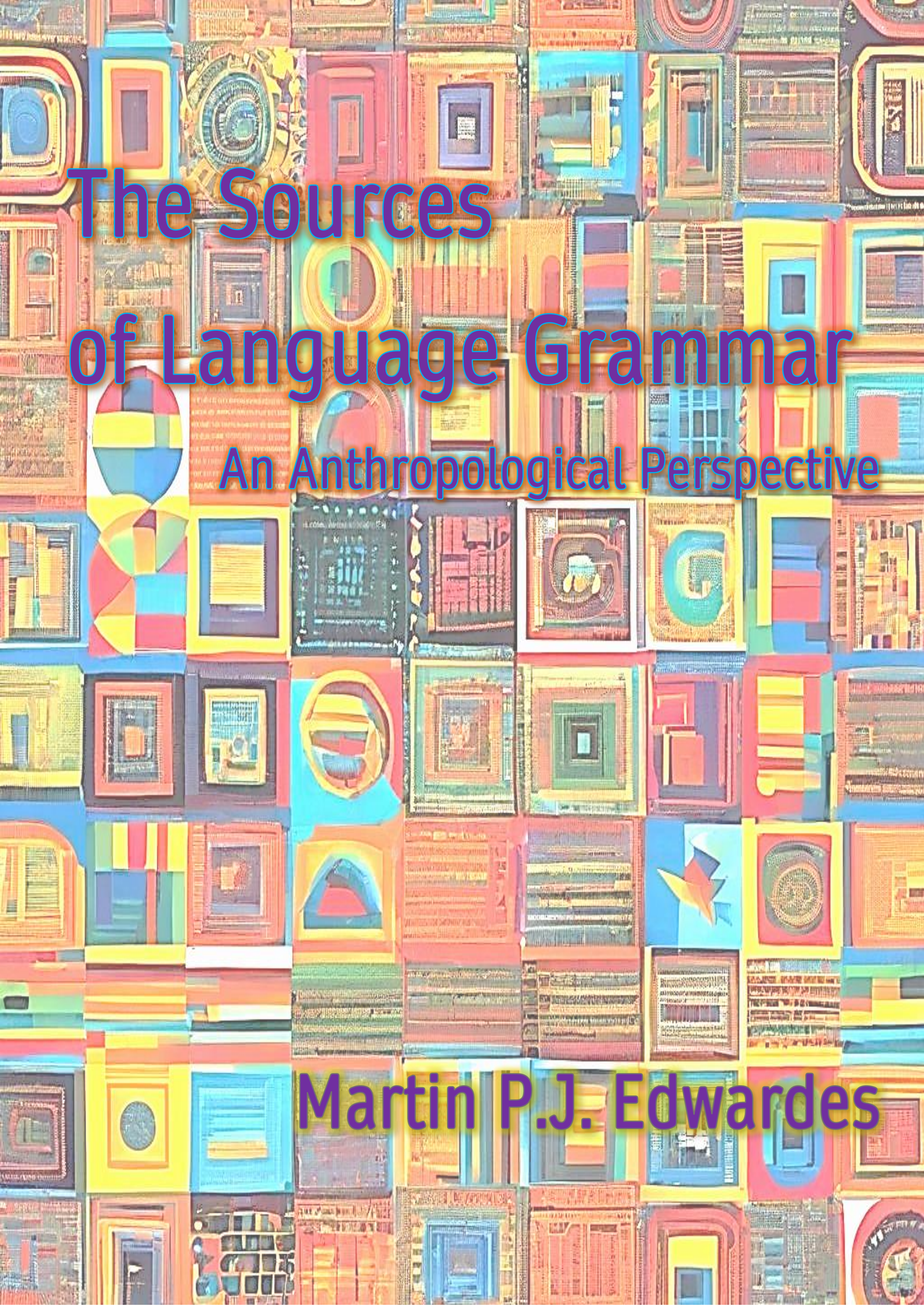

The Sources of Language Grammar

An Anthropological Perspective

Martin P.J. Edwardes

The Sources of Language Grammar

An Anthropological Perspective

Martin P.J. Edwardes

SCITSIUGNIL
PRESS

The Sources of Language Grammar
An Anthropological Perspective

Martin P.J. Edwardes

Published by Scitsiugnil Press
1 Maiden Road, London, UK, E15 4EZ
And produced in the UK

Product Design by Martin Edwardes
Scitsiugnil Press logo design by Jack Bruce
Cover illustration designed by Gencraft AI

ISBN: 978-1-9999369-4-5

Acknowledgements

Many thanks to all the people who have made this book possible. Especial thanks go to the Linguistics community at Kings College London, and to the members of the EAORC (Evolutionary Anthropology Online Research Community), who have inspired me to keep researching into retirement. What started as a passion became a paid pastime and now a retirement hobby – but never an imposition.

I must also thank my PhD tutors, Chris Knight, Camilla Power and Tom Dickins, who started me out on this search for the origins of our humanness. None of us knew where this project would go, but you had faith in my ability to make the journey, even when my faith in myself faltered.

Particular thanks are owed to Philip Rescorla, my life-partner, who has encouraged me and given me intellectual and emotional support through thick and thin. I'm sorry for all the times I was thick-headed and thin-skinned. Life with you has been golden; and now, gold.

The Sources of Language Grammar: An Anthropological Perspective

Martin P.J. Edwardes

Contents

Figure Contents

Prologue Language Grammar: Becoming Human

> ***If the human brain were so simple that we could understand it, we would be so simple that we couldn't.***
> *(George Edgin Pugh, quoting his father, Emerson M. Pugh)*

The quote above raises what Lyall Watson called "the Catch-22 of the biology of consciousness": to be able to explain itself, a brain needs to be bigger than it is. Fortunately, the human brain is capable of three tricks which circumvent, or at least mitigate, the impasse. First is aggregation: I need not know all the mechanisms of the human brain individually if similar mechanisms can be treated the same; understanding one is sufficient to understand them all. Second is multi-brain storage: I need not know everything myself if you and I are willing to share our knowledge. Third is offline storage: we need not keep our knowledge in our collective online memory if we can commit it to offline storage – write or draw it on a rock, or paper, or a computer. The first of these tricks, aggregation, is a cognitive skill, and probably not limited to hominins (the *Ardipithecus*, *Australopithecus*, *Paranthropus* and *Homo* clades). The last two tricks, online and offline storage, rely on a complex communication system which allows us to exchange a wide range of knowledge quickly and easily. We know this complex communication system as language; and the rules that drive the system we call grammar.

Fourteen years ago, I wrote *The Origins of Grammar: An anthropological perspective* (Edwardes, 2010) to investigate one question about language origins: where do the rules governing language come from? Language origins is itself only one aspect of a wider project to understand the origins of human behaviour; and this, in turn, feeds into the general investigation of *what it means to be human*. Anthropology is what we call the study of what it means to be human, so the anthropological perspective adopted in *The Origins of Grammar* endeavoured to identify the role language grammar plays in the story of being human. The objective for this book remains the same.

Since publication of *The Origins of Grammar*, our understanding of the role of language in human origins has changed considerably. Which means that revisiting the sources of language grammar is a task worth undertaking. Humans are still humans, and the accepted parameters of being human have shifted only fractionally; and language grammar as a cognitive capacity has not altered significantly in millennia – how humans use grammar within language seems unchanged since the beginning of recorded history (which, by definition, requires writing, so since at least 11kya). What has changed during the last fourteen years is our knowledge of the evolution of hominins into modern humans, and this in turn has affected our understanding of how language grammar developed. While structurally based on *The Origins of Grammar*, the content of this book is considerably different, reflecting the changes in our knowledge.

What we know of the sources and development of human language grammar relies mostly on the archaeological record from the Palaeolithic, which started about 3mya (million years ago) and is seen as fading away 12kya (thousand years ago). Recorded language, or writing, is a recent innovation in the story of being human: traditionally we see writing beginning after 12kya, although some archaeologists argue convincingly that our earliest rock art (more than 40kya) contains symbols which seem to carry information in a similar way to writing (Petzinger & Nowell, 2011). If we consider just the time since *Homo sapiens* – our species –began (at least 230kya according to recent redating of the oldest *Homo sapiens* fossils (Vidal *et al*., 2022), and probably closer to 300kya), then human communication, in a grammatically complex form most linguists would recognise as language, probably existed from at least 100kya, long before recorded history began. Yet for most of this period all we have is the Palaeolithic fossil record so far discovered, and from this we can draw implications, but not conclusions, about our ancestors' use of language.

The fossil record is, at best, an indirect marker of language because, as Fitch (2010) said, "language does not fossilise". Nonetheless, in the fourteen years since *The Origins of Grammar*, the fossil record has been enriched by the discovery of several new species of human. For instance, *Homo floresiensis*, also known as *the hobbit* because of its small stature, was identified in 2004 (Brown *et al*., 2004); but its role in the evolution of grammar-using humans still remains largely unknown. As an isolated species probably more closely related to *Homo erectus* than to *Homo sapiens*, it is likely to be, at best, only peripherally relevant to the sources of current language grammar.

Some archaeological finds have been genetic, as our ability to sample and analyse DNA from very old bones has recently made a major leap forward (Green *et al*., 2006). In 2013, Svante Pääbo's team at the Max Planck Institute for Evolutionary Anthropology showed that a single finger-bone found in the Denisova Cave in Russia belonged genetically to a new genus of human: not neanderthal, not sapiens, but closely related to both (Meyer *et al*., 2014). *Homo denisova*, to use its unofficial designation, introduces new complexity to the question whether truly complex language is available only to *Homo sapiens*, or was also used by *Homo neanderthalensis* and possibly other species of *Homo*.

In 2015, Berger's team described another new type of human, which they had discovered in the Dinaledi Chamber of the Rising Star cave system in South Africa (Berger *et al*., 2015). The fossils are approximately dated between 330 to 230kya, which makes them contemporaneous with the beginnings of the *Homo sapiens* species; but they had quite a different physical form. The find has provided an extensive range of fossils from at least 15 different individuals, with several almost-complete skeletons. The new species, labelled *Homo naledi*, had a cranial capacity similar to Australopithecines; it was better adapted to living in trees than *Homo sapiens*, and less able to carry out pursuit hunting (chasing prey until it is exhausted rather than fighting it to death). Yet the discovered fossils were from bodies that seem to have been carefully carried into the Dinaledi cave after death, which may indicate that this mysterious species had a cultural tradition of funerary rites. The commemoration of death could indicate a symbolic culture and the use of a *Homo sapiens*-like language – which, in turn, would have a major effect on our understanding of language and cognition in the *Homo* clade. However, while *Homo naledi* provides an intriguing alternative view on the development of Human language grammar, its story as a species is still at an early stage.

Another mystery surrounds the Dmanisi hominins, discovered in Southern Georgia in the Caucasus and first described in 1995 (Gabunia & Vekua, 1995). Their status, and even species, remains undecided – although the early date so far attributed to the fossils (1.8mya) means that they are unlikely to be of any significance in the sources of language grammar; so, from the viewpoint of this book, someone else's problem.

Two recent discoveries have tentatively added two new species to the *Homo* family tree. *Homo luzonensis* was discovered in 2007 in the Philippines and was initially treated as a modern human. However, further discoveries led to it being allocated its own species label (Détroit *et al*., 2019). We now have fossil remains from several *Homo luzonensis* individuals, and even signs of human butchery of a rhinoceros, which together indicate that there was a thriving population from at least 600kya to at least 100kya. What is remarkable about this species is that the only way they could have reached Luzon – an island – was by completing a sea voyage from mainland Asia. This would have required a level of technology we had previously assumed to be modern, available only to *Homo sapiens*.

The second recent discovery is *Homo longi*. Discovered in Harbin in China 1933, the almost-complete fossil skull was hidden in a dry well for 85 years before being brought to the attention of palaeoanthropologists in 2018 (Ni *et al*., 2021). There is some dispute about whether the skull represents a separate species or is that of a Denisovan or Dalian (*Homo daliensis* is another under-represented and under-described fossil species from China); and the long gap between discovery and analysis means that the skull has only an approximate date, between 300kya and 140kya, with the best guess being about 146kya.

Did *Homo luzonensis* and *Homo longi* have language? The technical skill required to reach Luzon would indicate a sophisticated culture able to undertake sea travel, so it would therefore be excessively cautious (or species-centric) to deny *Homo luzonensis* a language-like communication system; but we can make no useful judgement on *Homo longi*. Also, while it is not unreasonable to believe that any *Homo* species which existed after 300kya probably had some facility in sharing complex cognition, we cannot be sure they needed something like modern language grammar for that sharing, or even whether they had a rule-driven vocal communication system. The only grammatical language to have survived since 50kya is that of *Homo sapiens*, so this book focuses its attention on the sources of language grammar used by *Homo sapiens*.

Even our most recent forebear species has been questioned. *Homo heidelbergensis* was seen as the most likely precursor species for *Homo sapiens*, but new evidence has called for a re-examination (Wood & Grabowski, 2015). It is now more common to view *Homo heidelbergensis* as a precursor to *Homo neanderthalensis*, with *Homo sapiens* coming from different stock. A recent paper suggested a fossil discovered at Bodo D'ar in Ethiopia, previously designated as *Homo heidelbergensis* but now described as *Homo bodoensis* (Roksandic *et al*., 2021), could be the forebear species. However, evidence of interbreeding between *Homo sapiens* and *Homo neanderthalensis* has further complicated the issue: interbreeding has been significant enough to ensure that many modern humans have a sizeable chunk of Neanderthal and Denisovan DNA in their genetic make-up (Sankararaman *et al*., 2016). What we treat as a family of *Homo* species may be, in purely genetic terms, a family of sub-species in a single species.

Our knowledge of being human has improved in several other areas over the past fourteen years. Our understanding of the significance of cultural artefacts has improved as more artefacts have been found, and artefacts already known have been reinterpreted. For instance, the discovery in 2020 of an apparently-carved deer bone at Einhornhöhle in northern Germany, dated to 51kya – before *Homo sapiens* arrived in the area – shows that Neanderthals may well have understood and used symbolic reference, a prerequisite for language (Leder *et al*., 2021). By itself this is of marginal interest here because Neanderthals, despite the interbreeding, are not a *Homo sapiens* parent species; but it indicates that our last common ancestor with Neanderthals, about 800kya, may have been capable of symbolic reference.

Our knowledge of nonhuman cultural artefacts has also improved: contrary to interpretations accepted fourteen years ago, many nonhuman species seem to have local cultures. For instance, observing chimpanzee honey-dipping (stealing honey from bee colonies) in Guinea-Bissau has shown different techniques being used by different local groups (Bessa *et al*., 2021). Despite regular relocation of young females between groups, each local technique seems to remain stable. Harrod (2014) has even controversially suggested that behaviours around births, deaths and consortship may indicate a proto-religious culture among chimpanzees (*Pan troglodytes*).

