kinds of weird features.
  - **Con:**
    - Other apes have sex, too — why don’t they talk?
    - Mating system of chimps at least as complex as ours.
    - Language too obviously useful to be the random result of runaway sexual selection.
  - Sex may be involved with the origin of language — but more likely indirectly, in connection with the general sociopolitical point below.
Child care and teaching.

- Pro:
  - Language obviously useful for teaching.
  - Explicit teaching rare or nonexistent in other primates.
  - Humans have an extended childhood compared with other primates.

- Con:
  - Apes don’t teach their kids. Teaching must be common and important before it can drive language evolution.

- May well have contributed to the evolution of language to some degree.

Social relations.

- Pro:
  - Social relations is the primary topic of language use today.
  - Status and relations in the group is very important for individual fitness.
  - Human social relations have a complex structure with a network of friendships, alliances, political deal-making, and so on.
  - Human social groups are very large, much larger than those of other primates with complex group politics.
  - Better social communication obviously valuable in handling social relations — selective advantage for language users?

- Con:
  - Chimps have complex group politics, too — why don’t they talk?

- The communication needs for social relations and group politics appear plausible as the main driving force behind the evolution of language.

Further reading

CHAPTER 11

PROTO-LANGUAGE

The preponderance of the evidence reviewed so far points towards language as an adaptation, shaped by Darwinian selection in a process of interacting biological and cultural-memetic evolution. But *natura non facit saltum* — nature does not make sudden jumps, specifically not in evolution, which implies that language evolved in a gradual process, through several intermediate steps. The issue in this section is what those intermediate stages may have looked like.

There are several aspects of language for which intermediate stages are required. We need both proto-syntax, proto-semantics, proto-speech, and possibly proto-gestures (if gestures were important in language emergence). Furthermore, the interdependence between these components needs to be clarified, as a basis for discussing their possible order of appearance, as well as their viability at every stage as a communication system. Throughout the entire process, a Darwinian perspective needs to be kept in mind — is every step actually an improvement in the Darwinian sense, conferring a selective advantage on the individual taking that step?

This chapter is *not* intended to provide a full-fledged original scenario of language evolution. As stated already in the preface of this book, my purpose is to review the evidence and constrain hypotheses, not to ‘sell’ any theory of my own. That aim remains in force in this chapter as well, limiting its scope to a review of possible language evolution avenues, a sketch of some scenarios, and how they are constrained by the evidence presented in previous chapters.

11.1 Proto-speech

Only the mechanics of speech production and perception will be covered here, not its semantic and syntactic contents, which will be treated in Section 11.3 and 11.4 below. The evolution of the production and perception of speech sounds has already been dealt with in some detail in Chapter 5, the main results of which are reviewed here.
Speech perception is the easiest part, since chimpanzee perception is already adequate, as demonstrated by their comprehension of human speech in blind tests (Savage-Rumbaugh et al., 1993), reviewed on page 134. No proto-perception stage is needed.

Speech production is more difficult, since the vocal apparatus of our cousins (and presumably our ancestors) is inadequate to produce the full range of speech sounds. Furthermore, the neural control of vocalizations in apes is not well suited for speech production. With great efforts, chimps can produce sounds that vaguely resemble the vowel patterns of context-appropriate words — one gets the impression both from the anecdotal description in Savage-Rumbaugh & Lewin (1994) and from the more controlled study described in Taglialatela et al. (2003) that their chimpanzee student Kanzi (see page 132) is trying very hard but not very successfully to reproduce human words, in communicatively appropriate contexts. As already mentioned, Kanzi has demonstrated that he understands English well enough (Savage-Rumbaugh et al., 1993), but his production is severely handicapped.

Still, the sound repertoire of chimps is far from negligible — if only they had adequate neural control, so that they could combine freely the sounds that they do produce, they’d have a set of phonemes adequate for a simple language. Likewise, humans with damaged speech organs can often compensate enough to produce comprehensible speech (Wind, 1989), demonstrating that speech is possible with less than the full human equipment.

MacNeilage (1994; 1998) has developed a hypothesis for the origin of the first speech sounds, which is concerned with the evolution of neural control of jaw movements. He combines evidence from ontogeny (infant babbling) with evidence from universal sound patterns in human languages, to propose a set of simple syllables as the core of the ursprache. The exact set has evolved over time, from MacNeilage (1994) to MacNeilage & Davis (2000). In the latter, it consists of three CV syllables: coronal+front (e.g. te-te-te), labial+central (ba-ba-ba), dorsal+back (go-go-go), and one CVC combination (labial-vowel-coronal).

As is easily verified, the three CV syllables correspond to the same simple up-down jaw motions with the tongue fixed in three different places.\(^1\) These are indeed arguably the articulatorily simplest possible syllables. These syllables are also very common, and occur in both infant babbling and the world’s languages at frequencies well above those expected from the frequency of their component sounds, and also form a large fraction of Ruhlen’s (1994)\(^2\) proposed proto-world

\(^{1}\)Modern humans (or at least I) normally produce repeated coronal-front and dorsal-back syllables by moving the tongue rather than the jaws — but holding the tongue fixed and moving the jaw works perfectly fine as well, producing recognizably the same phonemes.

\(^{2}\)But note that Ruhlen’s proto-world reconstruction has not been favorably received among most linguists (Hurford, 2003a), and in any case the timing is wrong — MacNeilage’s syllables have to be hundreds of thousands of years older than Ruhlen’s proto-world, which must have been spoken by Homo sapiens with a fully modern phonetic repertoire.
vocabulary (MacNeilage & Davis, 2000). The findings of MacNeilage & Davis (2000) are further discussed in Locke (2000).

The neural control of jaw motions was available as an exaptation, well before the evolution of language, since we need that control for chewing and biting and have motor programs for it, and chimps use it also for lip-smacking sounds, so the step from a biting or lip-smacking motion to babbling does not appear large, involving mainly a steady airflow from the lungs (cf. the breathing control of MacLarnon & Hewitt (1999), discussed on page 82) and the maintenance of a steady tongue position with respect to the lower jaw.

And once speech got started, with these syllables or otherwise, it would be a matter of straightforward Darwinian selection to improve the vocal tract and neural control, in order to produce a richer variety of crisp and easily-understood phonemes. A rich repertoire of sounds is particularly valuable at this early stage, before the invention of combinatorial phonology, as each concept needs a sound of its own.

Given a vocal tract with reasonable articulatory capabilities, the sound systems of human speech can emerge through evolutionary self-organizing processes across generations, as shown for vowels by Bart de Boer (1999; 2000; 2001; in press) in a series of computer simulations. Demolin & Soquet (1999) discuss further the role of self-organization in the evolutionary emergence of sound systems.

As for combinatorial phonology, the stringing together of speech sounds into longer words, Nowak and associates (Nowak & Krakauer, 1999; Plotkin & Nowak, 2000) have shown that this is advantageous as soon as the number of concepts exceeds a certain threshold. According to Jackendoff (1999b; 2002), the switch from a system where each sound carries meaning, to a combinatorial system where meaning is assigned to combinations of individual sounds that become meaningless in themselves, represents a major cognitive step, enabling an explosive and open-ended growth in the number of symbols.

The combining of sound units into longer strings is observed in various kinds of animal communication as well, notably the songs of birds and whales, and there are hints of compositeness in some ape vocalizations as well (Ujhelyi, 1996; Ujhelyi & Buk, 2001), which may be a usable exaptation for proto-speech. The existence of composite vocalizations in three distantly related animal groups implies multiple independent evolutionary origins, from which follows that this stringing together of sound units is not evolutionarily difficult to acquire. The sequencing ability of Calvin (1993), discussed in Section 5.1.3, may be a useful exaptation here. But there is no evidence from other species of any emergent meaning in the sound combinations, which is the whole point with human combinatorial phonology.

3Robin Allott in his ‘Motor theory of language’ (Allott, 1991) attempts to build a more general model of language origins based on pre-existing motor programs, not just for speech but for all aspects of language. For speech and gestures, this is somewhat plausible, but hardly for the more abstract aspects of language, like syntax and semantics. Allott (2003) connects his theory with the discovery of ‘mirror neurons’; cf. Section 9.6.2.
Studdert-Kennedy (2000) and Studdert-Kennedy & Goldstein (2003) discuss how to proceed from the syllabic babbling of MacNeilage & Davis (2000) to combinatorial phonology, first breaking down the syllables into articulatory gestures. The gestures are then freely recombined into phonetic segments. Vocal imitation also plays a key role in scenario of Studdert-Kennedy (2000), further elaborated in Studdert-Kennedy (2002) with a discussion of the possible role of the mirror neuron system discussed in Section 9.6.2. Combining the hypotheses of MacNeilage & Davis (2000) and Studdert-Kennedy (2000) gives what appears to be one somewhat plausible model for the origin of proto-speech.

11.2 Proto-gestures

As in the preceding section, only the mechanics of sign production and perception will be covered here, not its semantic and syntactic contents. This results in a very short section here, since the issue is almost trivial — it is beyond reasonable doubt that our common ancestors with the other apes, or even monkeys, had all the dexterity needed for sign language production, as well as the visual processing capacity needed to perceive signs.

As discussed on page 123, chimpanzees and gorillas use gestures for communication (Leavens et al., 1996; Leavens & Hopkins, 1998; Savage-Rumbaugh & Lewin, 1994; Tanner & Byrne, 1999), also in the wild (Vea & Sabater-Pi, 1998; Jucquois, 1991). The gestures appear, at least to human eyes, to be largely iconic, unlike modern spoken language, but this is not unreasonable for a proto-sign language, as iconic features are present also in a significant fraction of modern sign-language signs (Pietrandrea, 2002). The evolution of sign language from the gestures of apes, possibly along the path discussed by Stokoe (1986), appears for several reasons much more plausible than the evolution of speech from animal vocalizations that Darwin referred to in the quote on page 173. The transition from iconic non-language gestures to a sign language has been observed in the spontaneous creation of sign creoles in deaf children without contact with sign-language speakers (Goldin-Meadow & Mylander, 1998; Saffran et al., 2001), an ontogenetic process that may have had a phylogenetic predecessor (Wilkins & Wakefield, 1995). Of course, there is the important difference that the children already possessed whatever innate language abilities that humans are equipped with — nevertheless, the parallel is interesting, and as far as I know no similar development of spoken creole from iconic sounds has been observed.

4A somewhat infelicitous name, given the important role of gestures in a totally different sense in the context of language origins.
11.3  Proto-semantics

Semantics, relating linguistic expressions with meaning, is inextricably intertwined with the ability to use symbols and symbolic thinking and communication. The ‘symbol’ concept is central here, but is not entirely trivial to define and operationalize, and the different definitions that can be found in the literature are not quite equivalent.

Symbols and signs were discussed already in Classical times, e.g. by Aristotle (Charles, 1994), Sextus Empiricus (Gladen, 1994), and St Augustine (Kirwan, 1994), though symbols were not clearly distinguished from other kinds of signs. Kant (1781) further analyzed the not entirely trivial issue of what it means for a sign to represent something, leading to an entirely new concept of representation (Sinha, 1988).

Bühler (1934) built a model of a sign with three essential components:
- *Symptom*, the sign as an expression of an aspect of the sender.
- *Signal*, the sign as it appeals to the receiver.
- *Symbol*, the sign as it represents a referent.

Peirce (1931–35) introduced another influential three-part classification of signs, but along totally different lines. While Bühler (1934) was concerned with different semiotic components of the same sign, as it related to sender, receiver, and referent, Peirce distinguished different types of relations between expression and content that a sign could have, different ways for a sign to acquire meaning:
- *Iconic*: having a non-arbitrary relationship between sign and referent, e.g., mimetic gestures.\(^5\)
- *Indexical*: having direct contiguity with its referent, e.g., pointing gestures.
- *Symbolic*:\(^6\) having conventionalized meaning, being related to its referent only through an arbitrary social agreement.

Implicit in the ‘conventionalized’ feature is the fundamentally social nature of symbols — symbols are used in communication to establish a field of shared attention, by virtue of the joint reference of sender and receiver. But a system characterized only by joint reference is regarded by Sinha (2001) as merely proto-symbolic, and ‘conventionalized’ goes beyond joint reference in that it entails not only that both sender and receiver attend to the same referent, but also that they both know that they both know, and both know that the other knows that they know (third-level intentionality *sensu* Grice (1957)), that a certain symbol means their joint referent. The emergence of shared meaning is thus a thoroughly social process, requiring that what is shared between speakers is not just that both use the same symbol for the same referent, but that they actually share the same *instance* of

\(^{5}\)But see Sonesson (2001) for some subtle complications surrounding the concept of iconicity.

\(^{6}\)It is unfortunate that both Bühler and Peirce use the word ‘symbol’ in their theories, as it means totally different things to them. In the following I will use symbol *sensu* Peirce, unless otherwise noted.
meaning (Itkonen, 1983). In Popperian terms (Popper, 1979, discussed in Section 3.4.2), this could be interpreted as both speakers having in their heads world-2 instantiations of the same world-3 object, or in memetic terms (Section 3.4.3) them sharing the same meme, not just identical but unrelated memes. In complex cases, this may require even higher levels of intentionality (Itkonen, 1997).

Another feature distinguishing symbols, according to Deacon (1997), is the ability of a symbol-user to disconnect the symbol from the presence of its referent; cf. the ‘detached representation’ of Gärdenfors (1996). Jackendoff (1999b) emphasizes also the non-situation-specificity of human symbols. Both indexical and iconic signs are obviously connected with their referents, either through contiguity or similarity, and are thus necessarily ‘cued representations’ sensu Gärdenfors (1996), but symbols have the potential for independent use, enabling a symbol user to communicate about matters absent in time or space, or even matters purely hypothetical and counterfactual.

The emergence of full symbolization sensu Sinha (2001) requires also construal, entailing the elaboration and conceptualization of the joint references of proto-symbolic systems. Construal is a matter of connecting a symbol with other symbols in a network of internal relations, not through relations between their respective referents; cf. Saussure (1916). This relational feature is sometimes interpreted as implying that symbolicity requires syntax, that true symbols can only be used in a grammatical context — but, as Jackendoff (1999b) argues, children at the one-word stage use symbols by any reasonable definition, and adult speech also contains a number of symbols, such as expletives, that are normally used singly, not in any grammatical context.

When signs are used for communication, the cognitive work required is distributed differently between sender and recipient depending on the level in the Peircean hierarchy (Deacon, 2003b) — indexes and icons are easier to interpret, but may be cumbersome to produce, whereas symbols are easier to produce but require a lot of learning and cognitive work in order to be interpretable. But symbols allow more semiotic freedom, without which the complex communication of humans would hardly be possible. The latter point may explain the trend towards symbolization, shifting the cognitive burden from sender to recipient, that must have been part of language evolution — but in order to keep the recipient’s task possible, semiotic constraints must be imposed on language, so that symbolic reference remains grounded (Deacon, 2003b). Predication is a central concept here — cf. Ghadakpour & Dessalles (2004), discussed on page 237 below.