Physical features of early humans have also been reinterpreted as our understanding of palaeoanatomy has improved. For instance, the musculature of the *Australopithecus sediba* thumb has been shown to be compatible with a precision grip (holding an object between thumb and fingers: Kivell *et al*., 2011). The precision grip, with its implications for tool use and manufacture, had previously been seen as exclusive to the *Homo* clade. Cebeiro & Key (2023) have recently shown that modern bonobos (*Pan paniscus*) can produce a precision grip between their thumb and the side of their hand, suggesting that a precision grip was possible even earlier than *Australopithecus*. Pursuit hunting (tracking and pursuing prey over long periods of time until it is exhausted), long known to be a feature of *Homo sapiens*, has now thought to be within the capabilities of earlier *Homo* species, indicating they had capacities to plan and delay gratification (Lieberman *et al.*, 2009). Some forms of planning and delayed gratification are now known to be present in chimpanzee cognition, indicating they may also have been present in the last common *Pan-Homo* ancestor (Cheney, 2011). Perhaps language grammar – or, at least, the precursor – was not limited to *Homo sapiens*.

Defining Language Grammar

Grammar is a term with a wide range of uses and meanings. In the UK, one of the primary meanings of the word *grammar* is to describe a type of school, supposedly teaching a more academic syllabus for 11- to 18-year-olds. They are called grammar schools because they were originally modelled on schools created in early Medieval times to educate priests and clerks in the Latin language – which was then (and occasionally still is) important in religious, legal, and governmental professions. Because Latin was seen as a prestigious language in Medieval Europe, and it has a closely described and quite fixed grammar, knowledge of Latin grammar became a useful skill to acquire. Understanding Latin provided a route toward a better-paid, more prestigious and more secure social position: a simple way to improve status in what was otherwise a heavily status-conserved social system.

This is only indirectly linked to the sources of language grammar; but it helps explain why an eclectic subject like grammar became so important socially; and it reminds us that local language grammars are often treated with some reverence in the human societies that use the language. Obeying local grammar rules is a marker of membership of the local group, and flouting those rules stigmatises the speaker as an outsider. Local grammar rules can even supersede standard rules: local dialects can be seen as indicating a “better” person than standard dialect in some contexts – and a “worse” person in other contexts. For instance, when I worked at a school in Bermondsey my received pronunciation and standard grammar were treated by some parents as a sign that I was an outsider; but they also helped to establish my credentials as a person suitable for work in an education environment. Grammar is just the principles, or rules, by which units of meaning – words – are organised into shareable meta-meanings. It is the way humans negotiate toward meaning and, theoretically, a tool available to, and known by, almost all of us; but it is also an effective social and cultural tool. While this book is concerned with grammar as the organisation of meanings, readers should not forget those meanings have social and cultural significance.

In this book, grammar is defined as a linguistic phenomenon. First, it contains the rules which an individual must attempt to follow if they are to fully take part in a particular community of discourse, the unspoken rules that define what is negotiation toward meaning and what is not. Second, grammar contains the rules which generate meanings in words and the ways of meaning allowed in the community of discourse. Third, it contains the set of rules which govern the use of words (units of meaning) in language constructs (words, phrases, sentences, utterances, and discourses), allowing meanings to be flexibly combined to create new meanings and composite meanings.

the first set of rules are rules of communication; the second set are rules of meaning creation and sharing; and the third set are rules of discourse construction. They are all needed for fully articulated human language grammar, and all three are the products of negotiation toward meaning: without agreement there are no rules, and without negotiation there can be no agreement. It is vital to remember this causality in any discussion of human language grammar.

The Roots of Language Grammar

To explain what caused language grammar, this book must address the question of what it is to be human. Particularly, it must look at how we make predictive models of the actions and interactions of others, and use them to enhance our cooperation and reputation. Cooperation and reputation are the root faculties from which we build our language grammar story: they are implicated in the evolution of our self-recognition, the development of the cognitive forms behind our language grammar – and, indeed, our capacity to construct stories, such as this one.

Why a story? Because language is a “soft tissue” problem, and its origins have left no obvious physical trace. Bones fossilise, and fossils can be dated; and we can tell a lot about a species from the hard facts they leave behind in terms of slaughter marks on bones, fire lenses, midden heaps, manufactured artefacts, coprolites, and other solid evidence. Language has, until recently, left no such traces. Instead, we must rely on the presumption that complexity in archaeological artefacts is a good indicator of social complexity, which in turn is a good indicator of signalling complexity. These correlations are assumed, however, and do not directly indicate the nature of signalling. Although it must be accepted that solid artefacts are the only evidence we have of what happened in the far past, we must not uncritically treat them as direct indicators of language and culture; and we should indicate clearly where we are presenting established knowledge about the physical evidence, and where we are making necessary speculations from the physical evidence – that is, where we are storytelling.

Storytelling has always been important in science. For instance, the story of the *Big Bang* remains one of our most illuminating explanations of how the Universe began and works. Yet the name was originally coined by Fred Hoyle, an opponent of the expansion model of the Universe, and a supporter of steady-state physics. *Big Bang* is a deliberate inaccuracy by Hoyle: the expansion model describes the sudden unfolding of space and time from a single point; and, because the single point was both the centre and the totality of the Universe at the start, the centre of the Universe is now everywhere. *Big Bang* is a misleading descriptor, but it still provides a hook on which to hang a larger and more accurate story. Similarly, the story told in this book is unlikely to be exact in all its details, but it hopefully provides a useful waypoint on our exploration of where we are and where we were as a species. The trans-disciplinary approach used in this book is part of a general, integrational approach to science, where stories told by different scientific disciplines can be merged into a single biography of the Universe. More parochially, I hope that this book offers important and lasting guides to the sources of language grammar; and, most importantly, that what follows tells a coherent and believable tale.

Studying Language Grammar

This book reviews language grammar from an anthropological perspective. Anthropology, however, is a wide discipline, overlapping with several other disciplines – including Sociology, the study of humans in groups, and Psychology, the study of human minds. Language, the social lubricant for our peculiarly human groups and a system of peculiarly human cognition, is therefore also of primary interest to anthropology. Exploring the sources of language grammar requires a varied range of academic approaches; but attention here remains focused on one key question: how does being human relate to having language grammar?

The journey takes a traditional approach: setting out the issues, suggesting a solution, and reviewing the solution in terms of available evidence. Chapter 1 formulates the problem this book addresses, while chapter 2 looks at current work on the sources of language grammar and investigates some of the routes that have been proposed to get from pre-human signalling to human language. Of particular interest is the significance these routes have for the sources of language grammar. Chapters 3 to 5 look at current linguistic theories, showing that linguistics is still a science of competing theories; the different emphases of these theories give us several approaches to language origins. In chapter 6 we look particularly at the anthropological perspective: the peculiar nature of human cooperation, and language grammar as a cognitive and communicative mechanism. The psychological perspective is presented in chapter 7: being able to recognise and work with models of ourselves and others is essential for human cooperation, so this chapter examines how we self-model and model others – and why it is, in evolutionary terms, weird. Chapters 8 and 9 then bring together the themes previously discussed to build a story of how we became human. They concentrate particularly on the human capacity to model both our self and others as generators of cooperation and language grammar. If you are particularly interested in this aspect of the grammar used in language, you may wish to consider this book's companion volume, *The Origins of Self: An anthropological perspective* (Edwardes, 2019).

In chapters 10 to 12 we review some of the evidence: what nonhumans, human children, and our understanding of time tell us about humans and language grammar; self-modelling and modelling others is shown in these chapters to be a determining principle for much of the grammar used in language. Finally, in chapter 13 the itinerary of the book is reviewed, identifying the major milestones and attempting to summarise the nature and roles of language grammar set out in this book.

Why This Book?

This book is about language grammar; specifically, where it came from. It should be a neat little package telling a simple tale about not much at all.

The book argues, however, that the story of the sources of language grammar is neither simple nor neat. Contrary to the view that language grammar is a coherent single system, I argue that language grammar is a hotch-potch of components, developed for a range of only partially related cognitive functions, which nonetheless came to work together as a single system in support of a novel communicative need.

Every book worth reading takes the reader on a journey from places known to places new, using intermediate landmarks – some familiar, some unfamiliar – to persuade the reader that the chosen path provides a reasonable description of the journey from source to destination. For some journeys, the start and end points are relatively unimportant when compared with the landmarks visited along the way: the landmarks provide the true story of the trip. This book is just such a journey. It does not always pass along well-travelled and pre-mapped roads, there are some less-travelled byways: the going may at times be difficult; and, in places, we make the map as we go along. You should also be prepared for controversy on this journey: the path lies across several academic battlefields – some historical, some still fought over. Prepare for incoming.

The journey should, nonetheless, provide an interesting excursion into the fascinating country of language grammar. So, don your walking boots and your personal choice of protective clothing; your journey starts here.

1 Why Is Language Grammar Important?

> ***We are glorious accidents of an unpredictable process with no drive to complexity, not the expected results of evolutionary principles that yearn to produce a creature capable of understanding the mode of its own necessary construction.***
> *(Stephen Jay Gould)*

What is special about being human? Camus (1951, II, 420) said that "Humans are the only creatures who refuse to be what they are,"[1] which makes humanity a complex and difficult subject to study. We reframe our existential world – the actual world that continues to exist regardless of our opinion of it – in terms of our individual interpretations of that world; and we then negotiate those interpretations with other humans until they become shared (and, hopefully, meaningful) models of the existential. Language is at the heart of this negotiation, which seems to make it a defining feature of human society, human culture, and being human.

This negotiation toward meaning makes humans a strange species; and, as individuals, we become aware of our shared strangeness when we first notice the similarities – rather than just the differences – between *you* and *me*. That noticing begins when I redefine myself as *my self*: I have an awareness that other people have selfhood, which means that everyone, including myself, is a self. Recognising myself as a self means that I can model *my self* in the same way that I model those other selves: as cognitive proxies for actual people.

[1] My translation of "L'homme est la seule créature qui refuse d'être ce qu'elle est".

Awareness of self and selfhood has led us to a second question, more complex than the first: what is special about the specialities that make us human? This second question is about the perspectives we take and interpretations we make of the skills that seem to define our species: do we view humanity from the inside looking out, using our skills as benchmarks against which to judge other species; or do we attempt to view humanity from the outside looking in, seeing our peculiarly human skills as instances within a range of natural specialisations? These two, largely exclusive, positions offer us different approaches to our understanding of human specialities; but it is only by contextualising the second question (what is special about the specialities that make us human?) in terms of the first question (What is special about being human?) that we can begin to understand both our special nature and our continuity with the rest of nature.

This chapter reviews the two questions in terms of some of our special physical and cognitive attributes. However, *special* has a particular meaning here: *the attributes of humans that are unusual in nature*. The word *special*, therefore, is not intended to create the impression that humans are somehow isolated from the rest of nature. In this book, our human specialities are viewed mainly from the outside looking in, taking the position that having language and grammar makes humans an atypical species – but not extraordinarily anomalous. The intention is not to identify differences that isolate humans from the rest of nature but to show that our species, like every other, has species-defining *speci-al* capacities.

When viewed as an indivisible system, human language grammar seems to be unique and, therefore, a marker of our species. This book proposes instead that grammar is not indivisible, it emerges from *speci-al* social, cognitive, and communicative capacities with their own, non-linguistic functions. Individually, these capacities seem unrelated to language or grammar; but collectively they require an integrated social, cognitive, and communicative structural complexity. The grammar used in human language is both partible and emergent: it is neither a monolithic system, nor is it without precedent in nature. Therefore, it cannot be the whole story of human language. The human capacities from which grammar emerges, by multiple routes, must be more deeply implicated in our humanness than language grammar itself. For evidence, this chapter looks at some of the capacities considered decisive in the evolution of our species, but which seem to be, at best, only indirectly involved in language; and it shows how they nonetheless can be implicated in the organisation and systematicity of human communication.

There are several species-specific physical attributes which enable *Homo sapiens* to do things that other species cannot. These attributes are proxies for significant cognitive differences because they show how our human mental world is qualitatively distinct from the mental worlds of other species; and, by reviewing these physical attributes, we can identify cognitive processes that define us as a species, and which may be implicated in language grammar. We look at some of these attributes next, starting with that often-cited but ill-defined quality of humanness: brainpower.

How Important Is Cognitive Capacity?

Many animals have brains. Certainly, all vertebrates have a nervous system with a primary control node on top of a spinal column. And all these vertebrate brains are involved in doing much the same thing: processing sensations from various parts of the body, producing holistic interpretations of the actual world outside the body, and generating whole-body reactions to those sensations and interpretations. In this approach, brains are viewed as processing mechanisms, using externally derived inputs to generate appropriate externalised outputs through the application of rules. We should remember that the brain itself is computational: it can interpret inputs and generate outputs, but it is only indirectly linked to the world beyond the individual's body, where the inputs are generated and the outputs expressed. The brain does not see, hear, feel, taste or smell, it only interprets electrochemical cues generated by the body, usually in response to other cues from beyond the body's boundaries; and it cannot act, it can only generate electrochemical cues which cause the rest of the body to act. The brains of different species vary enormously in their construction, the sensations they attend to, and the rules they use to generate bodily reactions; and, even within a single species, individual brains do not react exactly the same. The value of a brain, though, is in its capacity to produce efficacious responses that serve the organism of which it is a part. Having a wider range of responses is not necessarily better in the realm of brainpower.

Nonetheless, looking at brains from inside our own species, what seems to be important is logical and productive intelligence (where *intelligence* can be defined as knowledge gained, stored, and available for reuse); and this does rely on, and is enhanced by, processing capacity, where more is usually better. Human intelligence allows us to interact with the local environment in ways that can change it drastically: we do not just exist in a single niche environment on this planet, as many other successful species seem to do, we can live in a range of environments, and even make them more comfortable for us. We are not the only hominin to constructively interact with its environment: there is evidence that, at about 125kya during the Last Interglacial period, Neanderthals at Neumark-Nord (in Germany) changed the local vegetation using fire and tree clearance (Roebroeks *et al*., 2021). Laland *et al*. (1999) have used the term *niche construction* to describe the adjustment of the local environment by a species to favour themselves; and they show that it is a feature of many lifeforms, including the humble earthworm. It can even be argued that life itself is just arrangements of chemicals which use their local environment to make copies of themselves.