Sinha (in press) proposes a scenario for the evolutionary origin of symbols, based on “the epigenetic emergence and elaboration of symbolization”. Starting

---

7But see Sonesson (in press) for a critique of Deacon’s handling of the Peircean system. The intent of Deacon (1997), however, was not just to restate Peirce’s definitions, but to extend them in the light of post-Peircean discoveries in various areas (Deacon, private communication).
from non-symbolic communication between a sender and a receiver, symbolization may emerge in the following steps:

1. The receiver comes to pay attention to the sender as the source of communicative signals.
2. The sender comes to pay attention to the receiver as a recipient of communicative signals.
3. The receiver comes to pay attention to the evidential reliability of the sender’s communicative signals as a source of information, by checking what the sender is paying attention to, or doing.
4. The sender comes to pay attention to the receiver’s readiness to reliably act upon the information communicated, by paying attention to what the receiver is paying attention to, or doing. (Sinha, in press, p. 11 of preprint)

‘Meaning’, the other main component of semantics, is also a non-trivial concept to define. Zlatev (2003a) proposes a hierarchy of meaning systems, with four levels:

1. **Cue-based**: Meaning is innately assigned to predetermined cues. No learning or flexibility is involved.
2. **Association-based**: Meaning is assigned indexically to signals through associative learning of relations between environment and action.
3. **Mimetic**: Meaning is assigned to mimetic conventions, through social and imitative learning.
4. **Symbol-based**: The full human system of meaning and value, with social and cultural learning. Signals need no longer be indexical or iconic, but can be arbitrary symbols. Higher-order value systems, largely disconnected from the innate value systems of the lower levels, play a prominent role in the emergence of symbol-based meaning.

Each of Zlatev’s meaning-levels has a multi-component structure, with a level characterized by the type of signal, the value system and the type of learning involved, as well as perceptive and emotional components.

Communication systems among non-humans that may be relevant in this context were discussed in Chapter 6, particularly the alarm calls of vervet monkeys (Seyfarth et al., 1980). The monkey calls appear to have a mosaic of features, and do not fit neatly into any single level of either the Peircean or the Zlatevian hierarchy. But regardless of the label we assign to it, the system of the monkeys is far removed from the huge open multi-modal symbol-based meaning system available to humans. The openness of the human system, and our ability to acquire new symbols, and particularly our ability to establish joint reference, agreeing on the mapping between symbol and referent, is unmatched among wild animals.

The vervet-monkey system is not totally innate, but it shows no signs of being open either. Zlatev (2003a) assigns the vervet monkeys to the association-based meaning level, which would appear to be a reasonable assessment, even though the social component in their learning of the system may not be negligible.
Oliphant (1999) argues that the closed nature of animal communication systems, and the consequent difference in mapping abilities and in the abilities of juveniles to acquire the system through observational learning “is the primary factor limiting the evolution of language ability.” (p. 380). But Oliphant (1999) disregards here the learning abilities demonstrated by animals in captivity.

The two chimpanzees Sherman and Austin discussed on page 131 (Savage-Rumbaugh & Lewin, 1994; Deacon, 1997), make a much stronger case for symbolic communication in non-humans, with their clearly communicative use of a non-trivial number of arbitrary tokens. Sherman and Austin definitely achieved the joint-reference proto-symbolic level of Sinha (2001), as demonstrated e.g. by their ability to jointly assign a token to refer to a new item.

As for the ‘observational learning’ of the connection between symbol (or proto-symbol) and referent, large vocabularies have been acquired and are used communicatively by various apes. At least in the late stages their acquisition was more through observational learning than through formal training (Lyn & Savage-Rumbaugh, 2000). This indicates that at least some non-humans appear to have the ability to overcome what Oliphant (1999) regards as a major hurdle in language evolution.

Concerning the level of meaning displayed by these enculturated chimpanzees, Zlatev (2003a) assigns them to the mimetic level. Assigning them to lower levels is clearly untenable. It is, however, not clear whether higher levels of meaning can be excluded, or whether they even fit into a single slot in the level system. For one thing, the mimetic level is supposed to be characterized by iconic signals — but many of the chimps have successfully acquired non-iconic systems, notably both Sherman & Austin mentioned above, and Kanzi, the most successful of all non-human language users. While these chimps are definitely capable of handling arbitrariness, it is unclear to what extent they can manage full conventionality, with the third-level intentionality (Grice, 1957) that it entails, though there are episodes like the one described on page 131, where Sherman & Austin appear to display such intentionality.

Detached representation, the independence of symbols from the presence of their referent, has been difficult to establish in apes, though Menzel (1999) makes a fair case for it being within the reach of enculturated apes. There is no evidence that apes can build the network of internal systematic relations between symbols, that characterizes human language.

If iconicity were the main distinction between level 3 and 4, a strong case could instead be made for placing the chimps on level 4, together with humans. But there is no evidence in chimps for the higher-order meaning-value systems that, according to Zlatev (2003a), characterizes symbol-based meaning systems. Concerning learning, the distinction between level 3 and 4 is between imitative and cultural learning (the learning being social in both cases), and it is not clear whether the chimps display cultural learning. Chimps in the wild do have some signs of culture.
(de Waal, 1999; Boesch, 2003), but human-style cultural learning, the kind leading to the cumulative growth of cultural complexity (Tomasello, 1999a) appears to be lacking in the wild, and may well be absent also in enculturated chimps. The enculturated apes are left hanging partway between mimetic-based and symbol-based meaning, fulfilling some but not all of the criteria for level 4.

As stated by Zlatev (2001b), the Zlatevian level system is not intended to be more than a schematic classification, with no expectation that every species can be pigeonholed into a level, so having apes on level ‘3 1/2’ is not really a problem. With each level being composed of several separate requirements, that are not in all cases logically dependent on each other, intermediate stages like this must be possible. It must be possible to have e.g., a system that is conventional and arbitrary but not connected with internal relations, a system corresponding to the level Sinha (2001) calls proto-symbolic. This proto-symbolic level would appear to be a reasonable starting position for the evolution of fully symbolic communication.

It would appear, then, that cognitive abilities sufficient for proto-symbolic communication sensu Sinha (2001) were available in our common ancestor with the chimpanzees, as well as all the abilities required for Zlatevian level 3 and possibly some of those required for level 4 as well. This raises the questions of why the chimps have abilities that they apparently don’t use in the wild, and why chimps haven’t evolved a language-like communication system, at least at the level displayed by Kanzi and other enculturated chimps (see Chapter 7), if they already have such a central part of the abilities needed. Again, as on page 127, one may wonder whether chimp communication in the wild is richer than we have thought, possibly entailing enough aspects of symbolicity to motivate their proto-symbolic capacity? As for why they haven’t evolved full (or even proto-)language, Section 10.7 above attempted an answer in terms of different selection pressures. Another part of the answer may come from the studies of Nowak and associates (Nowak & Krakauer, 1999; Nowak et al., 2000), discussed in Section 11.1 and 11.4 respectively, who have demonstrated that the hierarchical structures of human languages are needed only when the number of concepts to be communicated exceeds certain thresholds — below those thresholds, a simple one-to-one correspondence between concepts and sounds is more economical, something which may apply to the arguably simpler world of the chimps.  

Alternatively, a more likely explanation may well be that while the chimps may have all the biological prerequisites for proto-symbolic communication, biology is certainly not the whole story in the ontogeny of higher-level meaning systems, and the biological differences between what the chimps need for their current lifestyle and what they’d need for proto-symbolic thinking need not be large. The emergence of meaning does not take place in a vacuum, but requires both a suitable biological development program, an appropriate sociocultural environment, and the ability to have a place in and be an active agent in this environment, a set of

---

8But see also King (2003), who objects to a similar simplicity argument from Corballis (2002).
Origins of language

requirements that is captured in the concept *embodiment* (Lakoff & Johnson, 1999; Zlatev, 1997). Zlatev (2003b; 2004) extends this idea with mimesis (Donald, 1997) as the vital link between embodiment and the sociocultural situatedness that is the key to the conventional and representational nature of language.

The wild chimpanzees would here be caught in an apparent Catch-22 (Heller, 1961): as long as they lack the right sociocultural situatedness during ontogeny, they won’t achieve symbol-based meaning — but if they don’t achieve symbol-based meaning, their children won’t have the right sociocultural situatedness during their ontogeny. Placing young chimps in a human sociocultural milieu will enable these chimps to break out of the loop, and achieve as high a level of meaning as they have the biological equipment for, which appears to be what has happened with Kanzi and his friends.

The ontogeny of symbolic behavior in human children, reviewed in DeLoache (2004), is not totally automatic either. Instead, the understanding of symbols is acquired gradually over the first few years of life, in social interactions with other humans. Understanding the intentionality of adult symbol-users is a key to acquiring symbols. It is interesting to note that children start using language before they have acquired full symbolic abilities, and according to Vihman & Depaolis (2000) the first words of some children “precede[s] any evidence of symbolic capacity...” (p. 135).

Looking at phylogeny rather than ontogeny, our hominid ancestors did manage somehow to bootstrap themselves out of the Catch-22 loop where the wild chimps are stuck. This is likely to have been a — or even the — major breakthrough in the origin of human beings, signaling the takeoff of the cumulative cultural evolution so typical of humanity, and so lacking in other species, even those that do have the rudiments of culture (Tomasello, 1999a; Deacon, 1997). This need not be the result of a major biological change — there may well be threshold effects involved, for example in innate value system (Zlatev, 2001b) or the fidelity of social learning (Rose, 1998; Gabora, 1997).

It has also been shown in computer simulations that, once a population has started to communicate symbolically, a system of arbitrary bidirectional signals, similar to the modern human system, will out-compete all others, assuming that accurate communication confers Darwinian fitness (Pinker, 2000). McArthur (1987) argues for an intermediate stage in which the only words were ‘names’: “Dans les premiers stades du langage il y aurait peut-être eu seulement des “noms” (...) : noms propres et substantifs, ...” (p. 161; cf. Winter & Gärdenfors (1998)), but given the conceptual capabilities of chimpanzees, and the ease with which Kanzi and others have acquired symbols for actions and other non-nominal concepts, this stage appears rather unnecessary. Hurford (2001; 2003c) argues the opposite, that names came *late* in the evolution of cognition and proto-language.

But there is more to human semantics than the ability to handle single symbols, which appears not far out of reach of apes. Many of the concerns of cognitive
linguistics, presented on page 9, belong here, as the interaction between our cognition and our utterances is highly complex and it is not immediately obvious how e.g., our creative use of metaphor can emerge from chimpanzee-level cognition.

Some non-humans have been trained to handle hierarchical concepts, like ‘color’ vs. individual colors (Pepperberg, 2001) or the functional categorization of objects that are perceptually dissimilar (‘tools’, ‘food’, etc.) (Savage-Rumbaugh & Lewin, 1994), but there have been no reports of their doing it spontaneously. In other experiments, apes have been found to understand quite abstract relationships and concepts, like the concept of a relation between abstract relations (Vonk, 2003). This abstract conceptualization ability would appear to be a promising exaptation for the evolution of the construal aspect of fully symbolic systems sensu Sinha (2001). But clear evidence of the network of internal relations between symbols required for full symbolicity is still lacking outside Homo sapiens.

As argued by Turner & Fauconnier (1999), linguistic expressions do not carry meaning in themselves — instead they prompt the listener to construct meaning, using his cognition and knowledge. All language use is filled with implicit meaning that needs to be reconstructed from context. Chimps can handle this in some simple cases, at least as described in Benson et al. (2002), but humans go much further. Even the grounding of a single nominal symbol, connecting it to its referent, may be far from trivial in normal human language (Langacker, 2004). This includes such seemingly purely linguistic concerns as the proper antecedent of a definite noun phrase, which is often not explicit in the conversation, but has to be reconstructed from our knowledge of the situation (Ungerer & Schmid, 1996).

A prominent part of human cognition and language use, vital for conceptualizing new domains, is our ability to identify — or construct — and exploit parallels and analogies between different domains, and to map conceptual structure between domains (Fauconnier & Turner, 1998). There is no evidence that any non-humans can handle e.g., metaphor, even in the most basic sense. The evolutionary origins of our capacity for handling multiple mental spaces and perform blending and other operations between them remains obscure. For some speculations on this issue, see e.g. Wildgen (2004).

But the ability to perform the cognitive operations needed to construct a common category for women, fire, and dangerous things (Lakoff, 1987) or to imagine a monk meeting himself on a mountain (Fauconnier & Turner, 1998) is hardly a prerequisite for the first simple proto-language. It may be noted that blending mental spaces is a fairly late development in ontogeny, well after children acquire basic syntax (Harder, 2003). This means that language at some level is possible without

---

9Incidentally, language training for non-humans appears to assume that they form concepts the same way we do, starting with basic concepts that are in the middle of the hierarchy; cf. page 10. But it is not self-evident that what comes naturally to us comes naturally to non-humans as well.

10Dolphins have similar abilities (Mercado et al., 2000), which may relate to their similar level of success in language learning (Herman et al., 1984). There are mixed reports on the abilities of monkeys, with Thompson & Oden (2000) and Fagot et al. (2001) arriving at opposite conclusions.
the full cognitive machinery of Fauconnier & Turner (1998). In an evolutionary context, it is more plausible that our cognitive and linguistic capabilities grew in parallel, with more advanced cognitive tools emerging gradually, for use in both linguistic and non-linguistic contexts.

The image-schemas of Lakoff (1987) are used as a link between perception and reason. They may well have a pre-linguistic origin, as links between perception and cognition must be phylogenetically old. This means that simple image-schemas may have been an exaptation that provided structures that could be used to bootstrap early proto-symbolic communication into something more language-like. Mandler (2004a) suggests such a process in ontogeny, with preverbal children using ‘perceptual meaning analysis’ to build image-schemas and concepts that underpin language acquisition. McNeil (1996) invokes image-schemas as the first step in her narrative model of language origins, with scripts (Schank & Abelson, 1977) as the second step. Concerning this transition, it is noted by Marcus (2004c) that mental maps and other tools for spatial cognition are present in many animals, and that spatial metaphors are ubiquitous in human language. Image-schemas, commonly spatial as well, may form a bridge between the mental maps of animals and the spatial metaphors of humans.

In conclusion, the change from a chimpanzee-level proto-symbolic system to a full symbolic system does not appear to entail any insurmountable difficulties, as the chimps already have the ability to fulfill, at least with the right upbringing, many of the requirements of a symbolic system, though the evidence of their capacity for detached representation and third-level intentionality is still limited. The remaining requirements that apes lack, notably higher-order value systems and mental space acrobatics, need not be impossible hurdles, and need not be present from the start. The ‘cognitive bootstrapping’ setting off the explosion of cultural and cognitive growth in human beings (Tomasello, 1999a) may even generate them as a byproduct — but a byproduct that then becomes a central part of the process of cultural and linguistic growth.

Concerning the timing of the transition to full human-level semantics, little information is available from the fossil record. But the appearance of art and other symbolic behavior, discussed in Section 9.4.1 may provide some clues.

11.4 Proto-syntax

Communication is certainly possible without syntax, as shown by the communicative abilities of both agrammatics and children at the one-word stage of language acquisition. But there are narrow limits on the complexity of the messages that can be transmitted without syntax. Nowak et al. (2000) have calculated the limits of non-syntactic communication, finding that there is a critical mass of words beyond which syntactic communication becomes highly advantageous. Similarly, there is
a maximum lexicon size imposed by the requirement that each new generation be able to acquire their parents’ lexicon, which again limits the number of possible messages in a syntaxless system (Komarova & Nowak, 2001; Nowak, 2000a) or in a system with list-based rather than rule-based grammar (Nowak et al., 2001). The maximum size of the lexicon depends on acquisition efficacy, which may have provided strong selection pressure towards improved lexicon acquisition, a conjecture supported by the remarkable rate at which human beings acquire new words, one word every several hours, around the clock, sustained for years and sometimes decades.\textsuperscript{11} This point is related to those raised by Oliphant (1999), discussed in Section 11.3, and by Wray (1998), discussed in footnote 8 on page 86.

Useful exaptations or precursors for syntax are difficult to find, though the cognitive capacity needed to handle relational concepts would appear necessary. Several authors seek to base this relational capacity in social interactions, e.g., as argued by Savage-Rumbaugh et al:

Clearly, prior to the emergence of syntax must be the emergence of the concept that one can request that A act on B, where the speaker is neither A nor B. (Savage-Rumbaugh et al., 1986, p. 223-224).

Aiello (1998) argues for ‘social scripts’, structured generalized patterns of social behavior, as syntax precursors, something which ties in nicely with the socio-political hypothesis of language origin (Section 10.5), and it can certainly be argued that the concept in the quote from Savage-Rumbaugh et al. (1986) above is a social script. The structure of the postulated social scripts is discussed in some detail by Worden (1996). Calvin & Bickerton (2000) make a similar argument, going from the social calculations needed for reciprocal altruism, to thematic role structure in meaning, to phrase structure.

The ‘tool-making scripts’ discussed on page 202 is another potential source of proto-syntax, in the same spirit as the social scripts, but with their basis in the structured action sequences of tool making, instead of in social interactions.

The cognitive tools involved in image-schemas and mental spaces, discussed both in Chapter 2 and on page 229 above may also provide useful exaptations for syntax, and may at higher levels coevolve with syntax, creating a feedback loop between language and cognition.

Carstairs-McCarthy (1999) proposes an alternative model in which syntax (as well as semantics) is derived from the combinatorial phonology made possible by the evolution of the human vocal tract. Uriagereka (2001) discusses some weaknesses with this idea (to which Carstairs-McCarthy (2001) responds). An additional problem is the apparent circularity: the evolution of both phonology and syntax presuppose a human vocal tract, according to Carstairs-McCarthy (1999) — but the human vocal tract is an obvious speech adaptation, and thus its evolution must have been driven by pre-existing speech needs.

\textsuperscript{11}I am about 400,000 hours old, and have an estimated (passive) vocabulary well in excess of 100,000 words, summed over languages. This gives an average learning rate of better than one word every four hours, sustained over four decades.
Origins of language

Ontogeny may give useful hints about a possible phylogeny of syntax. The one-word holophrastic stage can, at least in its early phases, reasonably be regarded as non-syntactical, with every utterance an unanalyzed whole (Elliot, 1981). But it is still usable for communication, albeit imperfect — all parents are aware of its strengths and weaknesses, from attempts to understand the will of their one-year-old child — and can thus be taken as a reasonable point of departure for the evolution of syntax (Jackendoff, 1999b; Catania, 2001).

The two-word stage is more interesting as a candidate for proto-syntax. There have been some attempts to argue, within a Chomskian innateness paradigm, that the two-word (and even one-word!) utterances of children ‘really’ have full-blown adult-grammar deep structure, but with all but two words dropped in the surface structure (Wanner & Gleitman, 1982). But these theories impute a complexity to the children’s utterances that goes far beyond that motivated by the available data (Elliot, 1981), and may be directly contradicted by more recent data, as reviewed by Tomasello (2000a) — but see also Pinker (1995) and Santelmann (1995). More plausible models of the two-word stage focus either on very simple grammars (like ‘pivot grammar’), or on child utterances as fragments of adult sentences, or on semantic relationships. Jackendoff (1999b) regards the two-word stage as a simple concatenation of symbols, with meaning determined by semantic and pragmatic considerations, rather than by anything resembling a grammar. This may be a reasonable position for child language, and is certainly a plausible next step in syntax evolution, after the initial one-word stage.

Furthermore, there is a fair degree of continuity in the ontogeny of language, with no clear gaps or sudden jumps separating the different stages:

...we recognize in these simple processes forms of language not separated by a break from our more complicated ones. We see that we can build up the more complicated forms from the primitive ones by gradually adding new forms. (Wittgenstein, 1969, p. 17, quoted in Zlatev (2001a)).

This continuity in ontogeny argues for at least the logical possibility of continuity also in phylogeny, with the gradual evolution of ‘more complicated forms from the primitive ones’.

Armstrong et al. (1995), who support a sign-language origin of language, invoke the temporal structure of signs as the roots of syntax, similar to the model for syntax origins proposed by Condillac (1746); cf. page 158 and Section 9.6.1. They raise an intriguing point, since an iconic sign sequence describing mimetically an action indeed naturally possesses a rudimentary structure that might be a reasonable syntax precursor. If nothing else, iconic sign sequences certainly facilitate the understanding of relationships among concepts (King, 1996), which, as noted above, is one step that must be taken on the road towards syntax. But Armstrong et al. (1995) may take their thesis too far, when they argue not only for signs as proto-syntax, which makes a fair amount of sense, but also attempt to derive the specifics of human syntax, notably the ubiquity of SVO word order among the languages of the world. This is an unnecessary and poorly supported extension that
distracts from their main point (Carstairs-McCarthy, 1996), though there are some possible parallels in the iconic aspects of constituent ordering in modern spoken language discussed by Langacker (2001).

Possibilities for the very earliest stages of proto-syntax have been discussed above. What happens if we look at the other end, at modern human grammar, and contemplate which components might be removable? Removability of a component from a modern grammar reasonably entails its addability to a proto-grammar.

As noted earlier, it is sometimes argued that modern human grammar is a monolithic system that cannot be built piece by piece (Chomsky, 1972). But I will argue that this stepwise construction (or deconstruction, from the perspective of modern syntax) is perfectly possible if the structure of grammar is looked at from an appropriate perspective, and the pieces are added in the right order — not all aspects of grammar are totally interdependent. Certainly, if any component of modern human language is removed, what is left is not equal to modern human language — but it may still be a functional language, if not as rich and expressive as what we’re using today. A language with proto-syntax, missing one or two principles of modern grammar, may not be adequate to write this book — but may nevertheless be adequate for the daily life of proto-humans.

A definitive analysis of the removability of different components or principles of grammar requires that we know what these principles are. However, as noted by Edelman & Pedersen (2004):

...we have, as yet, no comprehensive, psychologically real and neurobiologically grounded process model for language, and with a descriptive model there is a distinct possibility that the features we believe to be important are in fact immaterial (the description’s mathematical appeal notwithstanding). (p. 399).

The Chomskian paradigm is precisely such a mathematically appealing descriptive theory. And even though it dominates the scene, there are several competing theories of grammar that remain viable (Dooley Collberg, 1991), e.g. GPSG (Gazdar et al., 1985), HPSG (Pollard & Sag, 1994), RRG (Van Valin & LaPolla, 1997), or LFG (Bresnan, 1982). And Chomsky himself comes up with a new version of grammar about once every decade (1957; 1965; 1982; 1995), and there is no compelling reason to take for granted that the current incarnation, however elegant, is the ultimate one (for some ironic comments on this point, see Edelman & Christiansen (2003)). Instead, a slightly different perspective will be adopted here, focusing on a few features that modern human languages incontrovertibly possess, regardless of the details of grammatical theory, and see which of those features may be removable, and in which order.

Human languages are universally:

1. **Structured** in the sense that a sentence is not just a random juxtaposition of words, but a sequence ordered according to certain rules (whatever those rules may be).

2. **Hierarchical** in the sense that there are several levels of structures within structures.
3. *Recursive* in the sense that the same rules and structures may recur at different levels in the hierarchy, so that a structure may contain a substructure that is another instantiation of the same structure, in theory repeated *ad infinitum*.

4. *Flexible* in the transformational sense that there are many different ways to express the ‘same’ meaning by moving around words and restructuring sentences according to certain rules.

The list of features above is structured, in the sense of the first list item, in that the features are listed in a non-random order. The *Structured* feature is fundamental — it makes little sense to talk about syntax at all for a non-structured language, and all the other features presuppose *Structured*. But it is perfectly possible to imagine a language that is *Structured* without possessing any of the other features — the two-word stage of child language may be an example, and some pidgin languages may be clearer examples. This means that *Structured* must be the first syntax feature to emerge.

There are, however, at least two ways to get from a one-word stage to *Structured* language, either analytic/holistic or synthetic (Hurford, 2000a; Bickerton, 2003). In the analytic version (Wray, 2000; Arbib, 2003), the units of the one-word stage are holistic utterances, which are then segmented into parts that become independent recombinable morphemes in the next stage, whereas in the synthetic version (Bickerton, 2000, among others), units from the one-word stage are combined into structured utterances in the next stage.

It is not obvious to me, nor to Bickerton (2003), why the segmentation process envisaged by Wray (2000) would be expected to work. A similar process is certainly present in modern-day language acquisition — children first acquire some stock phrases as unanalyzed wholes, and later figure out their internal structure — but that works only because these stock phrases *have* an internal structure, given by the grammar of the adults from whom the child acquires them. As an analogy for the origin of grammar, this is unsatisfactory. Wray (2000) describes a scenario in which people already talking at the one-word stage at some point acquire a grammar from somewhere — apparently not from any linguistic or communicative pressures, but as an exaptation — and start applying it to their language, attempting to identify structure and constituents in their structureless holistic one-word utterances.

Gärdensfors (2003) considers a scenario with grammar as an emergent phenomenon, with the appearance of rules growing out of the patterns that we form when we try to make sense of each other’s utterances, in a process that somewhat resembles that of Wray (2000). But Gärdensfors (2003) places this process at the next step, when going from two-word or pidgin-like proto-language on to full grammar. At that stage, utterances are no longer unanalyzed wholes, but do have substructure in which patterns may be identified.

*A Hierarchical* language must be *Structured*, but need not be either *Recursive* or *Flexible*. The main breakthrough in achieving *Hierarchical* may be the grouping
of words into headed units, and the application of structural rules to headed units as a whole, rather than to individual words (Jackendoff, 1999b). It is possible to have a hierarchy of structures, without the same structure being allowed to recur in infinite recursive chains. I know of no examples of natural human languages like that, but some computer languages (e.g. FORTRAN) do not have full recursivity. Having years of experience of FORTRAN programming, I can testify that this is inconvenient but not fatal — there are always ways to get around the lack of recursivity.

In fact, nothing else in language requires the *Recursive* feature (whereas *Recursive* certainly requires *Hierarchical*). As argued above, it is quite possible to have a language with only partial recursivity, or even none at all, something further supported by the existence of children with SLI (see page 103) whose language apparently lacks recursivity (Bloom, 1999). Therefore, *Recursive* is a prime candidate for being a late evolutionary addition to human grammar. Hauser & Chomsky & Fitch (2002a) propose a closely related hypothesis, that recursion is the sole component of their FLN, i.e., the only component of the language faculty that is narrowly language-specific, with everything else being used also for non-linguistic purposes. Harder (2004), on the other hand, proposes that recursion is not part of the biological language capacity at all — instead recursion emerges culturally from social interactions — which is also consistent with it being a late addition to human language. Hurford (2003b) likewise proposes that the earliest languages lacked subordinate clauses, and thus presumably lacked (at least non-trivial) recursion. Kirby (2002) shows in computer simulations that recursion may emerge through cultural transmission, without being innate.

*Flexible*, like *Recursive*, appears to be an optional feature that can be removed without fatal effects. *Flexible* definitely requires *Structured* to be meaningful, and may require *Hierarchical*, but none of the other features require *Flexible*. Depending on the exact grammatical theory, there may be a lot of obligatory moving around of constituents in a sentence — but in modern human grammar there are also lots of optional movement possibilities, constituents that can be moved or not at the discretion of the speaker. Language would be perfectly functional, if less rich and nuanced, without these discretionary movements. Thus, *Flexible* is also a candidate for being a late evolutionary addition to human grammar.

This adds up to an allowed sequence of successive grammar elaboration, that may be a candidate evolutionary sequence:

1. One-word stage — semantics with no syntax
2. Two-word stage — *Structured*, but with none of the other features.
3. Hierarchical structure, much like a basic phrase structure grammar, but with no recursivity. This means a language without subordinate clauses and other forms of embedding.
4. Recursive syntax (alternatively, *Flexible* may be added before *Recursive*, since they are largely independent of each other).
5. Full modern human grammar.

Each step in this sequence corresponds to a functional communication system, if not as elaborate and rich as the modern human system. And none of the gaps that need to be bridged when going from one step to the next looks anything like the huge chasm commonly pictured between non-syntax and syntax. The steps roughly resemble the stages of child language acquisition, where both recursivity (Goldin-Meadow, 1982) and flexibility (Håkansson, 1994) are fairly late additions.

Jackendoff (1999b) presents a similar sequence, with a similar number of syntactic steps (he has more steps in total, but many of them do not concern syntax, and are covered in Section 11.1–11.3 above). The main differences appear to be that he concatenates symbols into longer strings than two words before adding Structured, and that he does not make a clear distinction between Hierarchical and Recursive.

Gärdenfors (1996) has a hierarchy of communication systems that also shows some parallels with the sequence proposed here. He distinguishes (a) Systems with single elements, (b) Compositional systems, and (c) Systems with grammar, which he then combines with his ‘cued/detached representation’ contrast, discussed on page 7. My stage 1 would in the scenario of Gärdenfors be the crucial step in which detached representation arose for its single elements.

The first step towards syntax, getting started on the sequence at all, might be thought the most difficult — but since we have compelling evidence that stage 1 and quite possibly stage 2 are within reach of chimpanzees, that step involves nothing but activating already existing capabilities, which cannot be an insurmountable problem.

Furthermore, depending on exactly what the underlying grammar looks like, it is not inconceivable that some of these steps can be subdivided even further, with e.g., simple additive recursion evolving before central-embedding sentence-level recursion. Expressed in phrase structure rules:

\[
S \rightarrow NP \ V \ P \quad \text{(11.1)}
\]

\[
NP \rightarrow Adj \ NP \quad \text{(11.2)}
\]

\[
NP \rightarrow NP \ Comp \ S \quad \text{(11.3)}
\]

Both rule 11.2 and rule 11.3 are recursive. But rule 11.2, which adds very little complexity in either production or comprehension, may well be an earlier development than rule 11.3, which is much more difficult to handle even for adult modern humans.

Harder (2004), citing Diessel & Tomasello (2001), discusses a three-step model for the ontogeny of phrase-level recursion:

- Frozen formulae
- Hedging attached to a central clause: ‘I think it’s nice’
The second point above also illustrates another important feature of recursion — in real language production, we often do not build recursive sentences top-down, as they are typically presented in formal grammars. Instead, recursion is typically a matter of embedding a central clause in more and more layers of predication, often as a result of social interactions (Harder, 2004). To take the example of Harder (2004):

1. *Mary to Jack:* George is impossible!
2. *Jack to Joe:* Mary thinks that George is impossible.
3. *Joe to Jack:* Are you sure that Mary thinks that George is impossible?
4. etc...