However, the effect that human intelligence has on the environment, both locally and globally, is greater than that of any other species. Indeed, it currently seems to be the dominant environmental influence on this planet (Lewis & Maslin, 2015). This ability to not just react to the environment but appropriate it is a direct product of our intelligence; and, as a species, we therefore tend to see intelligence as a marker of relative fitness, both within our own species and between species.

Unfortunately, bare intelligence does not really indicate anything useful in evolutionary terms. The evolutionary significance of something is not in its quantity but in the advantages it gives, and the abstract quality of intelligence by itself seems to be a poor gauge of evolutionary success: ants, with their individual brains of about 250,000 neurons (compared to *Homo sapiens'* average of 86 billion neurons), are nonetheless probably the second-most successful clade after *Homo* (Foitzik & Fritsche, 2021). Only when intelligence is socially applied (Gavrilets & Vose, 2006; Nettle & Pollet, 2008) does it become a useful marker of human evolutionary success.

Physical indicators of sophistication in the brain itself are even less reliable. If we measure brainpower simply by brain size then elephants have us far outclassed (Shoshani *et al*., 2006). If we rely on the surface complexity of the brain then whales and dolphins are ahead of us (Marino *et al*., 2007). If we see brainpower as the capacity for behavioural plasticity, then cephalopods (octopoi, squid, cuttlefish, *etc*.) can offer us significant competition (Godfrey-Smith, 2016, 50-59). If we choose brain/body ratio (the size of the brain compared to the size of the body) as Aiello & Wheeler (1995) propose, then our extinct cousins, the Neanderthals (*Homo neanderthalensis*), edge us out (Stringer & Gamble, 1993); nonetheless, they are extinct and we are still here. If we look at the absolute number of neurons in the cerebral cortex (the dome of grey matter brain cells which overlays the rest of the brain in mammals and reptiles) then modern humans, at 11.5 billion, come out just ahead of, but not significantly ahead of, whales and elephants (Roth & Dicke, 2005). There is even evidence that the cortex itself may be only a proxy measure of brainpower: the cerebellum, long seen as a primitive feature of the brain, has now been shown to have a complex a relationship with the cortex; and for every one neuron in the cortex there are 3.6 neurons in the cerebellum (Herculano-Houzel, 2010). All these measures are, however, crude reckoners of brainpower, equivalent to counting the value of notes and coins in circulation to calculate national worth: it is not the amount of cash available but how it is used that makes wealth; and it is not the number of cells available but how they are networked that makes intelligence (Bassett & Gazzaniga, 2011).

If physical measures of brains do not explain the difference between us and other animals, perhaps it is the computing power of our brains that sets us apart. Human brains are more complex in three important ways: connectivity, human brains have more connections between different parts of the brain; dynamics, human brains are chemically slightly faster and more sensitive to biochemical changes; and information, human brains have greater circuit redundancy and greater combinatorial power, creating greater synergy of information (Sporns, 2022). These abstract features of brainpower generate speed and versatility of cognition, allowing us to use tools such as IQ tests to compare intelligence between humans. Mostly, though, the tests cannot be applied to nonhumans: they contain elements which are human-centric and of no relevance to other animals. Our IQ tests cannot be applied across species, and they cannot therefore give a useful measure of evolutionary intelligence. The value of IQ tests in measuring even human intelligence has been questioned by some cognitive scientists (Gould, 1981).

The situation is complicated by the fact that modern humans are not even the brainiest *Homo sapiens* to have lived on this planet: the brains of our species have been shrinking since about 30kya, and average human cranial capacity is now about 10% less than at its peak in the Mesolithic (Henneberg, 1988). This reduction may not, by itself, indicate a reduction of brainpower: *Homo sapiens* brains use energy more efficiently than other primate brains, although there is no indication when in our evolution this improvement occurred (Beaulieu-Laroche *et al.*, 2021). Even without direct improvement in processing, the expansion of our internal memory systems with offline storage (*e.g.*, writing systems) has meant that we can access more information that we can carry around inside our heads. There is some evidence for a further reduction in human brain size after 3kya, concurrent with the rise of writing as an information storage medium (DeSilva *et al.*, 2021).

How Important Is Bipedalism?

Bipedalism has also been considered key to being human, with important implications for cognition. One early theory, however, illustrates the difficulty of working from physical difference through mental control structures to identify an evolutionary need for bipedalism. Hewes (1961) proposed that walking on two legs instead of four (physical difference) freed the hands to do other things (mental control structures), specifically to transport large amounts of food from where it had been found or killed to a home base (evolutionary need). This idea was attractive because it implied that early humans were highly cooperative, and it supported the origins of human manipulation: motivated by the need to carry things, manipulation preceded and opened the way for tool use. However, current evidence of australopithecine scavenging and hunting techniques indicates that, despite being bipeds, they went to the food and ate it where it was; they did not waste calories carrying it around the landscape (Domínguez-Rodrigo *et al*., 2005).

Harcourt-Smith (2013) presents the issues surrounding bipedalism in detail. First, the *Homo* clade are the only primates to have obligate bipedalism (we move more effectively on two legs than four). Other species, like chimpanzees, may sometimes move bipedally, but their stance and step placement make their walking more a shuffling stagger than a controlled stroll, and less efficient than quadrupedalism. While earlier hominids (*e.g.*, *Sahelanthropus tchadensis*, *Orrorin tugenensis* and *Ardipithecus kadabba*) were able to spend extended periods moving bipedally, they retained both quadrupedal movement and brachiation (using the arms to swing from branch to branch). In contrast, *Homo sapiens* adults find quadrupedal movement awkward and inefficient, while our wrists have evolved for increased manipulation, which makes brachiation dangerous for us. Australopithecines were more effective at bipedal movement than previous species, but were still competent quadrupedally; and even early *Homo* specimens, for example *Homo habilis*, seem to have retained some quadrupedal traits. Obligate bipedalism probably began with *Homo erectus*, about 2mya (Dunsworth, 2010).

Harcourt-Smith gives several reasons why bipedalism became a useful feature of the *Homo* lifestyle: manipulation of objects (food-carrying, toolmaking and use, food gathering and processing, *etc*.); acquiring food on the savannah (faster movement through the grass, increased visual range by raising eye level, *etc*.); keeping the body vertical to reduce exposure to sun heat; and more efficient locomotion (*e.g.*, long-distance running for pursuit hunting). However, he finds all these explanations to be individually insufficient, proposing instead that no single factor led to an evolutionary selection for bipedalism; they all developed together over millions of years and many species.

Treating bipedalism as the capacity that made us human is problematic because there are several different incentives for a species to become bipedal. For instance, chacma baboons use a bipedal stance to look for predators and, if they maintain a bipedal stance, it calls the attention of other baboons to what they are looking at (Byrne, 1995, 125-126); kangaroos are obligate bipeds because they use their back legs for a novel method of locomotion which is more efficient than quadrupedal locomotion; and birds are obligate bipeds because they use their front legs for a different type of efficient locomotion. Basically, if either the front or back legs are evolutionarily repurposed then bipedalism is inevitable; but this happens only if the repurposing increases the fitness of individuals in the species by improving food acquisition, predator avoidance, or reproductive success (Day, 1986). Bipedalism is probably implicated in the process by which we became human because freeing the hands from the job of walking was significant, and it created a niche around which several human cognitive functions seem to cluster; but it is unlikely to be the sole cause, or even the major cause, for the cognitive differences between humans and other animals.

What about tool use? If bipedalism by itself does not explain our humanness, perhaps one of its products, using our hands to manipulate tools, is what differentiates us from other animals. This idea is certainly plausible, humanity is clearly the most efficient tool-using animal on the planet. Tools extend our physical capacities, enhancing our strength, our reach, and our ability to hold and use objects. Modern humans use tools for many tasks, even where they serve no practical purpose, such as cutlery for eating.

For decades, tool use had been considered the most likely differentiator between humans and other animals. However, chimpanzee studies in the wild have shown that they use tools in activities such as termite fishing – eating the termites that swarm onto sticks poked into termite mounds (Sanz *et al*., 2004). Other studies show them cracking open nuts between a stone hammer and a stone or wood anvil – a skill requiring learning and practice, so not innate (Boesch-Achermann & Boesch, 1993). Two potential ways of differentiating *Homo sapiens* were thus ruled out by this one study: both tool use and in-life learning of tool-related skills are not exclusively human.

Bipedalism and tool use have been identified in a wide range of other animals, and they are no longer seen as solely human capacities. Nonetheless, both have been enshrined in our evolutionary taxonomy: the earliest *Homo* fossil discovered, in 1891 by Eugene Dubois at Trinil on the island of Java, is now called *Homo erectus* (upright human); and the most ancient fossil identified as a member of the *Homo* genus belongs to the species *Homo habilis* (skilful human). Despite the names, however, both were bipeds and both made and used stone tools. To understand the sources of language grammar we must move beyond bipedalism and tool use.

How Important Are Hunting and Culture?

Other behaviours have aspects peculiar to humans, making them features that could define humanity, although most of them (such as religiosity) require humans to already be in the universe of symbols created by language and human culture: they are products of, rather than sources of, being human. One non-symbolic behaviour, while not exclusively human, does have a peculiarly human signature, however: the way that we hunt.

Hunting is not limited to humans, many carnivores hunt. Group hunting is also not uniquely human: several species hunt in packs, and they have different sharing strategies to ensure that cooperation in the hunt is worthwhile for each participant. There is, nonetheless, something unusual in the way humans cooperate in a hunt, and in the way we share the resulting food throughout our community. When compared to the pursuit hunting of *Homo sapiens*, the hunting methods of *Pan* species, particularly chimpanzees, are quite different. Chimpanzee hunting is *ad hoc* and opportunistic; unlike human hunting, it does not seem to be strategically planned, while *Pan* food gathering often does. A chimpanzee hunt is usually initiated by a small number of individuals, with others joining in when they notice what is happening. It is a short affair, lasting minutes rather than hours; distribution of the kill seems to be controlled by the individual with the carcass; and the kill is consumed raw. (Newton-Fisher, 2014).

Human gatherer-hunters seem to separate specialisations by gender, in that men hunt and women do not (Balme & Bowdler, 2006). There are exceptions (Noss & Hewlett, 2001), and the role of women in hunting is greater than was previously believed (Anderson *et al.*, 2023), but this does seem to represent the situation for most of the gatherer-hunter cultures surviving today. Why this gender specialisation should occur is not immediately obvious, it is not based on abilities: the gatherer-hunter exceptions show that women are fully capable of organising communal hunts to regular schedules; and modern Western women, like Sarah Palin, can also be avid hunters. Mythology, too, is replete with goddesses who hunt: Astarte, Artemis, Diana and Freya (Indo-European); Yama-No-Kami (Japanese); Cihuacoatl (Aztec); and Achimi (Kabyle), to name a few.

One universal feature of human subsistence hunting seems to be sharing: hunting to primarily benefit others rather than the self seems to occur in all human gatherer-hunter groups. Sharing surpluses is not uncommon in nature, it generates a web of reciprocal debts which allow an individual to rely on others when personal foraging has been less successful (Wilkinson, 1984); but giving up the whole of your portion of foraged food, as many human hunters do, is rare. This food surrender is a ritualised act, surrounded by taboos, traditions, and punishments for those who do not share; and the hunters must also expect to receive some compensation for the food – sex or kudos or even a share in the food. Nonetheless, the food surrender is expressed both ostensibly and culturally as an act of altruistic modesty (Lewis *et al.*, 2014). This dislocation of the relationship between hunting and food is an important species difference, and it has significance for language.

The fact that food is subject to ritualised distribution implies that there is a mechanism to enforce this ritual – a system of moral and coercive processes we call culture. However, we need to be careful about what we mean by culture. One definition is that it *consists of the non-genetic capacities of individuals which are transmissible between individuals*, and it therefore relies on mechanisms which allow the sharing of information. This definition of culture can be divided into three key components: a capacity for innovation; a capacity for transmission; and a capacity for personal learning. These capacities are not limited to humans, though: apes (Lycett *et al*., 2007), macaques (Flack *et al*., 2006) and other monkeys (de Waal, 2004), as well as dolphins (Rendell & Whitehead (2001) and elephants (Bradshaw, 2004), have all been observed producing innovative actions that subsequently propagate through the community.

This definition of culture includes both what is accumulated by innovation and imitation within the group, and what is accumulated over generations by intergenerational teaching and learning. The first is usually simply called *culture*, the default way that culture is generated, while the second is called *cumulative culture*. Cumulative culture is a steady accrual of traditions which, over time, can lose their evolutionary fitness value, surviving only because they are habitual. They can serve as markers of group membership (we do this thing in this way in our group), and therefore costly signals of belonging, but the utility which originally caused them to be adopted has often faded away. Sometimes these markers can become genetically encoded, turning arbitrary physical features into markers of group membership, a process called *cultural evolution* (Boyd & Richerson, 1996). However, attempts to limit cumulative culture to humans, or even to *Homo sapiens* only, face a growing body of counterevidence: instances of cumulative culture can be observed in a range of other species – other primates, some birds, and some cetaceans (Mesoudi & Thornton, 2018). There is also some evidence that the division between culture and cumulative culture is not well-defined: they may form a continuum, where accumulation is a matter of degree (Davidson, 2016).

Nonetheless, there remains an important difference between transmitting advantageous technical innovations, like nut-cracking, and adopting belief systems that have currency only in terms of group identity (Premack & Hauser, 2001). Human culture is not just quantitatively different from other cultures, it has a depth and breadth unrepeated in nature. Our culture is not just knowledge which has not been encoded into our genes, not just the social compromise needed for living in groups, not just a way to differentiate group members from outsiders, and not just a costly signal of cooperation. It is all those things, but it is also arbitrary. It is a complex system for generating and exchanging meanings – both etic meanings, which are true by their nature but not necessarily obvious to everyone, and emic meanings, which are true only because we agree they are (Leerssen, 2021). The knowledge we exchange, the compromises we make, the groups we form, our flags of membership, and the costs we pay to signal cooperation, are often based around what, in practical terms, can be classed as pointless activities. Culture is not just signalling solidarity for solidarity's sake, it is what generates our mutually beneficial human societies (Powers *et al*., 2021).

If we look for the sources of language, the human cultural environment of cooperative hunting provides fertile ground. Language is symbolic, allowing arbitrary relationships between sounds and meanings; human culture is arbitrary, allowing symbolic relationships between individuals (such as nationality) to be treated as real. Language permits metaphor, allowing apparently incompatible meanings to be combined into arbitrary hybrid meanings; and metaphor allows different personal contributions to be compared using arbitrary hybrid measures, such as money. Language is segmented, differentiated and hierarchical, allowing specialised components to be combined into shareable meta-meanings; human culture allows individuals to specialise, working together within subgroups to create meaningful and stable enterprises within the group itself. These cognitive similarities between language and human culture point toward a strong relationship.