In this way, recursion can emerge from our ability to handle predicate logic in social interactions.\(^{12}\) Recursion flows naturally from an ability to handle nested predicates like \(P(Q(R(x)))\). This connects Harder’s argument with that of Ghadakpour & Dessalles (2004), who regard explicit predicate logic as a key innovation in the evolution of human cognition and language.

Hurford (2003c) traces the origins of predication to pre-linguistic ‘mental scene descriptions’ (cf. the image-schemas of Lakoff (1987) and the scripts of Schank & Abelson (1977)), and identifies a candidate neural basis for predication in our two parallel neural pathways for both visual and auditory information, one pathway telling us where our perceptions came from, the other telling us what we see. The ‘where’ pathway provides a deictic \(x\) and the ‘what’ pathway a proto-semantic property \(P()\), that are then merged into what amounts to a primitive predicate \(P(x)\). These dual pathways, and the merging of the information they carry, are present not only in humans, but in primates in general, and possibly in mammals in general, which implies that predication in this sense has an ancient history. Hurford (2003c) proposes the alarm calls of vervet monkeys (Seyfarth et al., 1980), discussed on page 120, as examples of such proto-predicates.

Dessalles & Ghadakpour (2003), however, carefully distinguish between predication *sensu* Hurford (2003c), and the predication involved in human cognition and language. They regard Hurford’s predication as mere object recognition, identifying an object holistically from all its features. Human-style predication, on the other hand, is isolating and contrastive, focusing on individual explicit properties of objects (Ghadakpour, 2003). It is not clear whether contrastive predication can be found outside *Homo sapiens*.

Recursivity is plausible as a late addition in phylogeny as well as ontogeny, also because adults have trouble with it — comprehension is poor on sentences with multi-level recursion (Christiansen & Chater, 1999), such as the following sentence built by triple application of rule 11.3 above:

\[^{12}\text{The ‘viewpoint chain’ of Langacker (2001) provides a very similar grounding of recursion, as may others of his ‘paths of mental access’, particularly causal chains.}\]
The rat the cat the dog the man hit chased caught squeaked. In theory, such sentences are grammatical — but they are commonly judged as ungrammatical (Christiansen & Chater, 1999), are difficult and time-consuming to parse (Bates et al., 1999), and are exceedingly rare in natural language. Even in simpler cases, we don’t handle recursion quite as automatically as we do most other syntax — we commonly have to stop and think consciously about the parsing of a recursive sentence with more than one level of recursion, which we rarely need to do in non-recursive cases. The central-embedding recursion discussed here appears to be the worst case for our language processing, which breaks down with no more than three or four nested levels (Marcus, 2004c), but with a sufficient number of levels of embedding most people find other types of recursion non-trivial to parse as well, as soon as the recursion amounts to more than simple concatenation.

Recursion was invoked by Chomsky (1957) as a language feature that was impossible to learn without an innate grammar (cf. footnote 28 on page 182 above), which may be true for infinite recursion. However, as noted above, human language does not in practice allow recursion more than a few levels deep. And recursion to the same depth that humans can handle, has been learned by a connectionist network (Christiansen & Chater, 1999). The network ‘understood’ sentences with the same structure as the rat sentence above, at about the same level as humans do. It can also be noted that Fitch & Hauser (2004) managed to train monkeys to master a simple structured grammar in a toy language, but failed with a recursive grammar.

Concerning earlier levels of grammar, it is worth noting that the postulated social scripts of Worden (1996) and Aiello (1998) have the features Structured and Hierarchical, and even rudimentary recursivity. Other hypotheses of the functioning of primate social intelligence imply similar levels of structure and complexity (Worden, 1996). This is a consequence of the structure and complexity of observed behavior, making it highly plausible that apes had a structured and hierarchical conceptual system available as a language exaptation. For example, baboons apparently recognize the two-tiered nature of their dominance hierarchies, with different kin groups having different rank, and each individual then having a rank within the kin group (Bergman et al., 2003). If the cognitive machinery of the social scripts of Worden (1996) is available and can be used for language, we immediately reach stage 3 in the evolutionary sequence above. Byrne (2000) goes

---

13 Seidenberg & MacDonald (1999) have a point in that such discrepancies between the theoretical competence grammar, and the actual performance of people, is a problem for the Chomskian paradigm.  
14 Networks of this kind have another interesting feature — the same network that’s used and trained for comprehension, can also be used for production of grammatical sentences within the same grammar (Christiansen & Chater, 1999). This circumvents the argument that language evolution is unlikely since both production mechanisms and comprehension mechanisms have to evolve in tandem as argued by e.g., Vanechoutte & Skoyles (1998) — cf. the ‘first speaker’ fallacy discussed on page 18. That argument is specious anyway, since coevolution of two coupled features is a common and fairly well understood phenomenon in biology, but it is nevertheless nice to have it undercut so cleanly.
one step further, and postulates more general cognitive structures for dealing with structured, hierarchical problems, which would be eminently exapt-able to syntax processing, whereas Bickerton (2000) invokes hierarchical semantic structures as an exaptation for syntax.

Pidgin languages, with their highly simplified grammatical structure, can also be invoked here, at the very least as examples of functional languages without all the features of full human syntax (McMahon, 1994), and possibly as a modern-day example of what an intermediate stage in the evolution of syntax might have looked like (Bickerton, 1995). Pidgins commonly lack both Hierarchical and Recursive, consisting of just a linear structure of words, illustrating level ‘2 1/2’ syntax, with the same basic complexity level as the two-word stage, but with more words and longer, more semantically complex utterances. Jackendoff (1999b) and Klein (2001) invoke ‘The Basic Variety’ (BV), a kind of minimalistic acquisition of a foreign language up to the level where basic communication can take place but no further, as another type of ‘living fossil’ proto-language. BV has no morphology and no hierarchical structure, and a very simple semantically based word order.

The selection pressures driving evolution from one stage to the next, can be related to the increasing complexity of proto-human society, along the lines of the calculations of Nowak et al. (2000) which indicate that a simple one-word language is optimal in a simple world, but that structured hierarchical language becomes advantageous once a complexity threshold is passed in the contents of the communication that the language is used for; cf. Kirby (2000). In a computer simulation, Hashimoto & Ikegami (1996) have shown that if reproductive success is based on communicative success, gradual evolution towards more and more complex grammars can result, whereas Kirby (2000) and Hurford (2000b) instead simulate the emergence of syntax through purely cultural evolution, totally disconnected from biological fitness. It would appear that both the biological and cultural routes are feasible, at least at in the simple worlds of these simulations.

For the last stages of the origin of modern grammar, Hurford (2003b) invokes the observed unidirectionality of grammaticalization processes. Since grammaticalization is a process of delexification, going from lexical stems to function words and inflections, it appears plausible that the earliest languages lacked function words and inflections. This would be a functional, if pidgin-like, language, that could then evolve into our present languages through normal processes of diachronic language change, without any further need for biological evolution. This also means that morphology need not be an issue in discussions of language origins, since it is largely the product of grammaticalization.

In conclusion, the gradual evolution of modern human grammar through several functional intermediate stages, appears perfectly possible. The exact path of evolution is speculative, due to the dearth of data on the structure of actual proto-languages, but no insurmountable obstacles are visible. Some of the required tran-
Table 11.1. *A summary of the starting position and possible exaptations and early evolutionary steps, in the four areas of proto-language discussed in this section.*

<table>
<thead>
<tr>
<th>Linguistic status of chimpanzees</th>
<th>Speech</th>
<th>Gestures</th>
<th>Semantics</th>
<th>Syntax</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rudimentary and poorly controlled sounds, but adequate perception</td>
<td>Good dexterity and perception</td>
<td>Can be taught many symbols, and use them for communication</td>
<td>Can be taught rudimentary syntax</td>
<td></td>
</tr>
<tr>
<td>Communicative gestures</td>
<td>Basic symbolic, conceptual and relational capacity</td>
<td>Social scripts?</td>
<td>Image schemas?</td>
<td></td>
</tr>
<tr>
<td>Neural control of vocalizations, e.g., MacNeilage-syllables</td>
<td>Spontaneous symbol acquisition, leading to one-word utterances</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Steps remaining

- Full neural and sequencing control, moved larynx, improved tongue motor control
- Hierarchical semantics, mental spaces, abstract multi-level classification
- Structured
- Hierarchical
- Recursive
- Flexible

Positions can actually be observed today, either in ontogeny, or in e.g., transitions from pidgin to creole language.

### 11.5 How can all the protos be combined?

As shown in the preceding sections, there are plausible gradual paths for the evolution of all the main aspects of language, starting from the current capacities of chimpanzees, the best available model for the capacities of our ape ancestors. But in order for language as a whole system to evolve, the various components have to
be combined. Some of them are independent, but others need to be assembled in a certain order, in order to maintain a workable and stepwise improved communication system. Two points are immediately apparent:

- At least one modality for production and perception needs to be up and running, at least at a primitive level, right from the start, otherwise language cannot be used for communication.
- Semantics must come before syntax. Semantics without syntax is a workable communication system (Bookless & Mortley, 1996; Bickerton, 2003), but syntax without semantics is pointless for communication.

Given those two conditions, there is considerable freedom to choose different evolutionary paths. In Table 11.1, I attempt to summarize the first steps. As stated earlier, the starting point for the evolution of the human language abilities is taken as the abilities of our last common ancestor with the chimpanzees, which, in the absence of parallel evolution, can be assumed similar to the abilities of modern chimpanzees.

Signing capacity is available right from the start, and primitive speech, at the level of MacNeilage & Davis (2000), is not far off, so the modalities need not concern us further here. As for semantics, the results from ape language studies, e.g., Savage-Rumbaugh & Lewin (1994), indicate that the conceptual and protosymbolic abilities needed for language at the one-word stage, the first step in my proposed evolutionary sequence on page 235, are within the reach of chimpanzees — full symbolicity, with its emphasis on relations between symbols, would hardly appear vital at the one-word stage, where symbols are processed one at a time.

After that, a co-evolution of semantics and syntax would appear likely, with the various aspects of syntax appearing in the proposed order. The evolution of syntax would be driven by more and more complex communication demands, very likely caused by a socio-political arms race. The same communication complexity would affect also semantics in the same direction, first simply towards more and more words, but later also towards a more structured semantics, with derivations and verb frames and hierarchical structures of concepts supplying the complex expressive abilities needed. The details here are of necessity speculative.

As the number of words grew beyond all bounds, increased demands would be placed on the modality used for language, providing selection pressure towards the production of clear and unambiguous and economical distinctions between large numbers of closely spaced sounds or signs. This is precisely what the modern human vocal apparatus provides today, so its evolution very likely took place at this stage, driven by these demands for communicative clarity and precision. Combinatorial phonology is likely to be another product of the same selection pressures, making it possible to form a vast number of distinct words from a modest number of sounds.
11.6 Summary

- **Proto-speech.**
  - Speech perception not a problem for apes — no proto-perception stage needed.
  - Relaxation of the automaticity of ape vocalizations needed.
  - Jaw motions may lead to proto-syllables.
  - A capacity for vocal imitation (present in humans, absent in chimps) would be a useful exaptation.

- **Proto-gestures.** No proto-gesture stage needed — apes have the equipment needed.

- **Proto-semantics.** Proto-symbolic level within reach of enculturated apes — no *biological* changes needed at the proto level.

- **Proto-syntax.**
  - Proto-syntactic stages in ontogeny may provide clues.
  - Syntax imperfectly understood, but some features uncontroversibly part of modern human syntax:
    - Structured
    - Hierarchical
    - Recursive
    - Flexible
    These features may in principle be added to a syntax-less proto-language, one after the other, in the order given above.
  - Recursion is a key, late-acquired feature of syntax.
  - Other sequential and hierarchical cognitive tasks may provide useful exaptations for syntax.

Further reading


The evolutionary history of us human beings remains murky in some details, but it is established beyond all doubt that we have evolved from ape-like ancestors, over the course of some millions of years. Those ancestors presumably had ape-like communicative abilities, from which we can conclude that an evolutionary transition to human language abilities must have taken place.

Our main bodily language adaptations are in the brain — but the neural organization of language is not well understood, beyond a gross anatomical level. Furthermore, it is not well understood to what extent language is innate and genetically encoded, much less what such encoding would entail at the neural and genetic level. The only thing that can be concluded is that it is not a simple matter of a dedicated ‘language organ’ with an innate Universal Grammar. A detailed understanding of how our brain has adapted to language remains a distant goal.

The human vocal apparatus is clearly adapted for speech in several respects, notably in the shape of the vocal tract and the neural wiring both of the control circuits in the brain and of the peripheral systems involved. But language, as distinct from speech, is modality-independent, and speech need not be the original modality. A gestural origin of language remains a distinct, if unproven, possibility.

Linguists commonly emphasize the gulf between human and animal communication. Hints of more language-like communication, with functionally word-like signals, can nevertheless be found among some non-humans, notably monkeys. Clearer examples of abilities partially bridging the gulf can be observed in apes (and possibly dolphins and parrots) that from an early age have been immersed in a human communicative and cultural environment, with the communication adapted to a modality the animals can use. Such enculturated apes demonstrate apparent language abilities that approach those of a human two-year-old. When a human child displays similar abilities, we call it language without the slightest hesitation — but many people are less eager to grant the apes the same recognition.

But we already know that we have evolved from apes, and that our language has evolved along the way — why should we be so surprised to find traces of proto-linguistic abilities among our nearest relatives? As ape-language pioneer
Origins of language


Do we need to be afraid? Would not our world be richer if they did have minds and language, not only for the light that would be cast on our own origins, but also just for the company — *Homo sapiens* would no longer be alone in this huge and mostly inhospitable universe.

But for company, we need not only beings with language, but also beings with minds of their own, which brings us to the next main question, concerning the relation of language and mind, and the possible presence of mind in non-humans.

In the case of ape minds as well as ape language, the preponderance of the evidence appears to be on the side of the presence rather than absence of at least the rudiments of mind, self-awareness, and theory of mind in our nearest relatives. The acquisition of self-awareness and a theory of mind appears to be a gradual process in both ontogeny and phylogeny, with adult apes on a par with humans around the age of two. This is remarkably similar to the language skills displayed by the bonobo Kanzi, who also matches a human two-year-old. This is not only evidence against any qualitative differences between humans and non-humans in either language or mind, but also evidence against there being qualitative differences between human language acquisition and other cognitive developments in human ontogeny. This is further supported by the conclusion reached on page 134, that the ape language acquisition results indicate that the postulated ‘language acquisition device’ of humans cannot be language-specific.

There certainly remain non-trivial differences between on one hand human thought and human language, and on the other hand the limited capabilities displayed by chimpanzees and dolphins and a few other non-humans. But as more animal studies accumulate, the differences appear more and more quantitative (however huge) rather than qualitative, and the evolutionary bridging of the gap between human language and animal communication appears far from impossible.

There are many conceivable paths of language evolution. But some hypothetical paths are more plausible than others, and some hypotheses are wholly untenable. Returning to the five dimensions of language evolution hypotheses discussed in Chapter 9, Table 12.1 summarizes the viability of their 32 possible combinations.

As can be seen in Table 12.1, the preponderance of the evidence favors language being an early gradual adaptation. All other possibilities are either totally untenable or highly unlikely, with the possible exception of early non-innate sign language emerging as a spandrel based on available exaptations.

Late language is excluded mainly because there is evidence of speech adaptations earlier in the fossil record than the proposed timeframe for late language. Sudden emergence of language is largely excluded as well, and totally excluded if grammar is innate, because complex genetically specified systems, such as our speech adaptations, and particularly an innate universal grammar, take time to evolve (Worden, 1995), which implies gradual emergence.
Table 12.1. Evaluation of the viability of different combinations of locations along the five dimensions of language evolution hypotheses discussed in Chapter 9. The salience of the text indicates the confidence level of the assessment. The evidence and arguments that the table is based on are spread out throughout Chapter 9.

<table>
<thead>
<tr>
<th></th>
<th>Adaptation</th>
<th>Spandrel</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Early</td>
<td>Late</td>
</tr>
<tr>
<td><strong>Gradual</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speech</td>
<td>Innate</td>
<td>Possible</td>
</tr>
<tr>
<td></td>
<td>Learned</td>
<td>Possible</td>
</tr>
<tr>
<td>Gesture</td>
<td>Innate</td>
<td>Possible</td>
</tr>
<tr>
<td></td>
<td>Learned</td>
<td>Possible</td>
</tr>
<tr>
<td><strong>Sudden</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speech</td>
<td>Innate</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Learned</td>
<td>No</td>
</tr>
<tr>
<td>Gesture</td>
<td>Innate</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Learned</td>
<td>Unlikely</td>
</tr>
</tbody>
</table>

Given an early gradual adaptive emergence of language, either innate or learned language are defensible hypotheses, as are either speech or sign language as the primary modality. The innateness debate will not be resolved here, but a few more things can be added concerning modality. Gestures first has the clear advantage that the production capacity was already available, but since speech is the dominant modality today, Occam would favour speech as dominant from the start. A hybrid solution is an attractive possibility, with the earliest stages of language mainly signed, but with the signs supplemented by the simple sounds that these early speakers could produce, a scenario much like the fictional one depicted by Auel (1980), though much earlier in history. Within such a hybrid system, selective pressure towards improved communicative clarity would drive the evolution of the modality which had room for improvement — speech, but not really sign, as the dexterity used for signing was already highly optimized. Our vocal apparatus would then be optimized for speech, the results of which we hear today.

Fossil evidence indicates that this speech optimization got started well before the emergence of *Homo sapiens*, almost certainly more than half a million years ago, and possibly more than a million, probably in *Homo erectus*. As the speech optimization, with its accompanying costs, would not occur without strong selec-
tive pressure from heavy demands for complex verbal communication, this implies that *Homo erectus* already possessed non-trivial language abilities.

There is no real evidence indicating just how complex syntax *erectus* had. They may have been at the one-word stage, or they may have had nearly full human language — it is difficult to imagine any way to tell. On one hand, *erectus* is the first hominid with a brain size approaching the modern human range — there are modern humans alive today with *erectus*-sized brains and excellent language skills — and they were also the first to spread out to many different habitats on different continents. But on the other hand their comparatively simple, static culture argues against their having modern human cognitive skills. In particular, it is quite clear that they lacked the cumulative cultural evolution that is so characteristic of modern humans. Given that they are different from modern humans in such fundamental ways, their having full modern human language appears unlikely.

Both modern humans and Neanderthals are descendants of *erectus* (with possibly one or two intervening chronospecies), which implies that both inherited the language abilities of *erectus*. We know the language skills of modern humans — but this common inheritance implies that Neanderthals also had language, at the very least at the *erectus* level.

Modern humans, after parting company with the Neanderthals perhaps half a million years ago, acquired the remaining features of modern language in parallel with acquiring modern human anatomy. Both aspects were finished before modern humans started spreading over the world, more than 60,000 years ago. The last common ancestor of all humans today, probably living in Africa not so long before this exodus, is the likely speaker of Proto-World, the common ancestor of all the modern language families, and the earliest language which we may have any remote hope of ever reconstructing. But there is no reason to believe that this Proto-World was the *first* language spoken — as discussed above, our ancestors may have had language for a million years already. The details of those earlier proto-languages are likely to remain opaque for the foreseeable future.

Concerning the causative side of the issue, the question of what drove the evolution of language, the socio-political hypothesis (see Section 10.5) appears the most plausible, though it is unlikely that any single factor dominated in the evolution of such a versatile tool as language — social, political, teaching, sexual, economical, and technological factors may all have contributed to a greater or lesser extent to the evolution of language. They are listed here in probable order of importance, though this is rather speculative.

There is much we still do not know about the origin of language, and much that we may never know. It is an issue of enormous complexity. But the field is not totally devoid of relevant data, and it is not open to unbridled speculation once all pertinent data are taken into account — the dictum of the Linguistic Society of Paris from 1866 is no longer motivated. Unfortunately, all the knowledge that bears upon the issue of language origins is spread over many different disciplines,
and workers within a single discipline rarely capture the full story. I have made an attempt here to integrate as much as possible across disciplines, and tried to determine which hypotheses of language evolution remain consistent with all available data, and which can be excluded as untenable or unlikely.

The result is that I can conclude with some confidence that the evolution of language is possible, in a process in which both biological and cultural-memetic evolution play prominent roles. There is no single uniquely determined history of language origins, but major avenues of speculation have been closed off, and the possibilities have been narrowed down considerably. We have an ancient history as speaking and probably signing apes, our language gradually evolving memetically over hundreds of millennia towards improved communication and learnability, and us evolving biologically even more gradually towards improved speakers and learners of language.
References


Agnetta & Hare & Tomasello (2000). Cues to food location that domestic dogs (Canis familiaris) of different ages do and do not use. Animal Cognition 3: 107-112


Alocot et al. (2000). Pitch and timing ability in inherited speech and language impairment. Brain & Language 75: 34-46


Ansaldo, U. (2001). The hybridity cline hypothesis: On language change and creolization. *presented at 15th International Conference on Historical Linguistics, La Trobe University, Melbourne, Australia, 13-17 August 2001*


Arnason & Xu & Gullberg (1996b). Comparison between the complete mitochondrial DNA sequences of Homo and the common chimpanzee based on nonchimeric sequences. Journal of Molecular Evolution 42: 145-152

Arnason et al. (1996a). Pattern and timing of evolutionary divergences among hominoids based on analyses of complete mtDNAs. Journal of Molecular Evolution 43: 650-661


Arsuaga et al. (1999). The human cranial remains from Gran Dolina Lower Pleistocene site (Sierra de Atapuerca, Spain). Journal of Human Evolution 37: 431-457


Baba et al. (2003). Homo erectus calvarium from the Pleistocene of Java. *Science 299*: 1384-1388


Barringer et al. (1965). *Social change in developing areas: a reinterpretation of evolutionary theory*. Cambridge, MA: Schenkman


Beall et al. (2004). Higher offspring survival among Tibetan women with high oxygen saturation genotypes residing at 4,000 m. *Proceedings of the National Academy of Sciences* 101: 14300-14304


Origins of language


Bjorklund et al. (2002). The generalization of deferred imitation in enculturated chimpanzees (Pan troglodytes). *Animal Cognition* 5: 49-58


Blumenschine et al. (2003). Late Pliocene Homo and hominid land use from Western Olduvai gorge, Tanzania. *Science* 299: 1217-1221


Boë & Maeda & Heim (1999). Neandertal man was not morphologically handicapped for speech. *Evolution of Communication* 3: 49-77


References


Cooper & Drummond & Willerslev (2004). Ancient DNA: would the real Neandertal please stand up? *Current Biology* 14: R431-R433


de la Torre et al. (2003). The Oldowan industry of Peninj and its bearing on the reconstruction of the technological skills of Lower Pleistocene hominids. Journal of Human Evolution 44: 203-224


Dean et al. (1998). On the phylogenetic position of the pre-Neandertal specimen from Reilingen, Germany. Journal of Human Evolution 34: 485-508


d’Errico et al. (2003). Archaeological evidence for the emergence of language, symbolism, and music — an alternative multidisciplinary perspective. *Journal of World Prehistory* 17: 1-70
Descartes, R. (1637). *Discours de la méthode de bien conduire sa raison et chercher la vérité dans les sciences; plus la Dioptrique, les Météors et la Géometrie, qui sont des essais de cette methode*. Leyden


Dogil et al. (2004). Where and how does grammatically geared processing take place – and why is Broca’s area often involved. A coordinated fMRI/ERBP study of language processing. *Brain & Language* 89: 337-345


Dupanloup et al. (2003). A recent shift from polygyny to monogamy in humans is suggested by the analysis of worldwide Y-chromosome diversity. *Journal of Molecular Evolution* 57: 85-97

Dyson & Eddington & Davidson (1920). A determination of the deflection of light by the Sun’s gravitational field from observations made at the solar eclipse of May 29, 1919. Philosophical Transactions of the Royal Society of London 220:291-333


Fagot & Wasserman & Young (2001). Discriminating the relation between relations: the role of entropy in abstract conceptualization by baboons (Papio papio) and humans (Homo sapiens). Journal of Experimental Psychology 27: 316-328

References


Farroni et al. (2002). Eye contact detection in humans from birth. Proceedings of the National Academy of Sciences 99: 9602-9605


Fernández-Jalvo et al. (1999a). Human cannibalism in the Early Pleistocene of Europe (Gran Dolina, Sierra de Atapuerca, Burgos, Spain). Journal of Human Evolution 37: 591-622


Ferrari et al. (2000). The ability to follow eye gaze and its emergence during development in macaque monkeys. Proceedings of the National Academy of Sciences 97: 13997-14002


Origins of language


Geiger, L. (1868). *Ursprung und Entwicklung der menschlichen Sprache und Vernunft* volume i. Stuttgart


References


References

Language acquisition: the state of the art chapter II.2, pages 51–77. Cambridge: Cambridge University Press


Goodall, J. (1971). In the shadow of man. London: Collins


Goren-Inbar et al. (2002). Nuts, nut cracking, and pitted stones at Gesher Benot Ya’aqov, Israel. Proceedings of the National Academy of Sciences 99: 2455-2460


Griffin, D. R. (2001). Animals know more than we used to think. Proceedings of the National Academy of Sciences 98: 4833-4834


Hagoort et al. (2004). Integration of word meaning and world knowledge in language comprehension. Science 304: 438-441


Hare & Call & Tomasello (2001). Do chimpanzees know what conspecifics know? *Animal Behaviour* 61: 139-151


Hare et al. (2002). The domestication of social cognition in dogs. *Science* 298: 1634-1636

Hare et al. (2003). Do capuchin monkeys, Cebus apella, know what conspecifics do and do not see? *Animal Behaviour* 65: 131-142


Henshilwood et al. (2002). Emergence of modern human behavior: Middle Stone Age engravings from South Africa. *Science* 295: 1278-1280

Henshilwood et al. (2004). Middle Stone Age beads from South Africa. *Science* 304: 404


Hickok et al. (2002). Role of the left hemisphere in sign language comprehension. *Brain & Language* 82: 167-178


Kaas, J. H. (2000). Why is brain size so important: design problems and solutions as neocortex gets bigger or smaller. *Brain & Mind 1*: 7-23


Kant, I. (1781). *Kritik der reinen Vernunft*. Königsberg

Kant, I. (1790). *Kritik der Urteilskräft*. Königsberg


Kappelman et al. (1996). Age of Australopithecus afarensis from Fejej, Ethiopia. *Journal of Human Evolution* 30: 139-146


Kimbel et al. (1996). Late Pliocene Homo and Oldowan tools from the Hadar formation (Kada Hadar member) Ethiopia. Journal of Human Evolution 31: 549-561


King, A. J. (2002). Neural plasticity: how the eye tells the brain about sound location. Current Biology 12: R393-395


Kivisild et al. (2000). Questioning evidence for recombination in human mitochondrial DNA. Science 288: 1931a


Krings et al. (1999). DNA sequence of the mitochondrial hypervariable region II from the Neanderthal type specimen. *Proceedings of the National Academy of Sciences* 96: 5581-5585


Kullmer et al. (1999). The first Paranthropus from the Malawi rift. *Journal of Human Evolution* 37: 121-127


La Mettrie, J. (1742). *Man a machine*. LaSalle, IL: Open Court, 1943


Leibniz, G. W. (1710). Brevis designatio meditationum de originibus gentium ductis potissimum ex indicio linguarum. Miscellanea Berolinensia, I-16


Lenski et al. (2003). The evolutionary origin of complex features. Nature 423: 139-144


Lorenzo & Arsuaga & Carretero (1999). Hand and foot remains from the Gran Dolina Early Pleistocene site (Sierra de Atapuerca, Spain). *Journal of Human Evolution* 37: 501-522
MacLeod et al. (2003). Expansion of the neocerebellum in Hominoida. *Journal of Human Evolution* 44: 401-429


Marino et al. (2000). Relative volume of the cerebellum in dolphins and comparison with anthropoid primates. *Brain, Behavior & Evolution* 56: 204-211


Natsopoulos et al. (1998). Do the hands talk on the mind’s behalf? Differences in language ability between left- and right-handed children. *Brain & Language 64*: 182-214
Natsopoulos et al. (2002). Differences in language performance in variations of lateralization. *Brain & Language 82*: 223-240
Newbury et al. (2002). FOXP2 is not a major susceptibility gene for autism or specific language impairment. *American Journal of Human Genetics 70*: 1318-1327


Nishimura et al. (2003). Descent of the larynx in chimpanzee infants. Proceedings of the National Academy of Sciences 100: 6930-6933


Nowak, M. A. (2000a). The basic reproductive rate of a word, the maximum size of a lexicon. Journal of Theoretical Biology 204: 179-189

Nowak, M. A. (2000b). Homo grammaticus. Natural History 12/00: 36-44


Petitto et al. (2004). Baby hands that move to the rhythm of language: hearing babies acquiring sign languages babble silently on their hands. *Cognition* 93: 43-73


References


Poelarends et al. (2000). Roles of horizontal gene transfer and gene integration in evolution of 1,3-dichloropropene- and 1,2-dibromoethane-degradative pathways. *Journal of Bacteriology* 182: 2191-2199


Potts et al. (2004). Small mid-Pleistocene hominin associated with East African acheulean technology. *Science* 305: 75-78


Povinelli, D. J. & Vonk, J. (2004). We don’t need a microscope to explore the chimpanzee’s mind. *Mind & Language* 19: 1-28

Povinelli et al. (2003). Chimpanzees spontaneously alter the location of their gestures to match the attentional orientation of others. *Animal Behaviour* 66: 71-79


Reno et al. (2003). Sexual dimorphism in Australopithecus afarensis was similar to that of modern humans. *Proceedings of the National Academy of Sciences* 100: 9404-9409


Rook et al. (1999). Oreopithecus was a bipedal ape after all: evidence from the iliac cancellous architecture. *Proceedings of the National Academy of Sciences* 96: 8795-8799


Rumbaugh et al. (1975). Conversations with a chimpanzee in a computer-controlled environment. *Biological Psychiatry* 10: 627-641


Saussure, F. (1916). *Cours de linguistique générale*. Ch Bally & A Sechehaye


Savage-Rumbaugh et al. (1985a). The capacity of animals to acquire language: do species differences have anything to say to us? *Philosophical Transactions of the Royal Society of London B* 308: 177-185


Scally et al. (2001). Molecular evidence for the major clades of placental mammals. *Journal of Mammalian Evolution* 8: 239-277