While cognitive complexity is clearly necessary to enable linguistic cognition, and bipedalism leading to tool use and tool manufacture all indicate cognitive complexity, it is hunting that provides the first signpost for exploring the important differences between humans and other animals. Our cooperative hunting and food sharing strategies opened the way for a complex cooperative culture which required a communication system – language – capable of sharing the complexities of that culture.

How Important Is Cooperation?

If human culture is quantitatively and qualitatively different from that of other animals, and if language is a reflection of human culture, then comparing language with other forms of signalling should identify important differences. The flexible communication system we call language occupies a signalling niche that other animals do not appear to have exploited. We can share information about events in the world around us, and we can speculate on those events; we can time-travel, treating events in the past as currently significant and forthcoming events as already completed; and we can make things up. In most of nature, communicating about non-existent events means deception, trying to gain advantage by misdirection; in humans, the misdirection can be shared with the listener, turning it into storytelling, a skill valued in all human cultures (Smith *et al.*, 2017). Other species have effective signalling systems which contain some language-like features as well as unique features of their own; but, when considered as systems, they work differently from human language in significant ways.

When we look at the features of language in detail, however, exclusivity begins to fray; several features once believed to be exclusive to language have been identified in other signalling systems. For instance, reference to non-immediate locations is used by honeybees (*Apis mellifera*): foragers can signal the distance and direction of food sources by performing a “dance” when they return to the hive; other foragers then use the dance performance as a map to find the food (von Frisch, 1973). Using multifunctional indexical signals has been demonstrated in chimpanzee play: The chimpanzee *play face* is used not just to invite juveniles to play, it also reassures mothers that any rough-and-tumble is not intended to hurt their offspring (Flack *et al.*, 2004). Diana monkeys (*Cercopithecus diana*) use segmented signals where different segments change the meaning of the signal; they can communicate type of predator as well as its distance, and thus allow the listeners to respond with a range of behaviours (Zuberbühler, 2000). In an experimental environment, one particular grey parrot (*Psittacus erithacus*) called Alex successfully understood naming and identified attributes of things using human adjectives (Pepperberg, 1999); and a collie dog called Rico successfully linked novel names to novel items (Kaminski *et al.*, 2004).

Possessing capacities for language seems a matter of degree rather than an absolute difference between species. That matter of degree is, however, significant: humans exchange considerably more complex messages than other animals. Recent attention has therefore concentrated on language grammar, the mechanism that makes complex messages possible.

Language grammar seems to represent a pinnacle of humanness. It is itself complex, greatly increasing the complexity of human signalling compared to other signalling systems; and that complexity appears to be missing from natural nonhuman signalling. Grammar provides a schema for naming objects and mapping them into a cognitive system of interrelationships; and it allows those mappings to be shared between minds by converting them into a one-dimensional stream of sound. This cognitive system is not a faithful representation what is actually happening in the world around us, it is an impression, or model, of that actual world: as well as actual things and events, it represents memories and what-ifs, imagined objects and fictional events, opinions and viewpoints, all in the same way; and language lets us share this cognitive system with others. This poses a problem for Darwinian signalling theory, which says that signals are only valuable to the receiver if they are reliable indicators of significant events in the actual world; so how can a signalling system relying on fictional representations become endemic in a species? What advantages accrue from what is, essentially, deception?

In Darwinian evolutionary accounting, being part of this world of fiction must have advantages for the individual: the individual's own reproductive fitness must somehow be enhanced by cooperating in Nietzsche's (1874 [1976]) lie of language[2]. This is odd: susceptibility to deception is usually an indicator of evolutionary inefficiency. How can speakers habitually generate untrue messages when a genetic resistance to duplicity should create distrust of known liars (Zahavi, 2003)? Basically,

[2] What Nietzsche said was, "wahrhaft zu sein, das heißt die usuellen Metaphern zu brauchen, also moralisch ausgedrückt: von der Verpflichtung, nach einer festen Konvention zu lügen, herdenweise in einem für alle verbindlichen Stile zu lügen". ["Being 'true' means using the usual metaphors; or, to put it morally, lying in an agreed way, cooperating in a lie that binds everyone" (my translation).]

because language is essential in human cooperative culture, and sharing agreed deceptions is essential in language. Cooperating in agreed lies advantages an individual sufficiently to make it a viable strategy, and that may be all we need to explain how language can be both deceptive and effective. This, however, only relocates the problem without solving it: we cooperate in the deceptions of language because of our cooperative culture, but what fitness advantage does an individual get from being part of a cooperative culture?

The problem of cooperation is vital for understanding the sources of language grammar; but the Darwinian paradox of cooperation must be explored using Darwinian tools: personal fitness, reproductive fitness, kin selection, reciprocal altruism and costly signalling. These tools all follow the basic scientific facts of evolution: that all life consists of a series of remarkable molecules which can copy themselves; and that those molecules, or genes, all access the same limited pool of raw chemicals to build copies.

To understand how cooperation emerges from this highly competitive, selfish-gene universe (Dawkins, 1989), we must look for a genetic explanation. Individuals cooperate because they are closely related (Hamilton’s *Kin Selection*, 1964), because they can expect cooperative behaviour in return (Trivers’ *Reciprocal Altruism*, 1971) or because cooperation makes them attractive to potential mates (the Zahavis’ *Handicap Principle*, 1997). Our explanations of the cooperative behaviour behind socialisation, culture and language should work within a selfish gene framework.

How Important Are Our Genes?

If language is a fitness solution effective at the genotypic or species level, then that fitness needs to be identified: what makes language so useful to humans? To answer that, we must define what we mean by language: what counts as language and what does not. Traditionally, linguistics treats language as a single system central to both communication and cognition: so, in terms of structure, communication and cognition must be similar, and answering the question of what makes language useful to humans also shows us what makes human cognition different. Unfortunately, a single solution has proved elusive, making it increasingly likely that we should be looking for a series of fitness enhancements rather than a single one. The evolutionary fitness of human cognitive controls, vocal controls, socialisation constraints, enhancements to cooperation, cultural mechanisms, Theory of Mind, and tool manufacture must be explained before we can properly describe how primates, signalling without language, became language-using humans. This does, however, allow us to approach the sources of language grammar as a series of small changes, a more likely evolutionary route than a single massive change.

In the traditional model, language consists of semantics (the meanings and intentions behind utterances, what the utterance is being generated to do) and grammar (the rules the utterance must follow to ensure that the meaning understood by the receiver corresponds with the meaning intended by the speaker). Grammar is further divided into morphology (how words are internally constructed), and syntax (how words combine into utterances). Between semantics and grammar there are lexis (the words themselves, which have both morphological and semantic content) and phonology (the sounds and gestures used to represent the words). Except for phonology, this model of language does not appear to be a system which could evolve incrementally.

In theory, semantics can exist as a cognitive mechanism without needing to be expressed through language: allocating meanings to regularly-encountered events is just an abstraction of concrete feelings about those events (Feeney & Edwards, 2021). What, however, is the fitness value of having an abstracted version of a fully-effective concrete mapping; and how could it work without tags – words, or cognitive precursors to words – to carry semantic content? Similarly, grammar can exist cognitively as a separate system: rules must exist to govern any task where a series of subtasks must be carried out in sequence. However, the subtasks must be conceptually segmented, requiring words or their cognitive precursors; and this type of open-ended cognition seems exclusively human (Heintz & Scott-Phillips, 2022). Semantics and grammar would seem to have no cognitive validity without lexis, implying that any incremental evolution of language must start with lexis, or a discrete, combinatorial form of cognition. What use, though, is conceptual segmentation with meaning? And, as the segments are to be combined, rules are needed to differentiate useful combinations from useless. Lexis cannot evolve without semantics and grammar.

A language utterance negotiates agreement between sender and receiver, and that meaning comes out of both the parts that make up the utterance and the way they are combined: semantics requires lexis and grammar, lexis requires grammar and semantics, and grammar requires lexis and semantics.

There are two ways out of this Mexican standoff. The first is to assume that all three – lexis, grammar and semantics – evolved in a single genetic change, or macromutation (Chomsky, 2006, 176-184). This approach does not need to explain the paradox incrementally: the problem of which came first is avoided, they all appeared together. However, successful macromutations seem to be extremely uncommon. The majority of individual genetic mutations that create changes at the phenotypic level are disadvantageous to the phenotype; only very occasionally does the appearance of a new gene allele create a phenotype that is fit enough to become established in a population. The chance of a cluster of advantageous mutations all occurring together (a genotypic macromutation) is highly improbable.

A second type of macromutation is phenotypic or homeotic: a small change at the gene level produces large changes at the individual level. This macromutation, or exaptation, is more common (Fitch, 2011): the change at the gene level is small, but it has a cascade effect on protein levels which can significantly affect the body form. Phenotypic macromutations that successfully produce large-scale changes are, however, still rare: the process of producing *hopeful monsters* more commonly creates unhopeful monsters. While some cases of successful phenotypic macromutation have been identified, they occur mostly in the plant kingdom (Theissen, 2006).

Macromutations do occur: the change from 48 chromosomes in apes to 46 in humans may be an example of a genotypic macromutation; and the appearance of different skin shades in many mammals is classifiable as phenotypic macromutation. However, genotypic macromutations tend to be deep coding changes which have very subtle, or no apparent, effects on the phenotype; and phenotypic macromutations tend to have obvious but cosmetic effects on the phenotype. Neither of these matches what we would need from a macromutation to generate language. We cannot dismiss a macromutation causing language as impossible; but other options must be fully explored before accepting it as likely.

Instead of a macromutation, the lexis-grammar-semantics paradox could have an incremental solution: The path to full language started with minimal lexis, semantic content and grammar, to meet a specific fitness need in a limited area of cognition (Jackendoff, 2002). There is still a single genetic change simultaneously producing lexis, semantics and grammar, but they are limited and specific, not pervasive and disparate as in modern language. Once the three rudimentary forms exist then, even though language itself does not exist, the hard part is over: the paradox disappears. There remain only two, simpler questions: which increments to which aspects of language made it into a generalised communication system; and in what order did they occur?

The "only" in the previous sentence is not intended to imply that these problems are trivial; the incremental changes, and their order, are crucial. In evolutionary terms, however, a series of small changes are easier to explain than a single large one. This does mean that language development took hundreds of thousands rather than thousands of years – there must be enough time to allow the many small changes to accumulate. However, as the changes proposed are cognitive, and not necessarily communicative, there is no need for a series of protolanguages to precede full language (although the model does not preclude this). The phonological problem, how the cognitive forms became expressed as signals, can have a separate solution to the lexis-grammar-semantics paradox, and is not examined further here.

What Is the Function of Language Grammar?

When investigating why language evolved, a good starting point is to review its current uses. It is unlikely that language was initially as versatile as today; but why such an unusual system of communication became so important is likely embedded in its current functionality.

On the broadest level, modern language occupies two particular niches: cognition and socialisation. The cognitive uses for language can be summarised as rational problem-solving and planning, which forms only a small part of the everyday cognition we undertake. Most of our problem-solving is not rational, it is what Kahneman (2011) describes as fast thinking, or system 1 thinking: we make most of our decisions based on immediately accessible information, and we are often unaware of the information used, or even that a decision has been made. There is no careful consideration of the specifics of the problem before us, there is just decision and action. Why did you have the breakfast you had today? Was your choice based on habit, or what was available, or what you could eat on the run, or just eating what was put in front of you? Or did that *why* question raise a more fundamental query: *what* did I have for breakfast today? Don't worry if you cannot remember today's breakfast, unconscious activity happens much more frequently than we believe possible; we sleepwalk through a good part of our lives and are none the worse for it.

Kahneman contrasts fast thinking with slow thinking, or system 2 thinking. This is rarer and more costly in terms of time and cognitive resources; and, mostly, system 2 thinking is not followed to a reasoned conclusion. Instead, we short-circuit the laborious cogitation process by arbitrarily choosing what feels right – we do not pursue our system 2 thinking to an exhaustive conclusion, we instead reach a system 1 decision based partly on whatever system 2 cogitation has already been done and partly on system 1 intuitive reaction.

System 2 thinking is often linguistic: it is propositional, asserting truths and falsehoods about the universe; it is evaluative, measuring and comparing different aspects of the universe; and it is systematic, relying on shared meanings and forms. System 1 thinking, by contrast, is rarely linguistic – although, weirdly, the actual production of spoken language is usually seen as a system 1 process. We convert the intention to share meanings into shared meanings mostly without being aware of the processes we are using. However, we do occasionally use system 2 thinking to consider the form our utterance should take before (or while) it is uttered. Language seems to be a liminal system: it is not wholly subliminal nor is it wholly conscious; instead, it moves bewilderingly between the two states – possibly a system 1.5 thinking process.

Yet not all system 2 cognition needs language, as a simple thought experiment shows. Imagine an object, any object; rotate the object so that you can see the other side; now turn the object upside-down. While this visual imagination experiment was activated by words, the transformational cognition itself required no language. During an average day, much of our cognition has linguistic content: we may feel we constantly conduct an internal monologue, but this *explains* and *describes* what is happening rather than *causing* things to happen. The internal monologue represents an ongoing linguistic reaction to what we are thinking; it is thinking about thinking, also known as *metacognition*, a system 2 cognitive process. Who is explaining or describing, and who the explanation or description is for, are problems reviewed in chapter 7. For now, we can say that, while language may not provide much to our cognition, metacognition is dominated by language. Metacognition is what we use when we analyse a problem or speculatively plan.

We can also consider language in terms of its social function. Some of these social uses have implications for cognition; but they began as solutions to social evolutionary needs. The social uses of language can be divided into four types: phatics, direction, negotiation, and information sharing.

The first two social uses are relatively simple. Phatics is the production of sounds and gestures to acknowledge and build our relationships with others. It does not require complex constructs and can rely on verbal gestures with simple meanings (such as *mmm*, a simple indicator that the listener is still engaged with the speaker), or utterances with very broad general meanings (such as *yes*); it is therefore closely related to non-verbal grooming used by other mammals (Dunbar, 1992). The second social use, direction, involves a willingness to engage in joint ventures, such as hunting or teaching toolmaking. It requires some cooperation between individuals, but it does not require complex language constructs – simple imperatives and non-verbal gestures, possibly accompanied by phatics, can suffice.