Schoenemann et al. (2000). Brain size does not predict general cognitive ability within families. *Proceedings of the National Academy of Sciences* 97: 4932-4937


Semaw, S. (2000). The world’s oldest stone artifacts from Gona, Ethiopia: their implications for understanding stone technology and patterns of human evolution between 2.6-1.5 million years ago. *Journal of Archaeological Science 27*: 1197-1214


References


Tattersall, I. (2000). Once we were not alone. *Scientific American 01/00*: 38-44


Tavaré et al. (2002). Using the fossil record to estimate the age of the last common ancestor of extant primates. *Nature 416*: 726-729


References


Tomasello & Hare & Fogleman (2001). The ontogeny of gaze following in chimpanzees, Pan troglodytes, and rhesus macaques, Macaca mulatta. *Animal Behaviour* 61: 335-343


Tomasello et al. (in press). Understanding and sharing intentions: the origins of cultural cognition. *Behavioral & Brain Sciences, target article*


Vonk, J. (2003). Gorilla (Gorilla gorilla gorilla) and orangutan (Pongo abelii) understanding of first- and second-order relations. *Anim Cogn* 6:77-86


Watve et al. (2002). Bee-eaters (Merops orientalis) respond to what a predator can see. *Animal Cognition* 5: 253-259


Westergaard et al. (2004). Tufted capuchins (Cebus apella) attribute value to foods and tools during voluntary exchanges with humans. Animal Cognition 7: 19-24


WoldeGabriel et al. (2001). Geology and paleontology of the Late Miocene Middle Awash valley, Afar Rift, Ethiopia. Nature 412: 175-177


Xitco & Gory & Kuczaj (2004). Dolphin pointing is linked to the attentional behavior of a receiver. *Animal Cognition* 7: 231-238


Zietkiewicz et al. (1998). Genetic structure of the ancestral population of modern humans. *Journal of Molecular Evolution* 47: 146-155


Index

a priori knowledge, 22
Aanen, 1
Abdellah, 171, 187
Abelson, 230, 237
Aboitiz, 89, 91
Abu-Akel, 150
Acheulean, 65, 85, 200
Acheulian, see Acheulean
Ackermann, 84
acquisition, 9, 102, 103, 179, 180, 185
Baldwin effect, 30
biased, 180
bilingual, 183
connectionism, 185
cultural, 185
ecological, 185
empiricist, 185
LAD, 101
language, 179
language–nonlanguage cor-
relations, 102
lesion studies, 101
lexicon, 231
modularity, 101
negative evidence, 181
socio-perceptual, 185
statistical, 186
usage-based, 185
Williams syndrome, 105
Acredolo, 175
Adams, 86
adapid, 45
adaptation, 163
Adcock, 70
Aegyptopithecus, 45
Africa
  Out of, 67
Afropithecus, 46
Agnetta, 150
agriculture, 1
Ahlberg, 89
Aiello, 60, 76, 78, 79, 114, 200, 202, 207, 208, 214, 231, 238
Aitchison, 79
Akeakamai, 136
Akre, 99
Alarcón, 172
alarm calls, 121
  evolution, 125
  iconicity, 121
  monkeys, 8, 124
  ontogeny, 120
  vervet monkeys, 120
Alba, 200
Alcock, 103
Alemseged, 59
Alex, 137
Alia, 135
Allard, 44
Allen, 59
Allman, 117
allometry, 91
Allott, 6, 221
Allwood, 9, 11
Almazan, 105
Almog, 145
Alonso, 70
Alper, 94
Altman, 186
Alvard, 21, 196
Alzheimer, 94
Ambrose, 60, 76, 78, 79, 112, 113, 197
anamensis, 57, 61
bahrreghazi, 58, 61
body size, 113
boisei, 58, 61
brain, 113
garhi, 60, 61, 200
robustus, 58, 61, 200
sex life, 204
tool use, 60, 200
autism, 148, 215
Awadalla, 68
Ayala, 51, 53, 57, 60, 68
Ayub, 70
Baba, 80
Babel
  Tower of, 161
baboon, 98
Backwell, 60, 97, 199
Bahn, 72, 169, 170
Bailey, 150
Baker, 28, 119, 126
Baldwin, 22, 30
Baldwin effect, 30, 189
Balter, 52, 58, 64, 66, 72, 84, 104, 168, 198, 215
Bar-Yosef, 168
Barbujani, 68
Barbulescu, 49
Bard, 115, 123
Bargmann, 88
Barham, 168
Barinaga, 187
Barnes, 169
Baron-Cohen, 148
Barrett, 29
Barthe, 105
Barton, 45, 91, 96
Origins of language

basal ganglia, see brain
Basho, 110
Basic Variety, 239
Batali, 133
Bates, 37, 101, 102, 106, 134, 161, 167, 175, 181, 185, 238
bats, see Chiroptera
Battell, 103
Bazter, 70
Bauer, 45, 176
Beall, 188
Beauchamp, 209
Bede, 182
Bednarik, 169
Beehner, 204, 208
Beerli, 71
Begun, 48, 53
Bennett, 103
Batz, 70
Bateson, 103, 104, 112
Bjorklund, 177
Black, 64, 66
Blackmore, 2, 29, 40
Blades, 148
Blake, 2
blending, 10
ontogeny, 229
Binkofski, 106
Bino’s, 105
Bipedalism, 53–56, 80, 209
Bloom, 3, 39, 80, 88, 111, 139, 144, 149, 150, 161, 162, 164, 166, 172, 193, 196, 202, 245
Blumenscheine, 63
Blumstein, 108, 110, 121
Bodnar, 7
Boe, 79
Boerlijst, 17
Boesch, 1, 20, 113, 121
Bogin, 207
Bolus, 72
Bonnefille, 97
bonobo, 48
language, 132–138
sex life, 203
Bonvillian, 98, 99
Bookless, 241
Boorman, 98, 99
Bossert, 53
Botha, 18, 161, 163, 164, 166
Botstein, 7
Bouhaddier, 57
bow-wow theory, 160
Bower, 199
Bowerman, 144
Bowers, 109
Bowles, 215
Boyd, 21, 28, 97
Boy, 44
Boysen, 122
Bradbury, 177
Braga, 59, 113
brain, 88
allometry, 91
and mind, 145
anteroior cingulate cortex, 94
arms race, 97
basal ganglia, 84, 104, 105
cerebellum, 84, 91
cognition, 94
cost, 94
diencephalon, 91
diet driven, 96
environmental variability, 97
epiphenomenal, 95
evolution, 89, 95
fetal development, 92, 111, 206
fossils, 113
imaging, 88
inferior frontal gyrus, 109
lateralization, 98
lunate sulcus, 113
mammals, 89
medial prefrontal cortex, 154
metabolism, 94
navigation driven, 96
neocortex, 89, 94, 97
occipital lobe, 113
origins of, 89
perisylvian fissure, 105
plasticity, 111, 134, 187
politics, 97
prefrontal lobes, 91, 95
Rolandi operculum, 110
size, 89, 94
socially driven, 97
specific language impairment, 105
striatum, 91
tool driven, 97
Brakke, 133
Brandt, 112, 113
Brannon, 122
Brantingham, 197
breasts, 205
nipples, 163
breathing control, 82
Bresnan, 233
bricolage, 16, 166
Bridgeman, 199
Bridgman, 73
Brinck, 195
Briscoe, 30, 37, 38
Broadfield, 64
Broca, 84, 98, 100, 109
aphasia, 106
fossils, 115
in apes, 115
mirror neurons, 176
monkeys, 125, 176
neuroimaging, 110
non-linguistic use of, 107, 110
sequencing, 84
Brockelman, 203
Bromham, 68
Brookfield, 70
Brooks, 49, 67, 69, 73, 168, 169

Index 333

Brown, 20, 35, 65, 68, 107
Brumm, 122
Bruner, 71, 113
Brunet, 52, 58
Buccino, 106, 176
Bühler, 223
Buk, 120, 127, 221
Bull, 190
Buller, 31, 88, 100, 101, 111, 187
Burke, 44
Burling, 19, 181
Butts, 29
Byrne, 17, 20, 94, 96, 97, 112, 121, 123, 135, 150, 151, 157, 194, 207, 211, 212, 222, 238
Caenorhabditis elegans, 88
Call, 123, 177
Callebaut, 22
Calvin, 15, 83–85, 103, 162, 192, 221, 231
Cameron-Faulkner, 181, 182
Campbell, 21–23, 37
Campbell monkeys, 124
Cane, 54
canine size, 204
Cann, 68, 70, 74
cannibalism, 73
cantalupo, 115
Capaldo, 197
Caporaal, 31, 209
Caramazza, 108, 109
Caramelli, 71
Carbonell, 66
Carey, 122
Carr, 84
Carretero, 67
Carroll, 51, 171
Carruthers, 143, 144, 201
Carstairs-McCarthy, 121, 162, 164, 177, 231, 233
Cartesius, 146
Caspari, 67, 69
Catania, 6, 31, 232
catarrhine, 45
categorical perception, 87
Cavalli-Sforza, 67–70, 74
CDS, see Child directed speech
Cela-Conde, 51, 53, 57, 60
cerebellum, 84
Chafetz, 181
Chaimanee, 47
Challis, 83
Chalmers, 52, 145
Chamberlain, 71
Changeux, 31
Chao, 109
Charles, 223
Chater, 185, 237, 238
cheating, 126
Cheek, 174
Cheney, 97, 122, 128
Chenn, 31, 91
Cheour, 87
Cherry, 7
Chi, 14
Chiarello, 84, 98
child care, 206
Child directed speech, 181
chimp, 48
alarm calls, 121, 123, 125
Broca, 115
communication in the wild, 127
comprehension, 134
critical period, 133
culture, 1
gaze following, 151
gestures, 123
hunting, 196
joint reference, 131
language, 129–140
language acquisition device, 133
larynx, 78
mathematics, 122
relation to humans and gorillas, 48
self-awareness, 153
sex life, 203
speech, 79, 130
symbol use, 131
syntax, 134
teaching by, 207
tool making, 198
Wernicke’s area, 99, 115
Chimpsky, Nim, 131, 139
Chinese Room, 147
Chiroptera, 44
Cho, 83
ching, 82
Chomsky, 3, 6, 8, 9, 11, 14, 23, 31, 39, 87, 100, 108, 131, 161, 162, 164, 171–173, 178, 179, 185, 186, 189, 193, 233, 235, 238
ape language, 129
resistance to evolution, 161, 162
Chowdhary, 83
Christiansen, 3, 18, 31, 84, 135, 180, 186, 233, 237, 238, 242
chromosome
X, 68
Y, 68
chromosome pattern, 49
chronospecies, 42
Churchill, 72, 200, 201
Churchland, 155
Ciochon, 48, 66
Clack, 86
clade, 48
Clahsen, 105
Clapham, 115
Clark, 18, 69, 91, 126, 144, 153, 174, 175, 183, 185
Clarke, 53, 58, 64
Claverie, 187
Clegg, 207
Clemmons, 119
Clever Hans effect, 131, 136
climatic variability
brain size, 97
CMP-sialic acid hydroxylase, 94
Coates, 89
Cocktail party syndrome, 105
Code, 98
Coffing, 63
Cogito, ergo sum, 145
cognitive linguistics, 9
cognitive reserve, 94
cognitive semantics, 9
Cohn, 96, 130, 135
Coleman, 23
Collado-Vides, 7
Collard, 57, 59, 60
Collins, 80
colorless green ideas, 9
Coltheart, 100
communication, 6, 120, 145
costly, 126
evolution, 125
honesty, 126
symbolic, 226
communicative intent, 122, 131
competence, 5
Conard, 72, 169
Condillac, 19, 158, 159, 174, 176, 191, 232
connectionism, 185
Connor, 97
Conroy, 51, 64, 66, 113
consciousness
hard problem, 145
construal, 8, 132, 224
conventionalized, 224
Conway, 84, 135
Cooke, 99
Coolidge, 73, 107, 169, 201
Cooper, 71, 150
Copland, 98
Coqueugniot, 207
Corballis, 18, 174, 175, 192, 201, 227
Corina, 84, 111
Corruccini, 49
Coski, 158
Cosmides, 100, 101
Cote, 48
Coupé, 74
Coward, 7
Cowen, 41, 86
Cowley, 120, 173
coyotes, 154
cranial base, 80
creationism, 14, 39, 42
Crelin, 79
creole, 37, 38, 179, 222
sign language, 174
critical period, 140, 187
apes, 133
Crockford, 121
Croft, 32, 33, 35, 37, 161
cross-fostering, 81
Crowther, 179, 181
cSányi, 17
cued representation, 7, 132, 224, 236
Cueto, 46–48, 73, 113
culture, 20
apes, 1
birds, 119
whales, 1
Cummins, 68, 101
cuozzo, 74
Currie, 95
curtiss, 101
cynx, 99
czikó, 31, 195, 202