Negotiation, the third social use of language, requires a more complex signal: as well as the information-receiver's action of *asking-for* or *demanding* (which together are referred to as *manding*), there is the information-giver's action of *offering*, often a *quid pro quo* for previous giving actions by the other party. Negotiation is based around two different informational actions, object identification and process description: the information-giver and information-receiver exchange cognitively linked pairs of action-object constructs. Negotiation does not require the complexity of modern language: it requires only simple one-argument grammar forms (verb-plus-noun) with constrained directionality and temporality limited to the present tense (non-verbal gesturing rather than labelling with words).

The final social use of language, information sharing, requires the full complexity of language. It gives the information-receiver access to the experience and viewpoint of the information-giver; it involves events no longer or not yet current; and they may not be geographically local. To share information, the giver and receiver must each cognitively model a *virtual stage* on which the giver's experience can be re-enacted. This requires shared metacognition, and a shared meta-signalling system. For instance, the simple information that "Alf likes Beth" is usually accompanied by meta-signalling indicating the information-giver's views about the information: how they see themself, Alf, Beth, the information-receiver, and the relationships between them. Information sharing is not just signalling information, it includes signalling about information: the signalling system must have sufficient complexity to signal multiple temporal relationships, share absent events, describe the signal sender and receiver as third parties within the signal, and co-ordinate a series of events into a story. Social information sharing is where language begins to find its full communicative purpose.

So, a plausible evolutionary explanation for language grammar should concentrate on the mechanics of social information sharing. As Mufwene puts it:

> … language is itself a technology developed through the domestication by hominines of their own anatomy to express their thoughts and feelings, to describe various states of affairs around them, to relate past experiences, to plan future states of affairs, and, as claimed by several students of the evolution of language, to cooperate toward the sustenance of their societies. (2013, 353).

The key question of this book, why we need grammar, is linked to the questions of how a social structure requiring the exchange of complex information could have evolved, and what that social structure was.

2 Speculation on the Sources of Language Grammar

> ***Storytelling is not what I do for a living – it is how I do all that I do while I am living.***
> *(Donald Davis)*

Language origins used to be a toxic subject for the scientific community. In 1866, the Linguistic Society of Paris declared debate on the genesis of language lacked scientific evidence (Aitchison, 1996, 5), after presentation to the Society of several papers owing more to theological speculation than scientific method. Darwin's *Origin of Species* (1859) had offended biblical literalists, beginning an ideological conflict between religion and science that continues today. Language origins was avoided in scientific discourse, becoming a topic less respectable than even paranormal phenomena.

This situation did not really change until 1996, when the first of a series of conferences on language evolution was held in Edinburgh (published proceedings Hurford *et al.*, 1998). This series, which has become known as Evolang, is held regularly every two years, and it is now one of several conference series with an interest in early human societies. Academic conferences across a range of disciplines now accept papers on language origins, and the search for the sources of language is beginning to produce viable answers backed by solid theoretical frameworks.

However, the reopening of the language origins debate can be traced back even earlier, to Noam Chomsky's publication of *Syntactic Structures* (1957). This revolutionised the way linguistics was done – a remarkable feat for such a short book. Before *Syntactic Structures,* attention was on what made languages different from each other (*e.g.*, Bloomfield, 1933). The approach was largely structural, viewing languages as coding systems existing between minds, and to which individual minds subscribed. Chomsky's Generativist approach, in contrast, looked at language inside minds. Despite the title of his book, he did not ask the structural question, *how do individual minds come to language?* Instead, he asked a question about the form of language, *what features do human minds have in common that make language possible?* Some of the proposed answers to this question have not survived critical analysis; but the fact that Chomsky asked this question – and so established a new way of looking at language – affected linguistic theory fundamentally.

Language origins became academically respectable again when the Generativist view of linguistics, that language is an internal cognitive mechanism, encountered the Darwinian approach to evolution. The physical characteristics most humans share must be products of our genetic inheritance: they survived because they made our forebears reproductively successful. Language is a near-universal characteristic of humans, so it probably has a significant role in reproductive fitness – it is either a successful evolutionary strategy, or a very useful adjunct (Dawkins, 1996, ch6).

Play as a Source of Language Grammar

Play, when most successful, is a negotiation toward shared enjoyment, just as language is a negotiation toward shared meaning: both involve the exchange of tokens which are mostly valueless in terms of fitness but valuable in terms of human culture, where a permissive environment of social exchange emphasises negotiation (Tomasello *et al.*, 2005). Play signals are reduced versions of earnest actions: a play bite is a reduced bite, a play fight should not result in actual injury, a playful submission does not indicate actual submission. In play, roles can be reversed while the game is underway, because they have no significance outside the game (Bateson, 1985).

Actions in play simultaneously represent reality and deny it: the play bite stands for an actual bite, but it also indicates that there is no intention to actually bite. This corresponds to the way words function: a word like *lion* both represents and denies the real world. I can talk about an actual lion in a zoo, or an imagined lion which only has existence in my mind; I can expand the meaning of the word *lion* to cover not only gregarious cats in Africa but also solitary cats in America, which are clearly a different species; I can describe First World War soldiers as *lions led by donkeys*; and I can use the term to refer to characters in *The Wizard of Oz* and *The Lion, the Witch and the Wardrobe*, although they are fictional and share only limited characteristics with actual lions – or, indeed, each other. Language and play both rely on all parties accepting that everything is negotiable.

For most animals, play is an activity of pre-adults. Pre-adult animals can play with other pre-adults, and with some adults, because dominance is not an objective of play: more powerful individuals self-handicap to allow the pre-adults to learn how to win as well as how to lose. Adults, however, do not play with other adults; or, when they do, dominance is never far below the surface (Bauer & Smuts, 2007). This is because the "games" between adults are interactions establishing and maintaining relationships, so they are serious and consequential. Young animals play to practice adult activities and relationships in relative safety, before puberty puts them into an ongoing competition to feed and reproduce. Unlike other animals, adult humans indulge in many different types of play, indicating that the imperatives which dictate adulthood in other animals have somehow been mitigated (Cook, 2000, ch4). One type of play specific to humans is language itself – our willingness to cooperate in a signalling environment where speculation, opinion and fantasy are treated as real. Language is not just *like* play, it *is* play.

Heraclitus (c.535-c.475BCE) is supposed to have said, "A person is most nearly themself when they achieve the seriousness of a child at play", an aphorism which encompasses key features linking language and play. First, both language and play are implicated in human selfhood: they both allow a person to access the concept of being themself by letting them explore being *other*. Second, both language and play are treated as activities with no serious consequences – speech can be ignored and play can be treated as trivial; but both are nonetheless carried out as if they are significant, even if just for the discourse or game in which they occur. Third, both language and play are developmental: children in a particular age group speak, listen and play differently from children of other age groups, expanding their ranges of language and play as they grow and mature.

Parten (1932) looked at children's play in terms of negotiated turn-taking. She proposed five types of play which occurred at different stages of childhood. Initially, a child does not play, they experiment with motor functions and observe the activities going on around them. Solitary Play emerges during the first year: the child plays by itself with objects they co-opt into the game; we do not know whether the objects are treated as symbolic or are just explored for what they are, but Parten suggests that symbolism is not part of play at this stage. From 24 months onward, Parallel Play occurs: the child plays its own game, sometimes using objects also being used by other children, although they need not represent the same things in the other child's game. At this stage the play appears to be symbolic, with objects representing whatever the child wants them to represent – a block of wood can be a car, a horse, a cake, and so on. Parten also describes a second type of behaviour at this age, which she calls Onlooker Play: one child watches another playing without joining the game, although they may comment on the game in progress.

At 36 months children begin to truly play with others, in what Parten calls Associative Play. They share objects and take turns using an object, sometimes attaching the same meaning or role to it; but, despite high levels of interaction, each child can still be following their own story line. At age four, Cooperative Play begins: children begin to negotiate together on how play is to be conducted, ensuring that everyone is playing the same game. This negotiation is ongoing, adjusting as play proceeds: the children are not using a formal script created before play begins, they are developing the script as the play develops. Around age six, cooperative play becomes more formalised. This stage was not recognised in Parten's original model but is now usually referred to as Intentional Play: children recognise the intentions of others and consciously acquiesce in a shared play script which is usually agreed at the start of play. Playground and party games like tag, hopscotch, musical chairs, and so on, are examples of early intentional play, which are replaced as children get older with increasingly formalised sports and activities.

While Parten's model remains a standard analysis of play, other models have been suggested. Piaget (1947 [1950], 138-140) took the view that play develops through four stages. At the Sensorimotor Stage (up to 24 months), the child is discovering the world through their senses, and also learning how to manipulate symbols as representations of reality; Piaget calls this Functional Play. Between ages two and seven the child is in the Preoperational Stage and using Constructive Play: constructing knowledge and solutions which help the child understand and navigate the social world around them. At about age seven the child enters the Concrete Operational Stage, and begins Symbolic-Imaginative Play: it is symbolic in that it creates its own internal meanings and representations; it is imaginative in that objects and roles used in play can represent other objects and roles, or can just be imagined; and it is cooperative in that the child's attention is directed at making the game work for other participants, rather than for the child only. For Piaget, symbolic-imaginative play is available only to humans: other animals indulge in functional play but have no access to the world of symbols. However, this is not the final type of play: humans also have Rule-Driven Play which becomes active in the Formal-Operational Stage. This corresponds with Intentional Play in Parten's model.

The main difference between Parten's and Piaget's models is that Piaget describes play in terms of the individual playing, while Parten sees play as an interactive socialising mechanism. This may seem like a small change of emphasis, but it reflects a key difference: for Piaget, childhood is a period of individual psychological development which follows a predetermined path toward conventional adulthood unless frustrated; for Parten, childhood is a process of acculturation, and the adulthood that emerges is more individual and less conventional.

Huizinga (1938 [1950]) took a different approach to play. He saw it as an agreed suspension of "everyday life" with the imposition and acceptance of an arbitrary set of new rules. Play is essentially either a contest or a performance, and therefore a formal, even ritualistic activity. Huizinga lists five features identifying play: it is undertaken willingly and can be suspended at any time; it is outside everyday life, allowing everyday life to be ignored; what happens in play stays in play, there is nothing to create gains or losses in everyday life; it happens in specific spaces and at specific times; and finally, play is rule-driven – what for Parten and Piaget is only the final type of play is for Huizinga the only true form of play, because it is the only form that is culturally and ritually defined, imposing order on chaos. Huizinga's model of how play works is more specialised than those of Parten and Piaget; he is interested in the role of play in culture, specifically the Western European culture of the mid-20th century. He is less interested in play as a behaviour, concentrating on play as a cultural mechanism; which means he ignores the way that play develops in a human lifetime, and therefore how it developed in human evolution.

Caillois (1958 [1961]) built on Huizinga's model, while disputing Huizinga's emphasis on competition. He divides play into four types: Agôn, or competitive play, where the existence of a winner automatically implies the existence of a loser; Alea, or playing against random chance, where the existence of a winner does not imply the existence of a loser or vice versa; Ilinx, where the individual's body is the opponent and playing is testing the limits of what the body can take; and Mimicry, or playing at being another self, where there are no winners or losers, only simulation. All these types of play require the player to model themself, others, and relationships between people; but each type of play requires a different type of modelling.

Agôn models other people and the relationships between them; and the question driving the modelling is, *what will they do next?* There is no need to model the self in this type of play, it can be undertaken by any socialised animal with awareness of others – at least, during its childhood.

Alea, in contrast, requires self-modelling, because it projects a self-image into at least two possible futures: where the gamble succeeds, and where it does not. Most other animals do not need self-modelling because they choose their responses subliminally, based on intuitive reactions. Humans can respond subliminally, too: attention to the self is not a default state for humans, only a small part of our cognition involves the recursive self-modelling implicit in Alea. The question driving the modelling is, "what do I do next if the universe does B when I do A?" This requires a modelled self who has done A, a model of the outcomes of doing A, a modelled self who has suffered the outcomes of the universe doing B, and a model of the choices available after A and B have happened.

Ilinx also requires self-modelling, to compare the pre-play and post-play selves. spinning until you are dizzy compares the induced dizzy state against the default undizzy state. Two models are needed: a model of the post-play self by the pre-play self to anticipate the enjoyment of the Ilinx play; and a model of the pre-play self by the post-play self to assess the outcome of the play. The questions driving the modelling are "How will I be different after playing?", and "what was I before playing?". As with Alea, there is a comparison between the current self-model and the future or past self-model.

Mimicry is different from other forms of play in that there is no winner, and no personal targets to be achieved. Mimicry is itself both the action and goal – there is no simple separation of form and function. Instead, the separation occurs at the symbolic level: the *I* doing the mimicry is not primarily modelling the *me* as mimic; instead, the *I* doing the mimicry is primarily modelling the person being mimicked, and only secondarily modelling me as mimic. The success of mimicry is judged by the similarity between the modelled other and the *me* modelled as mimic, while the *I* doing the mimicry is not part of the game. In the theatre they describe the intrusion of the *I* doing the mimicry as *breaking the fourth wall*.

What validates mimicry is not an internalised model of self, but recognition of the person being mimicked. I should be accurate in my mimicry not to entertain or inform me, but to entertain and inform you. Mimicry is, therefore, intrinsically social, not just in terms of engaging attention and establishing a compact of play. Ilinx and Alea clearly need no social element, while Agôn requires another person only as a target for the game – a non-animate object often serves almost as well. Mimicry is, therefore, the most sophisticated of the forms of play, and requires an iteration, or recursion, between the questions *what would they do if they were me?* and *what could I do if I were them?*

Mimicry for play is not the same as mimicry for learning. When a wild chimpanzee attempts nut-cracking or termite fishing after having seen another chimpanzee do it, the action and the goal are quite different. The action is related to the adult, non-play versions of Ilinx (*e.g.*, refining a practical skill) and Agôn (*e.g.*, competition for food) rather than to Mimicry: the importance of the goal gets in the way of learning the action through imitation, and it becomes almost impossible to dispassionately divide the task into parts and to master each part separately. In fact, it seems that each chimpanzee learns a new skill mostly by trial and error, rather than by repeating the methods behind the skill (Tomasello *et al*., 1993).

Langley *et al*. (2019) take a different view of the relationship between play and language, arguing that play led to increasing cognitive capacity and increasingly complex language, facilitated by the extended childhood of *Homo sapiens*. The extended childhood was, in turn, a product of self-domestication, which increased docility and group cooperation. How this came about is not fully explored, even though self-domestication has been identified in very few species. For Langley *et al*., self-domestication increased cooperation, making childhood longer and increasing cognitive capacity; and this led to a ratchet effect between play and language complexity: more complex play led to more complex language, which led to more complex play, and so on.

The theory has two difficulties, however. First, the role of play in the complexity ratchet may not be as direct as the theory suggests: if play is a product of cooperation, then the ratchet could be between cooperation and language complexity, with increasingly complex play being a side-effect of increased cooperation. It would still seem that play is ratcheting with language complexity, but the correlation would lack causality. The second issue is the order of events: was self-domestication the initial driver, creating increased cooperation which generated extended childhood which generated complexity of play which generated language complexity, or is another causal chain likely? Langley *et al*. offer a useful list of components implicated in the sources of language grammar, but, as they say, more evidence is needed to establish a chain of causality.

For Christiansen & Chater (2022), humans are by nature playful and creative, and this is driven by a capacity to identify analogies between objects and events. Human cognition uses analogy extensively: for categorisation, for play (through homology), and for language (through metaphor). Humans have a deep-rooted capacity for identifying arbitrary connections between objects and events.

Christiansen & Chater see language as a collection of connected communicative games, which together form the language game. Just as a decathlon consists of ten different sports, so language consists of many different communicative activities; and, just as the nature of a decathlon cannot be properly specified by describing cycling alone, so language cannot be specified by describing only one aspect of the language game: the language game is not a set of rules, it is negotiation toward meaning. Despite the way we understand and teach languages, there is no such thing as a rule book for any language (and probably no rule book for the phenomenon of language); instead, a language is constantly growing and changing through interaction between its players; and it is this constant negotiation toward meaning that explains the wide variety of languages in the World. New language learners do not need an innate language capacity, because they are born into a culture of shared linguistic games. They learn their language by playing the games.

From this viewpoint, language is just putting words (signs which consist of gestures, sounds and meanings) together to create shared understanding. There are no stable rules about how the words must be put together, there are only conventions – which are often deliberately flouted to enhance understanding. Nor, contrary to the views of some linguists, is there a special cognitive system dedicated to language; there is nothing to be activated by exposure to language, there are only the communicative conventions to be followed or ignored. In *Star Wars* to an English speaker completely comprehensible Yoda is because speak English he does, despite of his words the unusual order.

The lack of a dedicated language system, or language faculty, in the human brain means that no ideal language is needed, neither as a target for a language learner nor as a template for Universal Language. Actual languages are not limited instantiations of a perfect language, whether that perfection is dictated by culture or by genes, they are attempts to share information. The communication of information is paramount, and to get it done we improvise, negotiate, reformulate, and generally tailor our language production to the needs of the communication event. We play the games of language when and if we need them, with the target being to get a useful facsimile of an idea from my head into your head. Language may be the most important human artefact, but it is not a pre-planned faculty; it is an unplanned side-effect of the need to share information, or simply to communicate.

Christiansen's & Chater's Language Game theory harmonises with the basic approach of this book. However, two unspoken assumptions in the theory need to be discussed here, because they raise issues that this book addresses. The first is the assumption that language is an interactive event, a sharing of experience, and therefore essentially for communication: there remains a valid counterargument that language is not primarily a communicative faculty, it is primarily cognitive. In contrast to the Language Game theory, Reboul (2015) suggests that language originally evolved for cognition and was only later exapted for externalised communication. Most language does not happen in the communicative space between minds, it happens in the heads of speakers and listeners, or writers and readers; and even this is only a small fraction of our thinking in language. The most common instantiation of language, the personal internal monologue, seldom makes it to the external language production system (Fodor, 2008). "What lunch tomorrow, that red funny, Napoleon horse Marengo what Wellington horse, Putin really look ill, what lobscouse mean, curiouser and curiouser …": each of us is constantly conducting a monologue as we move through the World; and hardly any of that monologue is communicative, unless talking to myself is communicative. But then, if talking to myself is the same as external communication, what language games am I playing with myself, and why? The communicative interaction of two or more brains is a form of play, the meanderings of a single brain less so.

The second assumption in the Language Game theory is that language handles all forms of information equally well. There is no need to consider the different structures of information because they are essentially all the same: no type of information is more privileged in terms of communication, they can all be shared with equal facility. Yet we know that, even today, language is better at some communicative tasks than others: for instance, accurate visual description is harder to convey than interpersonal relations. This points toward differences in the need for, and importance of, sharing different types of information, with the possibility that language was initially more specialised, generalising only later. If language started as a game for sharing a particular type of information, then the nature of that first information is significant.

Does treating language as play have any implications for the sources of language grammar? The key feature of play is its arbitrariness: rules are conditional and adaptable. Even where rules are a key feature of a game (for example, Monopoly) we find a plethora of *house rules* being generated for situations with no pre-agreed rule, or where the pre-agreed rule is inconvenient. What happens when the Bank runs out of money, or hotels, or houses? Can a player collect rents if they are in jail? Can more than six people play? The rules of Monopoly are negotiated as the game progresses, and the game can be seen quite differently by different players: if the current agreed rules match my idea of Monopoly, we are playing Monopoly; if they do not, we are playing a variant of Monopoly (Edwardes, 2002).

The only basic rule of play is that there should be agreement between players; every other rule can be negotiated or renegotiated. This allows play to take a range of different forms, from solitary, functional Ilinx and Alea to intentional, rule-driven Mimicry. Language grammar has this level of flexibility, but there are uncompromisable features. At a minimum, language grammar differentiates between object and action roles (nouns and verbs); and it seems to have a structure which lets objects be related together by actions as the causes and receivers of those actions – the subject and object roles of nouns (Allen & Saidel, 1998). Structure and ordering are important for sharing meaning in language, less so in play; and this is reflected in the timescales for comprehending and applying the rule systems of language and play. Grammar, the rule systems of language, is grasped sufficiently well for a child to use language effectively from about age four, while the rule systems of play are not usually mastered until about age six. If there is any dependency between acquiring these two rule systems then language formalism precedes formalism in play, and not the other way around.

There is another important way in which language grammar differs from the rules of play. Play is often a levelling event, encouraging self-handicapping to make everyone equal despite strength and skill differences; grammar, on the other hand, relies on hierarchies. Noun phrases can take the place of nouns and can themselves contain noun phrases; verb forms can contain other verbs to create nested meanings, such as *I began to try to find an answer*. Grammatical hierarchy creates the need for structured rules; hierarchy is not a basic requirement of play.

Language as a shared activity is certainly a form of play: individuals interact in a shared imaginary space where meaning is negotiated, not predetermined. Yet the rules of language are both less arbitrary and less coherent than the rules of play. Grammar emerges from a need to negotiate meanings into messages, easiest achieved by using cognitive forms for formulating meanings which are already shared between brains. However, the cognitive forms come from multiple sources, they do not rely on a single, coherent system; so, the language grammar they generate does not need a single, coherent system, either. This relationship between cognition and communication does not apply for play: language grammar uses systematic rules which play does not need.

Making Tools as a Source of Language Grammar

In chapter 1 we saw that tool use is unremarkable in nature: many primates crack nuts by picking up a rock and hitting a nut on a tree root, and some apes fish termites by picking up a stick and pushing it into a termite mound. Neither activity requires the tools to be made, only used; so perhaps it is toolmaking that distinguishes us from other animals. Unfortunately, once again, nonhumans have disproved our uniqueness. At first, there was evidence only of tool refinement: rather than picking up any stick, chimpanzees break off suitable twigs and strip off the leaves before using them for fishing (Goodall, 1988, 34-36). However, evidence of other tool creation began to accumulate. Among other examples, chimpanzees were seen chewing twigs to make brush-like tools to fish for honey in bee nests (Brewer & McGrew, 1990); chimpanzees have also been observed sharpening thin branches with their teeth and then using the branch to spear bushbabies in tree hollows as food (Pruetz & Bertolani, 2007).

From the archaeological record, it seems likely that the last common *Pan-Homo* ancestor was a toolmaker (Rolian & Carvalho, 2018). However, Toth and Schick (2009) show that there are four ways that human toolmaking differs from that of chimpanzees. First, humans knap stones to make sharp edges and thus improve the tool, where chimpanzees use stone tools as-is – any improvements are accidental. Second, both species transport tools to where they are to be used; but where chimpanzees transport them for tens of metres, humans transport them for tens of kilometres, which implies greater selectivity in the tool material. Third, chimpanzees mostly use tools to extract food from plants, where most early human tools involved extracting food from carcasses. Finally, chimpanzee tool curation (saving a tool for later reuse) is lackadaisical, where human tool curation is longer-term and more deliberate.

Toolmaking is not limited to humans and chimpanzees. For instance, beavers construct their environment by bringing down trees and tree branches to dam rivers and create large pools; but in one example there seems to be evidence that a beaver (*Castor canadensis*) cut and moved a willow branch log to use as a step, enabling it to work at a greater height to bring down smaller, more manoeuvrable branches (Barnes, 2005). Impressive examples of nonhuman toolmaking have been demonstrated by New Caledonian crows (*Corvus moneduloides*). These birds have been observed making various tools, both in captivity and in the wild. In the wild they cut pandanus leaves into various types of insect spears with their beaks, some for catching grubs in the ground, some for grubs inside trees (Hunt & Gray, 2004). In laboratory conditions they were observed bending metal wires into hooks to lift a container of food out of an otherwise-inaccessible pipe (Weir *et al.*, 2002). New Caledonian crows show that planning and manufacturing tools is not exclusively human, nor even exclusively mammalian.

Yet tools made by humans are identifiably different from those of other animals in complexity of form, purpose and manufacture (Byrne, 2004); and this is especially so for stone tools (Davidson & McGrew, 2005). We make composite tools, often out of multiple materials; we make multi-functional tools; and we make tools to make other tools, sometimes involving multiple levels of manufacture to create the final product. This capacity to plan a tool by breaking its production down into discrete steps does seem to be peculiarly human; but it also seems to be something we developed over a considerable timescale (Henshilwood & Marean, 2003). It is not, therefore, a single genetic event creating a new primary capacity; instead, it emerges slowly from other cognitive capacities. It is possible that toolmaking emerged from the same group of primary cognitive capacities as language, but it did not follow the same trajectory: toolmaking is an essentially solitary activity, relying at each stage on the genius of one brain; language is social, relying on negotiation between brains, and it should therefore be developmentally slower. In practice, language seems to innovate frequently and change rapidly, while innovation in tool creation was, until recently, a slow process.

Moore (2010) describes how the strategies used in organising utterances are the same as those used in tool manufacture. He calls these, in the terminology of Greenfield (1991), *grammars of action*, and shows how, in lithic manufacture, low-order actions (*e.g.*, flake removal) are combined in the correct order to produce a high-order tool – in the same way that lower-order words are combined to make higher-order utterances. However, this hierarchy of strategies is not limited to lithic technology and language, it is common throughout human cognition – which makes it difficult to identify a primary cause. Using a hierarchy of strategies in lithic technology very likely preceded its use in language, simply because lithic technology happened first; but whether lithic technology was the first cognitive activity to use a hierarchy of strategies is not clear.

To make any tool you must have an idea of the shape and size of the finished tool, and a knowledge of how to alter the base object into the final object. This conversion from raw material to finished object is a complex skill involving planning, visualisation, intention, and the ability to constantly triangulate from the current state of the unfinished tool to the desired outcome; it is a skill that has clear implications for general cognition. In contrast, the material used to make the tool requires cognitive capacities specific to the material, and which are not readily transferable to other materials or other cognitive processes. Knowing how to bash rocks together effectively is a remarkable skill, but can it provide direct help in solving other lifestyle problems?

Figure 2.1: Oldowan knapped stone tools, Early (about 2.5mya –from Guelmim-Es Semara region of Morocco; Álvarez, for the National Archaeological Museum of Madrid, 2007) and Later (about 1.9mya – Duval *et al.*, 2021).

Hovers (2012) points to one significant feature which does, indeed, affect other lifestyle skills: creativity. The development of lithic skills in the Oldowan is incremental, both at the individual level through learning and at the group level through teaching. However, while there is a clear increase in toolmaking skills through the Oldowan, making later lithic assemblies both more effective as tools and more elegant as objects, it is difficult to identify this process of sophistication at work. By comparing the tools produced at the beginning of the Oldowan with those produced toward the end, we can see that creativity must have occurred; but the biological evolution of creativity throughout the period is a matter of invention, loss, and reinvention of skills; and the cultural evolution is a matter of innovation, attrition, and rediscovery of method. Nonetheless, we can detect important crossovers from tool skills to other cognition when we look at dietary changes and systematicity of behaviours. The cognitive skills required to bash rocks together effectively do, indeed, seem to have helped in solving other lifestyle problems.

Despite evidence for some technological advance in toolmaking during the Oldowan period (which included both Australopithecines and early *Homo*), the next period of stone tool manufacture, the Acheulean, involved a stone-knapping style that seemed to remain unchanged for a million years (Ambrose, 2001). The ideal stone tool shape produced in the Acheulean period seems to have been a tear-shaped oval, knapped out of a flint cobble or similar hard stone. Quite why that shape was so favoured remains a mystery, but during those million years, the Australopithecines disappeared and many species of *Homo* (*habilis*, *ergaster*, *erectus*, *heidelbergensis*) came and went; brains expanded from 600cc to modern sizes (1350cc) and beyond; and the first *Homo sapiens* individuals appeared. Either the Acheulean handaxe was so useful that improved cognition could find no better tool, or the improvements in cognition during the Acheulean were not tool related.

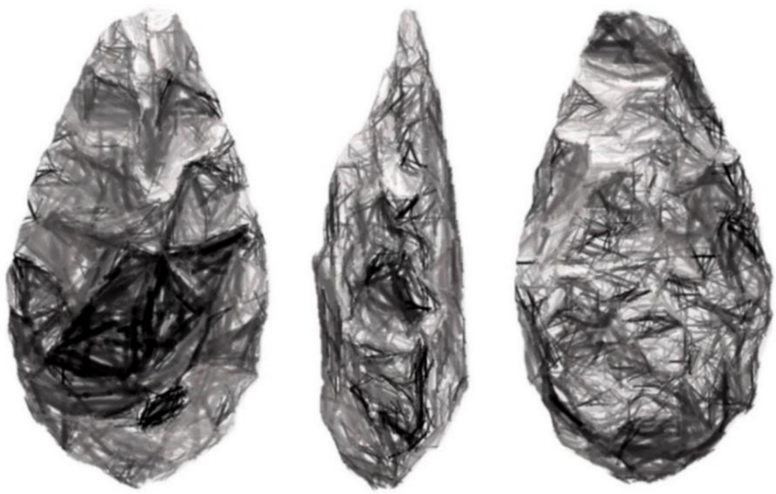

Figure 2.2: A prototypical Acheulean teardrop handaxe. Front, side and back views (Modern re-creation, Key & Dunmore, 2018).