dainton, 48
daly, 194
damasio, 91, 109, 156
dangerous things, see lakoff
d’Aouût, 54, 56
daroff, 108
dart, 58, 112, 196
darwin, 13, 18, 32, 36, 40, 41, 52, 160, 173, 174, 202, 206, 215
davidoff, 144
davidson, 73, 200, 215
davies, 72, 103
davis, 34, 109, 125, 220–222, 241
dawkins, 15, 22, 25, 27, 29, 33, 40
de boer, 221
de bonis, 52
de jong, 44
de la torre, 200
de pomerai, 88
de veer, 152, 153
de waal, 1, 97, 150, 203, 211, 227
deacon, x, 8, 30, 31, 37, 38, 77, 80–82, 91–95, 105, 106, 111, 112, 129–131, 133, 165, 171, 177, 180, 187, 189, 190, 198, 204, 218, 224, 226, 228
dean, 72, 76, 78, 114, 207
deepke, 1
deer
larynx, 80
deffeur, 73
deGraff, 179
defGusta, 74, 82
dehaene, 144
dehaene-Lambertz, 87
deinard, 48, 49
defour, 153
delisle, 30
DeLoache, 228
delson, 58, 64, 67
DeMarais, 35
demolin, 221
dennell, 65
dennett, 26–28, 32, 40, 145, 146, 155, 161
dent, 185
depaolis, 176, 228
depew, 30
dermoptera, 44
d’Errico, 60, 73, 74, 79, 168, 169, 199, 201
descartes, 145, 146
desmond, 84, 130
dessalles, 126, 212, 214, 224, 237
detached representation, 7, 132, 224, 236
deumert, 26, 28, 32, 147
development
plasticity, 111, 187
devolution, 30
deVore, 216
diamond, 201, 202, 218
diana monkeys, 124, 127
Dick, 175
diessel, 236
diller, 32
dilley, 146
ding-dong theory, 161
dinosaur extinction, 44
Dionysius Thrax, 8
dirlam, 28
dirven, 11
disotell, 68, 70
placed reference, 7
placement, 93
DNA, 26, 49
as a language, 7
louse, 70
mitochondria, 68
neanderthal, 70
Dobzhansky, 13
Dolgopyat, 110
dogs, 73, 150
dollo, 18
dolphin, 87, 120
brain, 91, 92
communication in the wild, 127
language, 136
self-awareness, 153
signature whistle, 87, 137
Domínguez Alonso, 89
Domínguez Rodrigo, 196–198, 208
Donald, 86, 156, 175, 201, 228
Dooley Collberg, 233
Doolittle, 35
Dor, 30, 164, 188
Dorit, 68
double standard, 140, 154
Douzery, 45
Drozd, 182
Dryer, 218
Dryopithecus, 48
dual causality, 146
dualism, 146
Duarte, 71
Dubois, 64
Duboule, 166
Duchaine, 100
Ducrocq, 45
Dufour, 111
Dunbar, 94, 95, 97, 122, 150, 195, 199, 202, 208, 210–214, 218
Dupanloup, 204
Durham, 30
Durrant, 166
Dyson, 23
ear, 86
Eastale, 44, 49
ecological inheritance, 21
economy, 1
Edelman, 31, 147, 155, 233
Edey, 57, 76
Edgar, 76
Edwards, 71
EEA, see Environment of Evolutionary Adaptedness
Einstein, 23, 33, 144
Eldredge, 13, 170
elegance, 107
elephant, 87, 98
Elliot, 181, 232
Elliot Smith, 41
Ellis, 185
Embick, 110
embodiment, 228
Emlen, 202
Enard, 104, 166
encephalization, 89
encephalization quotient, 94
endocast, 112
Enquist, 205
Environment of Evolutionary Adaptedness, 112, 194
Eosimias, 45
Epicurus, 158
Epstein, 101
EQ, see encephalization quotient
ERP, 88
Escalante, 68
Etler, 64, 67
Euarchonta, 44
Euclid, 22, 23
Evans, 121
Everson, 7, 18, 145, 158, 212
evolution, 13
as an abstract process, 14
bricolage, 16, 101, 166
complex features, 18
cultural, 20, 31
fallacies, 17–20
goal-directed, 20
good of the species, 17
irreducible complexity, 18
ladder, 19, 41
progress, 19
goal-directed, 20
horizontal transmission, 35
idiolects, 36
in ontogeny, 31, 39
limitations of, 16
multiregional, 67
mutations, 15
of idiolects, 35
of language, 32
of religion, 14
of scientific hypotheses, 23
progress, 19
resistance to, 39, 161
sexual, 202
social, 31
species, 63
speed of, 39
timescale, 39
universe, 14
evolutionary epistemology, 23, 32
evolvability, 173
exaptation, 18, 31, 163, 164, 166, 171, 231, 239
aquatic ape, 82
bipedalism, 175
brain, 95
Broca, 125
chewing, 82, 221
composite vocalizations, 221
comprehension, 19
conceptualization, 229
grammar, 234
hearing, 87
hierarchy, 238
image-schemas, 230, 231
in chimpanzees, 157
intelligence, 112
language, 164
larynx descent, 80
mimesis, 175
mirror neurons, 177
play, 214
sequencing, 85, 221
tool making, 202
Fabbro, 84
Faculty of Language
Broad sense, see FLB
Narrow sense, see FLN
Fagot, 229
Falguéres, 67
Falk, 79, 95, 113
false beliefs, 150
family resemblance, 9
Farrell, 1
Farroni, 149
Fauconnier, 9, 10, 229, 230
Faumuina, 79
Faupl, 69
Feinfeld, 149
Feldman, 68, 70, 74, 98, 185
Fellows, 144
Felsenfeld, 104
Fernández-Jalvo, 73, 197
Ferrari, 14, 151
ferrets, 111
fetal development
brain, 111
Fichtel, 121
Fiez, 84
Fieg, 112
figure/ground, 10
Finlay, 93
Finlayson, 72
fire, 65, 66, see dangerous things
Fischer, 86, 87, 120
Fischman, 68
Fish, 96
Fisher, 104
Fitch, 3, 6, 60, 78, 80, 81, 162, 235, 238
FLB, 6
Fleagle, 42
Flinn, 97
FLN, 6, 235
flying lemur, see Dermoptera
Fodor, 100, 161, 179, 181, 185
Fogassi, 176
Foley, 65, 67, 68, 76
footprints
fossil, 53
Formicola, 73
Fortey, 41
FORTRAN, 235
fossil
apes, 46–49
birds, 89
brain size, 112–114
DNA, 70
early hominids, 51–60, 82
early mammals, 42
early primates, 44–45
footprints, 53
formation, 42, 58
gaps in the record, 42
handedness, 115
hearing, 88
humans, 60–74, 82
in language, 166
species, 42, 63
symbol use, 168
Foster, 214
Fouts, 2, 129–131, 139, 142, 175
fovea, 45
Fowler, 86
Fox, 86, 111, 198, 203
FOXP2, 104, 172
autism, 148
Fragaszy, 21
frames, 10
Franciscus, 79
freckles, 30
Frith, 105, 148
Fritsch, 86
Froehlich, 71
frugivory, 45
Fruth, 203
Fu, 68
Fuentes, 47, 203
Fukuchi-Shimogori, 112
Fuller, 30
functionalism, 8, 9
Furness, 129
Futuyma, 40, 165
Gabora, 15, 21, 22, 26, 27, 201, 228
Gabunia, 64
Gagneux, 68, 203
Heeschen, 215
Heesy, 44, 198, 208, 209
Heim, 79, 84, 110
Heinrich, 28, 97
Helgen, 44
Heller, 228
Helmuth, 94, 110
Henderson, 168
Henshilwood, 168
Hepper, 99
Herbert, 49
Herder, 159, 160
heritability
language, 104, 105, 172
Herman, 130, 136, 139, 229
Hertrich, 84
Hewitt, 49, 68, 70
Heyes, 149, 153
Hickok, 84, 110, 111, 174
Hickson, 68
hierarchical, 233
Hilgeman, 29
Hill, 31, 185
Hillix, 56
Hirsch, 200
hobbits, see Homo floresiensis
Hoekstra, 134, 135
Honda, 86
Hopkins, 1, 99, 115, 123, 222
Horwitz, 89
Hosaka, 209
Höss, 71
Houghton, 79
Houle, 45
Hovers, 168
Hsu, 126
Huber, 177
Hulbin, 71, 72, 201
Hulsenbeck, 70
Huffman, 112, 196
Hull, 15, 26, 32, 35
Humboldt, 8, 119
Humle, 199
Hunt, 43, 45, 198
hunter-gatherers, 194
hunting, 85, 195
by apes, 196
of apes, 208
Hurford, 18, 149, 220, 228,
234, 235, 237, 239, 242
Hurley, 154
Hussey, 23
Hutslers, 99
Huxley, 23, 41
hyaena, 87
Hyatt, 1
Hyde, 215
hyoid, 78, 79
hypoglossal canal, 82
Iacoboni, 110, 176
iconicity, 137, 222, 233
monkey calls, 121
idiolect, 34, 35, 190
evolution, 36
IFG, see brain, inferior
frontal gyrus
Ik, 208
Ikegami, 239
image-schemas, 10, 230,
231, 237
Imenov, 70
imitation
vocal, 81
imitative learning, 176
Immodino-Yang, 98
immunology, 49
Indefrey, 110
indexical, 223
inferior frontal gyrus, see
brain
innateness, 178
modularity, 101
Insel, 91, 112
Insley, 87
intelligence
brain size, 94
gorilla, 135
social, 97, 208
verbal, 172
intentionality, 8, 122, 145,
223, 226
fossils, 212
interactor, 15
memetics, 27
Inuit, 195
Irish, 70
Irons, 194
Isaac, 98
Isbell, 54, 209, 216
isocortex, see brain, neocortex
Itakura, 131
Italian, 38
Itonen, 5, 224
Iverson, 175, 182
Ivy, 83
Izpisua Belmonte, 98
Jablonska, 30, 164, 188
Jablonski, 46
Jackendoff, 6, 9, 173, 178,
179, 187, 221, 224,
232, 235, 236, 239
Jacobson, 29
James, 23
Janik, 87, 120
Java man, see Homo erectus
Jaynes, 143
Jeannerod, 147
Jeafares, 212
Jenkins, 42
Jennions, 97
Jenny, 209
Jespersen, 215
Joanisse, 83, 103
Joffe, 206
Johanson, 57, 63, 76
Johansson, x, xi, 33, 36, 38,
164
Johnson, 10, 17, 228
Johnstone, 34
joint reference, 124, 131,
223
Jolly, 97
Jonker, 143
Jorde, 70
Jucquois, 2, 123, 222
Jusczyk, 87

Origins of language

Kaas, 86, 87, 89, 93, 115
Kacelnik, 111
Kaessmann, 49, 68
Kahn, 70
Kaminski, 150
Kant, 13, 22, 223
Kanzi, 132–135, 137, 138, 140, 220, 226–228
Kaplan, 104, 115, 195
Kappeler, 208, 210
Kappelman, 57, 73, 113
Karmiloff-Smith, 101, 105
Karten, 89, 91, 92
Kattman, 326
Kave, 103, 105, 107
Kay, 44–46, 82
Ke, 70
Keil, 144
Keith, 41, 67
Keller, 32
Kelley, 46
Kellogg, 129
Kenyapithecus, 48, 49
Kenyanthropus platyops, 58, 61
Keno, 70
Kellman, 79
Kem, 144
Khoratpithecus, 47
Kidd, 48
Kim, 108
Kimbell, 60, 63, 113, 200
Kimura, 200
King, 81, 111, 128, 130, 133, 167, 199, 207, 227, 232, 295, 299
Kiple, 30
Kipling, 126
Kirby, 3, 18, 31, 180, 186, 190, 235, 239, 242
Kircher, 153
Kirk, 44, 45
Kirwan, 223
Kissinger, 202
Kitchen, 152
Kitten
Chomsky’s, 178, 189
Koko’s, 189
Kittler, 49, 70
Kitzing, x
Kivisild, 68
Klein, 71, 76, 109, 199, 200, 239
Knight, 71, 177, 214, 242
knuckle walking, 48, 54
Knudsen, 112
Koch, 154, 155
Koenig, 209
Koeslag, 210
Köhler, 46–48
Koko, 135, 153, 189
Komarova, 231
Kooij, 134, 135
Kordos, 48
Korey, 63
Kosso, 108
Koufos, 52
Krakauer, 86, 221, 227
Krakow, 145, 146
Krause, 208
Kraytsberg, 68
Kreis, 160
Kreiter, 86
Kullmer, 59
Kuper, 29
Kusakawa, 86
La Mettrire, 129
Lacopithecus, 47
Lachapelle, 30
Lachmann, 126
lactase, 30
lactose intolerance, 30
Ladefoged, 83
Lagar Velho child, 71
Lai, 83, 104
Laitman, 79
Lakoff, 9, 10, 228–230, 237, see women
Laland, 21, 31, 39, 94, 97
Lamarck, 16, 23, 27, 30, 36, 39
Lamb, 5, 9
lamb, 117
Lana, 131
Landau, 41, 195
Langacker, 9, 121, 165, 229, 233, 237
Langdon, 82, 215
language, 5–10
acquisition, see acquisition
and thought, 144
as adaptation, 163
complexity, 164
diachronic change, 38
discourse, 7, 221
DNA, 7
dolphin, 120
dynamic, 9
early, 167
genes, 104
handedness, 99
heritability, 105, 172
late, 167
learnability selection, 37
memetics, 31
mind, 144
modularity, 101
normative, 5
right hemisphere, 98
semiosis, 8
social, 5
syntax, 9
teaching to non-humans, 129–140
uniqueness, 2, 7
universals, 179
written, 174
language evolution
memetic, 32
language, 5
LaPolla, 233
Larick, 66
laryngeal air sacs, 80
larynx, 78
adaptation for speech, 80
bipedalism, 80
chimpanzee, 78
descent
ontogeny, 78
phylogeny, 79, 80, 86
dry habitat, 80
exaptation, 80
sexual selection, 80
Lasnik, 39, 179, 181
lateralization, 98, 109
birdsong, 99
fossil, 115
non-human, 115
ontogeny, 98
phylogeny, 99
sign language, 99
vocalization, 99
Latin, 38
Lau, 154
Launey, 8, 9
Laurence, 179
Le May, 79
Le Prell, 87
Leakey, 53, 57, 58, 60, 64, 216
learnability
selection for, 38
Leavens, 99, 123, 222
Lebel, 74
Lee, 149, 195
Lee-Thorp, 54, 197, 216
Legate, 179, 181
Leibniz, 158
Leigh, 93, 206
lemur, 45
Lennie, 94
Lenski, 19
Leonard, 54, 72
Levin, 98
Levinson, 144
Levy, 101, 103, 105, 107
Lewin, 42, 48, 52, 54, 58, 60, 64, 76,
Index


Lewis, 27

Lewontin, 18, 163

Li, 6, 94, 96, 119, 166–168, 171, 107, 110, 117, 188, 193

Libet, 147

lice, 70

Lidz, 179, 181

Lieberman, 58, 63, 71, 78–80, 84, 89, 106, 107, 110, 117, 188, 193

Liégeois, 104

Lightfoot, 2, 165, 170

Liljenström, 146

Limber, 129, 159, 160

Lin, 68

Ling, 135, 139, 226

MacA-Meyer, 70

Macchiarelli, 53

MacDonald, 178, 182, 183, 238

Mace, 21

Machiavellian intelligence, 97, 208, 211, 212

Macho, 48

MacLarnon, 82, 204, 221

MacLachy, 46, 53

MacLeod, 91, 92

MacNeileage, 82, 125, 175, 220–222, 241

MacWhinney, 37, 185

Maddox, 72

Maestripiere, 1

Majoros, 28

Maliukova, 122

Mallez, 65

Malmberg, 209

Mameli, 21, 31, 100, 101

mammals

origin of, 43

placental, 44

transitional fossils, 43

Mandler, 10, 144, 185, 230

Manley, 86

Mann, 1, 46, 48

Manning, 185, 206

Manser, 121, 122

Manzi, 65

Maratos, 185

Marchant, 115

Marcus, 103, 104, 112, 166, 181, 186, 187, 230, 238

Marean, 168, 198

Marentette, 174

Margolis, 179

Marino, 92, 153

Markson, 111

Marlar, 73

Marler, 119–121

Marlowe, 202

Marsden, 29

Marten, 153

Marth, 68

Martin, 45, 91, 95, 96, 109, 124

Martinet, 7

Matata, 132

mathematics

chimpanzee, 122

Matsumoto-Oda, 209

Matsuzawa, 122, 199

Maturana, 143

Mauk, 84

Maupertuis, 13

Maurus, 86

McArthur, 7, 8, 228

McBrearty, 67, 69, 73, 168, 169

McCrone, 78, 79

McCormick, 58

McCormack, 87, 98

McCowan, 87, 120

McCrossin, 46

McElreath, 28, 97

McGee, 86

McGrew, 54, 115, 198

McHenry, 49, 53, 54, 58, 59, 63

McKee, 67, 76

McKinney, 92

McLennan, 49

McPherron, 200

Meaburn, 104

Meagher, 121

meaning, 120, 223

medial prefrontal cortex, see brain

MEG, 88

Melichar, 82

Mellars, 168, 201

meme, 25

language as a, 35

semantics, 27

virus, 27

world 3, 26

memetics, 25

interactor, 27

language, 31, 32

replicator, 27

Meng, 86, 105

Menon, 89

mental events, 145

mental maps, 230

mental spaces, 10, 230, 231

Menzel, 80, 132, 143, 226

Menzies, 146

Mercader, 199

Mercado, 229

Mercola, 98

Merzenich, 111

metacognition, 155

metaphor, 10, 229, 230

metonymy, 10

Meyer, 2

Miera, 107

Mietto, 53

Mikkölä, 133, 150, 174

Milberg, 108, 110

Miles, 130

Milinkovitch, 48

Miller, 45, 115, 202

Mills, 174

Milo, 198

mimesis, 159, 175–176, 215, 223, 225, 226, 228, 232

memetics, see mimesis

mind

and memory, 153

animal, 122

epiphenomenon, 145, 146

hard problem, 145

language, 144

robots, 24

spandrel, 147

theory of, 8, 148–151
autism, 148
def children, 148
Williams syndrome, 148
mind-body dualism, 146
Mindell, 2
minimalist program, 39, 187
Mintz, 185, 186
Mirazon Lahr, 65, 68
mirror neurons, 176
mirror test, 152
Mitani, 120, 196, 197, 211
Mitchell, 153
Mithen, 65
mitochondrial DNA, 68
Mitochondrial Eve, 68
Miyamoto, 44
modalities, 173
Model/Rival paradigm, 137
modularity, 100
Big Modules, 101
conclusions, 111
evolution, 101
innateness, 101
small modules, 101
Mogie, 19
Molnar, 54
Molotova, 122
Monboddo, 159, 160
monkeys
alarm calls, 8, 124
awareness of memories, 153
causal reasoning, 124
gaze following, 151
New World, 45
Old World, 46
origin of, 45
syntax, 127
Montgomery, 149
Monuki, 112
Mooneyham, 143
Moore, 88, 91
Morell, 58, 72
Morford, 7
Morgan, 82, 215
Morotopithecus, 46
morphemes, 33
morphology, 239
Morris, 14, 39, 99, 115, 186
Mortley, 241
Morwood, 65, 74
motherese, see Child directed speech
motor theory, 221
Mousterian, 72, 200
Moya-Sola, 46–48
MP, see minimalist program
MRI, 88
Mueller, 1, 39, 82, 86, 87, 98, 99, 106, 161
174, 186, 187
Mufwene, 34–36, 179
Muller, 98, 110, 159, 160, 169
multiregional evolution, 67
Mungo Man, 71
Murphy, 44, 189
music, 214