Making tools and making language both involve the construction and assembly of components to produce complex objects and outcomes. In language-making, like toolmaking, the products of one level become the tools or components at the next level; and both require modular and hierarchical planning to complete the intended final product. Neither toolmaking nor language-making are themselves target products; they are, instead, the way that target products (things like consumable food or shared information) are achieved. In this respect, language-making is not just *like* toolmaking, it *is* toolmaking (Stout & Chaminade, 2009). Human toolmaking is certainly special in evolutionary terms, and it does seem to use similar, if not the same, cortical systems as language grammar (Gabrić *et al*., 2018); but, while the cognition of toolmaking can be convincingly implicated in the forms that language grammar takes, its role as a driver of language grammar development remains tenuous.

Treating language as toolmaking has important implications for the evolution of grammar. If the cognitive processes of toolmaking are applicable to general cognition and communication, then the genesis of grammar is unremarkable: those processes allowing us to construct complex tools are the same processes allowing us to construct complex language. Language as toolmaking does seem to provide all the features we need for language grammar: the capacity to analyse a problem into its different components; the capacity to solve the individual components separately; and the capacity to resolve the main problem by integrating the individual solutions in correct order. The same cognitive processes generate both complex toolmaking and complex information-sharing.

The fossil evidence places the appearance of complex tools (that is, tools composed of multiple components made from multiple materials) at about 40kya, at the start of the Aurignacian (Gravina *et al*., 2005) – or earlier, at about 50kya, if Neanderthal evidence is included (Niekus *et al.*, 2019). Before this, the constructed tool set consisted mostly of carefully knapped and shaped stones, essentially limited to hammers, cutters and scrapers; in the Aurignacian, it became a mix of bone and stone punches, drills, saws, even needles, as well as an expanded range of specialised hammers, cutters and scrapers. Very few wooden artefacts have survived, so we can only speculate on the portability of toolmaking methods between materials, but the same level of dexterity would have been available for both. However, as Stout (2011) points out, making composite tools requires considerable dexterity, and that relies on the development of appropriate physical and cognitive skills. effective dexterity cannot appear overnight, it is the product of a long chain of genetic changes which probably began about 700kya.

The Aurignacian was not the first change in tool technology after the million-year technological standstill of the Acheulean, there seems to have been another technological event between the Acheulean and the Aurignacian. This event changed the quality and range of tools made, and seems to have propagated relatively quickly through existing human groups. This technology, known as Levallois-Mousterian, was in use during the period from 350 to 40kya. Traditionally, Mousterian tools were produced by Neanderthals while Levallois tools were produced by *Homo sapiens,* but this is a somewhat arbitrary division which is not supported by all archaeologists. They are treated here as a single technology, using the neutral term *prepared core technology*.

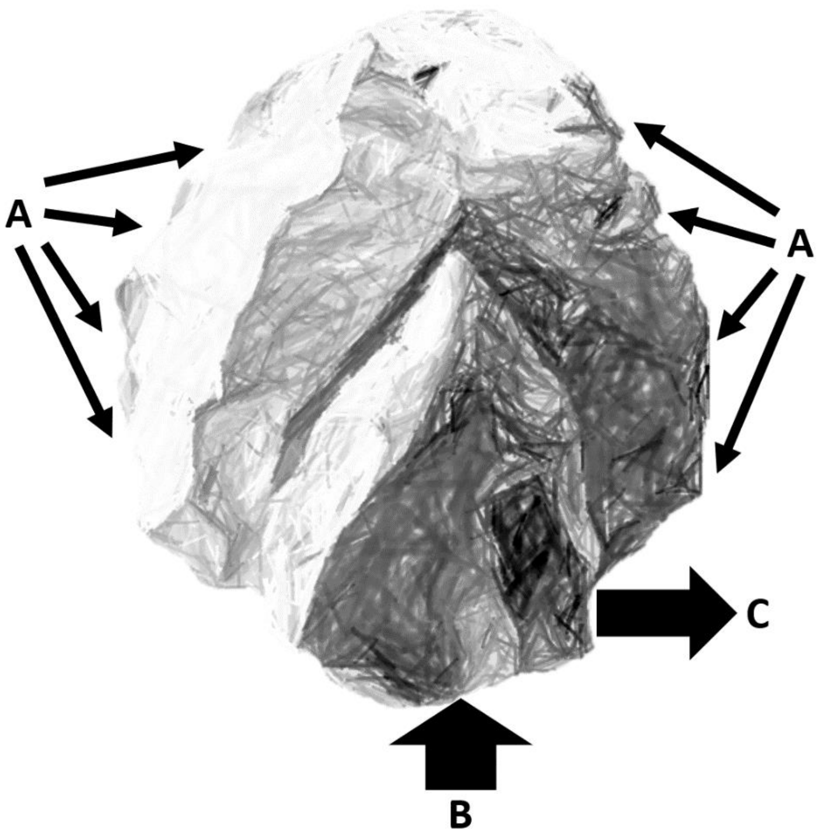

Figure 2.3: The Prepared Core method. Flakes are struck from the raw cobble (A). This leaves a bulge which is given a single, placed blow (B), detaching a blade or point (C).

Prepared core technology was essentially a stone tool industry, although with increasing use of other materials. Unlike the Acheulean stone axe technology, the prepared core method of toolmaking was a two-stage operation: first, flakes were struck from the edges of the raw cobble, leaving a bulge in the centre of the modified cobble; the edge of the bulge was then given a particular type of blow at a particular position, and a large bladelike flake, or a sharp pointed flake, was detached from the core cobble (see Figure 2.3). The removal of the flake often revealed a second bulge from which a second blade or point could be struck, and so on (White & Ashton, 2003). There is evidence that earlier knapped and discarded Acheulean axes were reworked into Levallois-Mousterian blades and points, indicating that the two traditions may be linked, and that the prepared core technique could have been a logical extension of the Acheulean method (Shimelmitz *et al*., 2016).

Prepared core techniques propagated through the human population too fast to have been evolutionary. They seem more like a teaching-and-learning event because the speed of transfer between individuals reflects a contagious rather than an inherited pattern. Adler *et al*. (2014) suggest that the technological propagation probably involved multiple, independent, almost-simultaneous innovations rather than just one; prepared core technology was, more a logical progression from existing technology than a *de novo* event. This hypothesis is supported by the number of local cultural technologies we designate as prepared core. Carbonell *et al*. (2016) show that prepared core technologies produced new types of tools, leading to multiple tool sets and cultural differences in technological solutions.

Contagious propagation implies that mechanisms for widespread knowledge transmission must have existed before prepared core technology. The communication system must have been complex enough to both negotiate the exchange of knowledge and to demonstrate the technological complexity of the new tools – in short, it would have required a pre-existing language-like grammatical complexity (Szathmáry and Számadó, 2008). Even if we claim that this communication system was not initially “proper” language, it is unlikely that the communication system available at the start of the Aurignacian was not already complex. The hypothesis that humans underwent a technological and linguistic revolution in the Upper Palaeolithic of 40kya has been replaced by a longer timescale (Hovers & Belfer-Cohen, 2013) and a slower technological timetable (Burdukiewicz, 2014) which envisions a complex and braided incremental development of language, technology, culture, cooperation and domestication (McBrearty, 2007). Language is a form of toolmaking, yes; but that does not necessarily mean that language descended from toolmaking.

Fitness Signalling as a Source of Language Grammar

A problem that perplexed Darwin was the peacock's tail: how could something so elaborate and without practical function have evolved just to make the peacock look good to the peahen? According to evolutionary theory, attractiveness should be demonstrated by obvious signs of practical fitness, not by non-functional arbitrary beauty (Darwin, 1859, 87-90). And yet there he was, the male peacock, dragging around a train that seemed to make him less able to escape predators. Why were peahens so interested in it? The answer was finally provided over 130 years later: the peacock's tail is a cost to the male and therefore reliably indicates to the female that this male can pay the cost and still thrive (Zahavi & Zahavi, 1997). Natural selection favours females who choose males with elaborate tails because their offspring will tend to be fitter; and the tendency toward elaborate tails will therefore be positively selected for by both the choices of females and the necessary fitness needed by males to successfully carry the tails.

Could language be a similar costly signal of fitness? Were talkative partners favoured because a capacity and willingness to talk represented an arbitrary cost demonstrating the reproductive value of the talker? It is certainly true that language represents a high cost in many ways. Cognitively, language requires the capacity to use both rules of language (grammar) and component labels (words); it requires the capacity to quickly formulate acceptable utterances to express intended meaning; and it also requires both control over the rapid and complex movements needed for speech articulation, and the capacity to interpret a stream of sound into a stream of meanings. Being able to meet these cognitive costs is significant even today: effective talking and listening are valued in careers which involve leadership, entertainment, teaching, and a range of other interpersonal activities.

However, being a good talker is not the only route to social (and therefore reproductive) success. The traditional markers of fitness – physical skill, motor dexterity, creativity – are valued as highly or, in some contexts, more highly than speech skills – as the incomes of some sports people, musicians and entrepreneurs demonstrate. Talking, unlike the peacock's tail, does not by itself seem to be a certain route to reproductive success.

If language really were a costly signal of fitness then we would expect it to be a one-way signal: it would be unequally distributed between the sexes, and it would predominate in the sex that has the lowest reproductive cost. The peahen does not need an ornate tail because she pays all the costs of incubating and protecting the young – effectively, she has control over the process of reproduction. In contrast, the male only contributes sperm, which is plentiful and cheap. The fitness of the female is of minor importance to the male – impregnating one female does not limit his opportunities to impregnate others. The fitness of the male is, however, of paramount importance to the female. If the wrong male impregnates her then she has an invidious choice: she can make the best of the bad job, hoping for a better male next time; or she can abandon the eggs or offspring and move immediately on to the next time. If she cuts her losses, however, she is disregarding the costly commitment of time and resources she has already made to reproduction (Campbell, 1999). It is in the female's interest to discover the male's fitness before copulation, and therefore in the male's interest to display it to gain access to reproduction; and that is why male peacocks have magnificent trains, and females do not.

This separation of roles by gender is not something we see in human language. While subtle statistical differences have been found between male and female language use (Eckert & McConnell-Ginet, 1998), there is no marked separation of language by gender. Both male and female humans seem able to use the full range of language, and there are no forms that are clear to one gender and opaque to the other.

If we look at the ways we share meanings in language, rather than the ways that sharing itself works, then we are in even more difficult territory. By sharing meanings the sender is making information available to the receiver; information which, if true, could be valuable to the receiver because it may enhance their fitness. However, the value of the information to the receiver is a reason why the sender should not share it: if the receiver is a reproductive rival then giving information enhances the rival's fitness by removing the sender's advantage of knowing something the receiver does not. This means that sharing information could be considered a costly signal, like the peacock's tail; but to be a costly signal it must be done ostentatiously in the presence of prospective reproductive partners. Sharing information as a costly signal only makes sense if the receiver is the same sex as the sender and there are members of the other sex to see the cost being paid. This is a very specific way of signalling which explains only a fraction of what humans do with language: we often share information privately in pairs, or in groups containing only individuals of our own sex; and, most importantly, we share not just valuable true meanings but apparently valueless fictions.

The fictional nature of many language utterances poses a problem for utilitarian approaches to language origins, such as costly signalling. Why would a receiver be happy to be told lies? Humans like being told stories, and good storytellers are socially valued, as George R.R. Martin can attest. It seems likely that storytelling was one of the first uses to which language was put, and it may even be one of the primary sources of language. Certainly, the entertainment industries nowadays form a significant component of the world economy. This is explored further in chapter 9, because any theory of the sources of language grammar should explain this apparent fitness anomaly.

The stability of honest signalling in human language has recently come under scrutiny. Dessalles (2014b) reviews how sociability creates a stable environment for sharing honest information, and he finds that asymmetrical societies – groups where rivalries outweigh friendly cooperation – tend to produce two types of individual: those who signal frequently and those who seldom signal. The fitter individuals are willing to signal their fitness, while the less fit individuals cannot compete so do not. This is not what we see with human language; instead, we have symmetrical societies, where friendships drive an informal sharing of information because friends share time together. With shared time, friends build relationships and alliances, work together in joint ventures, and use language to negotiate toward meaning, rather than trying to individually dominate meaning. In a symmetrical society, signalling does not need to be extrinsically honest, but each utterance (or negotiation) should be contextually honest to build and preserve the friendship: the receiver must believe that the sender believes they are being honest, even if the receiver knows otherwise.

Another approach to the honesty problem looks at altruistic punishment as a way of imposing costs on deceptive signallers. Altruistic punishment is a feature of organised societies which involves an otherwise uninvolved third party punishing a deceiver on behalf of the deceived. A study by Clark & Kimbrough (2017) modelled a society with individuals of three different levels of aggressiveness, and it found that deception can be eliminated where altruistic punishment is common and severe. However, where it is less severe and the society is stratified into groups of low, medium and high aggression, deception can persist. It tends to die out in the middle stratum of the society, but continues in the lower and higher strata, although at a reduced level. As human societies are often stratified, and altruistic punishment is often delegated by the many to the few (*e.g.*, to police forces, security forces, lawmakers, *etc*.), we would expect that deceptive signalling is a stable part of human communication, especially in the privileged and disempowered social factions – and that does indeed seem to be the case in human societies.

Barker *et al*. (2019) take the view that emphasising costly signalling as a source of honesty has skewed our understanding of human language: the costs of being discovered as dishonest are often greater than the costs of being honest, because reputation is a key feature of human societies. There may be non-negligible short-term advantages to dishonesty; but reputations are built over time by a series of individually trivial actions, and loss of reputation can wipe out years of reputation-building in days. It is better to default to honesty, not because it is a signal of fitness, but because discovered dishonesty is a powerful signal of unfitness.

These studies show that, while costly signalling is important for signalling in general, its role in human communication has been muted by social factors favouring honesty over dishonesty. Perhaps, though, the complexity of human language grammar is itself the costly signal: just as the complexity of some birdsong seems to act as a viable costly signal (Vehrencamp, 2000), so the capacity to use complex language may itself be a signal of fitness.

What does this tell us about the nature of grammar? If complexity of form is an indicator of cost then producing complex utterances should be an indicator of fitness. We would expect humans, especially males, to use complex forms whenever possible; complexity should be valued and rewarded with reproductive opportunities, as happens with songbirds. We would expect talking competitions to be commonplace, and we would expect the ability to produce complex utterances to be positively correlated with breeding success. This does not seem to be the case. Instead, we seem to value clarity over complexity, and prefer simpler utterances over complex ones (Denton, 2006). Inappropriate complexity is seen as an unattractive trait, and unavoidable complexity is viewed as a necessary evil.