cosmic, 74
language origins, 214
Neanderthals, 74
Musso, 110
mutation, 15
Mylander, 174, 222
Nadler, 203
Nagel, 145
Nagell, 1, 177
Nakahara, 111
Nakamichi, 198
names, 228
Nanay, 15
narrativity, 215, 230
Natsopoulos, 99
natural selection, 13
Neanderthals, 66, 82, 178, 195
art, 169
brain, 113
DNA, 70
extinction of, 72
hearing, 88
hybrids, 71
language, 246
music, 74
religion, 73
speech, 79, 85, 167
tools, 200
vocal tract, 79
Neergaard, 99
Neff, 204
negative evidence in acquisition, 181
neither nature nor nurture, 180
Nekaris, 45
Nelson, 155
neocortex, see brain
neoteny, 92, 94
Nesse, 194
Nettle, 37
Neufeld, 51
neural control
breathing, 82, 221
jaws, 221
sequencing, 83
tongue, 82
neural events, 145
neuroimaging, 108
Broca, 110
semantics, 109
syntax, 110
Neville, 174
New World monkeys, 45
Newbury, 104, 148
Newmeyer, 2, 6, 162, 166, 181
Newport, 38, 176
Ng, 89
Ni, 45
Nicastro, 126
Nielsen, 89
Niewoehner, 73
Nilsson, 94
Nim Chimpsky, 131, 139
Nimchinsky, 94
nipples, 163
Nishikawa, 89, 91
Nishikimi, 189
Nishimura, 78, 174
Noad, 119
Noble, 125, 126, 215
Novacek, 44
Nowak, 20, 86, 125, 221, 227, 230, 231, 239
Nuckolls, 121
Oakley, 199
objective knowledge, 25
occipital lobe, see brain
O’Connell, 196, 197
Oden, 229
Odling-Smee, 21
Ohta, 189
Okanoya, 119, 206
Old World monkeys, 46
Oldowan, 65, 200
Olendzenski, 35
Oliphant, 226, 231
Oliwenstein, 53
Olson, 88
O’Malley, 198
omonomy, 45
Oms, 64
O’Neil, 208
Onion, 74
onoma, 8
ontogeny, 102
opposable thumb, 47, 60, 200
optimality, 185
orangutan
culture, 1
hunting, 196
language, 129
origins, 47
sex life, 203
Oreopithecus, 47
Origin of Species, 13
Ornelas, 30
O’Rourke, 71
Orrorin tugensis, 52, 61
Osborn, 41
Osterhout, 89
Otavipithecus, 48
Ott, 83
Otte, 201
Index 341

Ouranopithecus, 52
Out of Africa, 67, 69
Ovchinnikov, 71
Owens, 179
Oxnard, 91

Pääbo, 49, 68, 70, 187
Paciulli, 115
Pagel, 21
Paging, 149
Palmer, 99, 115
Palmer-Brown, 186
Palmour, 104, 105
Pan paniscus, see bonobo
Pan troglodytes, see chimpanzee
Pandich, 103
Papineau, 145
Paradis, 108
Paranthropus, see Australopithecus
parasites
idiolects, 36
lice, 49, 70
worms, 49
parentese, see Child directed speech
Pares, 64
Paris Linguistic Society of, 2
Parisi, 185
Parisse, 148
Pariser, 200
Parks, 70
Parnell, 203
parole, 5
parrots
language, 137
Partridge, 58
Pascual-Leone, 154, 176
Passingham, 84
Paterson, 105
Patterson, 96, 130, 135, 169, 189
Pauling, 49
Pavesi, 70
Pearl, 80
Pearson, 70
Pedersen, 233
Pei, 160
Peirce, 25, 223
Peking man, see Homo erectus
Pendleton, 89
Penin, 92
Pennisi, 68, 70, 93, 95, 150
Pennock, 23
Pepperberg, 130, 137, 139, 140, 142, 213, 229
Pepys, 129
perceptual meaning analysis, 230
Peretz, 215
performance, 5
perissodactyls, 44
PET, 88
Peters, 198
Peterson, 148
Petersson, 110, 183
Pettertto, 101, 174, 184
phenotype, 15
Phoenix, 136
phonology, 221
Piaget, 23, 31
Piattelli-Palmarini, 23, 161
Pikford, 48, 52, 53, 69
pigdin, 37, 38, 239
sign language, 174
Pietrandrea, 222
Pietroski, 23–26, 32, 36, 139, 224
world 1/2/3, 24
Piemont, 125
Pinker, 2, 3, 5, 6, 8, 39, 80, 88, 103–105, 108, 138, 139, 141, 144, 161, 162, 164, 166, 172, 173, 178, 181, 182, 192, 193, 196, 202, 211, 216, 228, 232
Pinto, 22
Pithecanthropus erectus, 62
Place, 146
placental, 44
plasticity, 111, 134, 187
Plato, 7, 24, 143
platyrrhine, 45
Plecan, 204
Plotkin, 221
Plummer, 196, 197
Plunkett, 101, 185
Poelarends, 35
Poeppel, 98, 100, 110, 111
Poltro, 233
Polygyn, 204
Pongidae, 51
Ponter, 2
Ponter, 54
pooh-pooh theory, 160
Poule, 87
Popper, 23–26, 32, 33, 36, 139, 224
Poverty of the stimulus, 179
Povinelli, 123, 151–155
Praeanthropus africans, see Australopithecus afarensis
Prat, 63
preadaptation, 18
precision grip, 200
Australopithecus robustus, 60
Oreopithecus, 47
predation, 208
predication, 8, 224, 237
pro, 237
prefrontal lobes, see brain
Prell, 87
Premack, 8, 148, 149, 175, 207
Price, 17, 109
primates, 44
alarm calls, 120, 121
lesion studies, 125
origin of, 44
social calls, 122
vision, 45
Prince, 185
Pritichard, 70
Proconsul, 46
protein sequence, 49
proto-speech, 85
proto-syntax, 230
protosyntax, 172
prototype, 9
Pullum, 178, 179, 181, 183
Pulvermüller, 117
Pütz, 144
pyramidal neurons, 99
qualia, 145
quantum physics, 26
quarks, 112
Quiatt, 198
Quinney, 57
Radick, 19, 32, 98, 174
Rae, 46
Rankir, 95, 113, 164, 195–197, 214
Ragsdale, 112
Raichle, 94
Rakic, 112
Ramapithecus, 47
Ramirez Rozzi, 207
Ramsdell, 98
Ramus, 87
Ran, 145
Rancourt, 171
Randerson, 80
Rasmussen, 45
Reader, 94, 97
Reaney, 207
Reby, 80
Recer, 168
Tattersall, 58, 64, 66, 69, 71, 76, 79, 80
Taung child, 58, see Australopithecus africanus
Tavaré, 44
taxonomy, 42, 51
splitting, 63
Taylor, 9, 154, 158
Taylor-Parker, 153
teacher talk, 182
teaching, 207, 208
teleology, 14
Templeton, 69, 70
Teotonio, 18
Terrace, 122, 130, 131, 135
territoriality, 209
Tesar, 185
Théoret, 176
theory of mind, 8, see mind, theory of
Thieme, 85
Thier, 154
Thomas, 179
Thompson, 147, 229
Thompson-Schill, 109
Thomson, 70
Thorne, 66
Thornhill, 19, 39, 202
Thornton, 179
throwing, 83
sex differences, 85
Tilkens, 72
Tillier, 79
Tishkoff, 70
Tobach, 153
Tobias, 53, 63
Tocheri, 200
Tolkien, 65
Tomasellos, 21, 28, 38, 97, 123, 149–152, 154, 175, 177, 180, 185, 207, 213, 227, 228, 230, 232, 236, 242
Tomblin, 103
tongue, 81
neural control, 82
Tononi, 94
Tooby, 100, 216
tools, 1, 63, 198
Acheulean, 65
animal use of, 198
Australopithecus, 200
brain size, 97
Mousterian, 72, 200
Oldowan, 65, 200
Toro, 87
Toth, 52, 199
Toulmin, 23
Tower of Babel, 161
Trabant, 2
Trace Deletion hypothesis, 107
trade, 1
trajector/landmark, 10
tree shrews, see Scandentia
Treisman, 67
Treves, 210
Trinkaus, 67, 71, 73, 74, 76
Trombetti, 161
Tsonis, 7
Turing test, 148
Turnbull, 208
Turner, 10, 229, 230
Turney, 30
Turtle dove, 65
Tyack, 87, 97
Tyrrell, 71
Uddin, 49
Uemliain, 162
Ungulapithecus, 46
Ujheelyi, 120, 127, 221
Ungar, 52, 54
Ungerer, 9, 10, 229
universal grammar, 9, 172, 179
universe
elegance, 108
evolution, 14
Uriagereka, 100, 108, 179, 181, 186, 231
Ussery, 19, 39
Utumi, 196
utterance selection, 33
Valian, 185
van den Bos, 152, 153
van der Hulst, 174
van der Lely, 103, 104
van der Meer, 19
van Hoof, 196
Van Peer, 168
van Schaik, 1, 198, 199, 203, 208
Van Valin, 233
Vancassell, 39
Vanderwolf, 146, 147
Vanechouette, 85, 155, 214, 215, 238
Varki, 88
Vea, 123, 222
vehicle, 15
Vehrencamp, 177
Vekua, 64
velopharyngeal closure, 79
Verhaegen, 82, 215
Vierspoo, 11, 144
Vertut, 169, 170
vervret monkey, 120, 225
alarm calls, 120
vestigial features, 163
Victoriapithecus, 46
Videan, 54
Vigilant, 203
Vigliocco, 98, 107, 110
Vignaud, 52
Vihman, 176, 228
Viki, 130
Vila, 72
vitamin C, 189
vocal chords, 81
vocal tract, 78, 86
apes, 79
Neanderthals, 79
vocalization, 77, 81
Voelkl, 177
Vogel, 1, 54, 199, 216
Vollmer, 22, 23
von Humboldt, 8, 119
Vonk, 151, 154, 229
Vrba, 18
Waddington, 30
Wagner, 182, 184
Wakefield, 3, 84, 85, 115, 222
Walczak, 185
Walker, 57, 58, 64, 82, 115
Wallace, 13
Wallman, 142
Walsh, 31, 91, 98, 112, 154
Wang, 35, 54, 87
Wanner, 179, 232
var, 2
Ward, 47, 57, 71, 204
Washburn, 156
Washoe, 130
Wassink, 148
Waters, 175
Watkins, 104
Watson, 83, 85
Watts, 2, 196, 197, 211
Watte, 150
Watwood, 1
Wawood, 1
Wernick, 5
Weinreich, 106
Weiss, 14, 46, 48
Wells, 158–160, 174, 192
Wermter, 186
Wernicke, 100
Westergaard, 1, 83, 115, 199
whale song, 119
whales
culture, 1
self-awareness, 153
Whaling, 119
Wheeler, 101, 111
White, 25, 48, 53, 63, 64, 69, 119, 205
Whitehead, 1, 119, 252, 269, 296
Whiten, 1, 97, 151, 177, 198, 208
Whitfield, 52
Whitham, 1
Whitmeyer, 39
Wible, 43
Wiener, 119
Wilcox, 121
Wildgen, 97, 167, 202, 229
Wildman, 51
Wilkins, 3, 13, 15, 16, 23, 26, 28, 32, 35, 39, 84, 85, 115, 166, 222
Williams, 40, 49, 105, 194, 209
Williams syndrome, 101, 105, 148
acquisition, 105
mathematics, 105
Wilson, 2, 17, 194
Wimsatt, 16, 22, 26, 28, 36, 39
Wind, 220
Winter, 228
Wirtz, 35
Wittgenstein, 9, 26, 232
WoldeGabriel, 216
Wolf, 45
Wolpoff, 51, 52, 67–69, 71
wolves, 150, 154
Womble, 186
women, see fire
Wong, 111, 187
Wood, 42, 48, 52, 53, 57, 59, 60, 63, 67, 68
Wood Jones, 41
Woodruff, 8, 148, 149
Worch, 199
Worden, 33, 39, 187, 231, 238, 244
world 1/2/3, see Popper
Wozniak, 23
Wrangham, 2, 52, 54, 197, 199, 204, 211
Wray, 86, 231, 234
writing, 174
Wu, 66–68
Wuetrich, 66
Wuketits, 20, 165
Wundt, 160
Wyman, 198, 199
Wynn, 54, 64, 73, 85, 107, 169, 200, 201, 212, 257, 287, 304, 316, 318
Xitco, 123, 136
Y-chromosome Adam, 68
Yamei, 66
Yang, 179, 181, 186, 188
Yellen, 168
Yerkes, 129
yeti, see Gigantopithecus
yo-he-ho theory, 161
Yost, 98
Young, 46, 49, 52, 54, 209, 216
Yurk, 119
Zemel, 209
Zhang, 74
Zietkiewicz, 68
Zilhão, 71
Zimmer, 148
Zinjanthropus, see Australopithecus boisei
Zipf, 120
Zlatev, x, 24, 26, 33, 143, 144, 149, 176, 225–228, 232
Zoloth, 86
Zonneveld, 86
Zuberbühler, 121, 122, 124–127, 209
Zuckerkandl, 49