An important feature of language is that we use it to work with people rather than against them – we value cooperative dialogue over competitive argument. If there is cost to the speaker in a language utterance it seems to lie in the meanings offered – the information being giving away – rather than the complexity of form (Buzing *et al.*, 2005). While there is a physical cost in using grammar to produce utterances containing complex meanings, it is the complexity of the meanings that drives the need for complex grammar – the one-dimensional stream of speech must be capable of expressing ideas which are often multidimensional. In fact, expressing complex meanings in simple forms is what we consider most cooperative.

If language is a costly signal of reproductive fitness then complexity is an indicator of cognitive cost, and duration of utterance is an indicator of physical cost; and these are what we would value in a speaker. This, however, is birdsong, not language. Instead, there seems to be a different costly signal, cooperation, for which language and grammar are necessary tools; and it is the complexity of signalling required to facilitate human cooperation that drives the complexity of grammar (Kirby, 1998).

Embodiment as a Source of Language Grammar

Language is a function of physical bodies. It is tied to bodies both by the sensory input needed to generate meaning, and by its role as a communication system (Tibbetts, 2014). Human language is intimately linked to the human body, and it calls upon a range of embodied systems to support its existence (Arbib, 2008).

The first of these embodied systems is sound production. Language is primarily seen as an instrument of sound, which we produce using a series of orofacial gestures; these gestures convert meanings in the speaker's brain into sounds which can then be interpreted in the listener's brain back into meanings. Language meanings involves conscious cognition at both ends; but producing and apprehending sounds is mostly subliminal and automatic. When we hear language we hear the sounds subconsciously – accents and local pronunciations are somehow standardised in our brains before we consciously "hear" what has been "said". This standardisation is an inherent part of our vocalisation and auditory systems. For instance, we may consciously hear the words "I'm in position" when what was actually produced and heard was "I mim perzishun"; or, if in Devon, "Oi min purrzizhon". We may be aware that the two versions sound different, but we apprehend the meanings of the sounds, and therefore the words, as identical.

The vocalisation systems used in language developed slowly over a long period, which started before the first humanlike clade, *Australopithecus*, appeared. The gestural facility of the lips, cheeks and frontal mouth parts began to develop early in the evolution of ape vocal communication: orangutans use a volitional kiss-squeak call to warn a predator that it has been seen, and to warn any other orangutans in the area that a predator is around (Lameira & Call, 2018); and chimpanzees use lip smacks (as well as kissing noises, plosives and nasals) in their social sound-making (Fedurek *et al.*, 2015). *Ardipithecus ramidus*, an intermediate species between the last common *Pan-Homo* ancestor and *Australopithecus*, had several adaptations which made them more effective vocalisers than chimpanzees: reduced face and jaw, a flexed cranial base, and probably a larynx which was deeper in the neck (Clark & Henneberg, 2017). All these species – orangutans, chimpanzees, *Ardipithecus* and *Australopithecus* – seem to have had some volitional control over the lips, cheeks and frontal mouth parts used in sound-making.

Bipedalism, discussed in chapter 1, was the next evolutionary feature to undergo selective pressure. The Australopithecines seem to have been the first clade of habitual bipeds (they regularly walked on two legs), but whether they were obligate bipeds (walking on two legs was their most efficient form of movement) remains unknown; their longer arms and chimpanzee-like wrist joints mean they probably used brachiation (swinging from branch to branch using their arms) frequently (Crompton *et al.*, 2008). Bipedalism did, however, lead to increased conscious management of the chest muscles controlling the lungs, along with greater control over airflow; and these became major drivers for increasing complexity of vocalisation and, eventually, language in later clades.

The chest cavity, lungs and air flow are key components of the chimpanzee pant-hoot call; but chimpanzees have less control over their breathing than humans, partly because they have retained the capacity to breathe and swallow at the same time without choking (Lewin, 2005, 222-223). This is one reason for the failure to teach chimpanzees – most notably Gua (Kellogg & Kellogg, 1933 [1967]) and Viki (Hayes & Nissen, 1971) – to vocalise in the same way as humans: humans have greater conscious breath control, and therefore greater control over the loudness and duration of our vocalisations.

Control of the tongue & vocal cords seems to have developed relatively late in our evolution. For chimpanzees, the tongue and vocal cords just add volume and tone; their main articulators for sound-signalling are the lips, cheeks, frontal mouth parts, chest cavity and lungs – although they have less control over these articulators than modern humans. Like chimpanzees, Australopithecines are unlikely to have had refined control of the tongue and vocal cords, simply because their cognitive motor systems seem less developed than those of the later *Homo* clade. However, we cannot be certain of this because brain, tongue & vocal cords are soft tissue and do not fossilise. There is some indication of the range of movements possible from the muscle anchor points in the jaws of fossils, but these are sketchy and open to interpretation. We have only one good clue to the level of vocal control: in both *Homo sapiens* and Neanderthals, the hyoid bone is reduced, which indicates greater muscle control (Nishimura *et al.*, 2006). This may mean that their common ancestor, probably *Homo erectus*, could have spoken in a language-like way (Dediu & Levinson, 2013); but whether they did so remains disputed. Belyk *et al.* (2021) have recently shown that there seem to be two laryngeal motor control areas in the human brain, and that each area coordinates both laryngeal and respiratory control. Whether this was true of *Homo erectus* is unknown, but a lack of this dual control system would have affected vocalisation.

Why human vocalisation developed in the way it did is another issue. One proposal is that it developed for singing (Mithen, 2005). Singing is used by many species, including some apes (Geissmann, 2000), and for many purposes: it is used to indicate personal fitness (arguably the main purpose of birdsong), as a way of signalling location (the main reason for solo calls), to warn predators and rivals of the size of your group (the main reason for chorusing – Knight & Lewis, 2017), and to build and maintain social relationships (the probable reason for anitiphonal singing – Jordan *et al.*, 2004). If singing is a costly signal indicating fitness then there is likely to be an evolutionary race toward song complexity – the more complex the song, the fitter the individual, and the more breeding success they have (Locke, 2017). This is supported by a recent study of the highly variable and multimodal displays of birds of paradise, which involve singing, dancing, and a decorated bower (Ligon *et al.*, 2018).

Just as important as the production systems for signalling are the receiving systems. These are often neglected, but there are important differences between chimpanzee and human hearing which seem to be language-related (Quam *et al.*, 2012; Belin, 2006). There is even some evidence of hearing differences at the genetic level: as Clark *et al.* (2003, 1,962) state, "The gene with the most significant pattern of human-specific positive selection is alpha tectorin, whose protein product plays a vital role in the tectorial membrane of the inner ear." This seems to affect hearing acuity, and possibly tonal discrimination, giving humans a better capacity for differentiating between subtle sound contrasts, such as voiced and unvoiced. Differences have also been found between the auditory systems of *Homo sapiens* and Neanderthals, although it is less clear what they signify (Gómez-Olivencia *et al*., 2015).

It seems likely, therefore, that the changes to the auditory system occurred continuously over the whole evolutionary period from the last common *Pan-Homo* ancestor to *Homo sapiens*, closely following the increasing sophistication in the vocalisation systems. There was not a single mutation that made our hearing system speech-friendly, there was cumulative, non-neutral evolution driving us toward speech perception.

In language, context is everything: the same signal delivered by a different individual can have a different meaning because the sender and receiver are important parts of the signal context. The ability to quickly identify individuals is an important capacity in any social group, and particularly so if the social group has developed language. Language is not just about the signal, or the meaning of the message in the signal, it involves the status and reputation of the sender and the receiver, and the historical relationship between them (Molleman *et al*., 2013). This means that it is important for the sender and receiver to identify each other, so they know who they are dealing with; it pays both parties to be able to recognise other individuals. In the case of humans, it is useful to recognise individuals by their vocalisation as well as their face.

Fortunately, both facial recognition (Parr, 2011) and vocal recognition (Belin, 2006) seem to be ancient cognitive features in the primate clade, although there does seem to be increasing sophistication of recognition from monkeys through apes through the *Pan* species to humans. Primates do not just hear a signal and respond to it; there is evidence that they relate the signal to the signaller and respond appropriately depending on their social relationship with the signaller (Engh *et al*., 2006). There is evidence for a primate facial recognition system in the superior collicus (Le *et al*., 2020), a brain area which, in primates, is organised differently from other mammals. However, caution is needed when comparing the recognition systems of different species: for instance, bottoms, not faces or voices, seem to be the primary way that chimpanzees identify each other (Kret & Tomonaga, 2016).

Facial expression is another embodied gestural system used to signal intention. Darwin (1897) thought anger, fear, disgust, happiness, sadness, and surprise were the six core human expressions with universal meaning (i.e., they are genetically based and non-volitional). However, there are other facial signals which seem to have universal form, like laughter (Sauter *et al.*, 2010), yawning (which is common in many species and may be a subliminal signal to increase alertness – Gallup, 2007), the eyebrow flash of recognition and the quizzical furrowed brow (Grammer *et al.*, 1988), and the knitted brow of pain or grief (Morris, 1994). There are also some facial expressions (often deceptive or group-based signals) which are culturally defined (*e.g.*, covering the face with an open-palmed hand can have different meanings, depending on culture and context). Different cultures also read different parts of the face to interpret expressive meaning (Gendron *et al.*, 2014). Rather than Darwin's six core expressions, we now recognise four universal expressions: happiness; sadness; fear/surprise; and anger/disgust. The expressions for fear and surprise, and for anger and disgust, seem to be culturally interchangeable (Jack *et al.*, 2012).

We used to believe that human facial expression was much more communicative than that of our nearest relatives. We now know that, in terms of musculature, we are very similar. Chimpanzee expressions may appear to us to be limited, but that is largely because differences in bone structure affect muscle responsiveness (Burrows *et al.*, 2006) – and also because, while they use expressions similar to ours, they do not always mean the same thing. Most notably, the chimpanzee exposure-of-teeth display may look like a smile; but it is, like the dog exposure-of-teeth display, a sign of anger or fear (Parr & Waller, 2006).

Gesture is commonly used in communication systems throughout nature; but, while many animals use simple iconic gestures to communicate, the use of semantically complex gestures seems to be limited to humans. Chimpanzees produce a limited number of physically complex gestures, but they usually have simple and invariant meanings; which may be because producing reliable complex gestures relies in large part on obligate bipedalism freeing the hands for greater communicative activity (Schmitt, 2003).

Chimpanzees do use a range of kinaesthetic signals, of which Hobaiter & Byrne (2014) have identified at least 66. Of these, many seem to be genetic conventions (the same across all groups); but some, like the Mahale grooming handclasp (McGrew *et al.*, 2001), are group-specific and therefore cultural conventions. In contrast, human signals are mostly cultural and therefore symbolic, and liable to be interpreted differently in different cultures. This is not because humans have fewer non-volitional signals, we have largely the same range as chimpanzees; it is because we have many more volitional signals (Cartmill *et al.*, 2012). Because our gesture system is mostly cultural and volitional, humans have a gestural signalling channel which can be as rich as speech; and this is why deaf sign languages are now treated in linguistics as full languages with an integral gestural phonology (Mann *et al.*, 2010).

Modern deaf community signed languages are languages by other means; and there should be no doubt here, they are actual languages. Signed utterances can convey the same meanings with similar complexity of form as spoken utterances. In terms of grammar, sign languages have constructs to express time relationships and non-present and non-existent events, and they can link utterances together into metaconstructs. Signed languages are segmented (they use word forms); they are differentiated (they use different types of words, such as nouns, verbs and adpositions[3]); they are hierarchical (they have adjectives dependent upon nouns, adverbs dependent upon verbs, and noun phrases which contain noun phrases); and they are rule-bound (the same basic form is used for most utterances, and order is significant) (Kyle & Woll, 1985).

Gesture, however, allows sign languages to have real place-marking. There is no "over-thereness" about the spoken words *over there*, but the sign

[3] There are two types of adposition: prepositions, which occur before its noun phrase; and postpositions, which occur after the noun phrase. Many languages use one or the other as the dominant form. The term *adposition* is used here to indicate both forms.

language equivalent can indicate approximately where *there* is. *Up* is up and *down* is down, *in front* and *behind* are where you would expect. Sign language puts events onto a virtual stage in front of the receiver, something that spoken language cannot easily replicate – and that is why spoken language is often supported extensively by gestures (Hanks, 2005).

This use of gestures to fill the gap between utterance and actuality has caused some commentators to see speech accompanied by gestures as a possible halfway house between non-symbolic pre-human signalling and symbolic language. Ape vocalisations are heavily constrained, used mainly to express intuitive reactions, and seem to be under only sporadic volitional control. In contrast, the dexterity of apes is notable, and they appear to have quite proficient control over their gestures (Arbib, 2005). Could it be that the volitional symbology of language first appeared in the gestural mode, and only later converted to the vocal mode?

This idea has been intensively explored (Iverson & Goldin-Meadow, 1998; Corballis, 2002; Arbib *et al.*, 2008, among others), and a canonical developmental process seems to have emerged. Initially, volitional gesture developed slowly in Australopithecines to meet a range of different signalling needs, such as hunting; and gestural signalling remains important in human hunting cultures today (Lewis, 2009). At some stage the vocal channel became subject to greater conscious control – possibly because, like birds, singing became a signal of fitness (Mithen, 2005, ch9) – and it became possible to generate meaning-rich vocal gestures as well as manual gestures. At that stage, vocalisation became the primary channel for exchanging meanings, with gesture remaining important, but secondary (Steklis & Harnad, 1976).

What does all this embodied gesture imply for grammar? Unfortunately, not much: if both gesture and vocalisation can use grammar of equal complexity, the channel used does not indicate when or how the complexities arose. The mode of the signal is transparent when we look for sources of language grammar – and, to a certain extent, it is unimportant. The cognitive capacities that allowed vocalisation to become volitional and controllable seem unrelated to the cognitive capacities behind language grammar. Language is embodied and therefore, inevitably, partly gestural; and it may well have evolved from a completely gestural signalling system; but the grammar used in language is modality-independent, and its development is unlikely to have been affected by changes in the signalling channel.

Multimodal Signalling as a Source of Language Grammar

Because language is embodied it also has the potential to be mode-independent: the plasticity of the human brain (Willis *et al.*, 2009) means motor gestures can express meanings in any channel, and the listener can apprehend and interpret those motor gestures back into meanings (Corballis, 2003, ch9). Your reading of this page is evidence of this: you are interpreting meanings from ink marks on paper or contrast effects on a screen, produced in the first instance by my fingers pressing keys, or by me speaking into a microphone. By themselves the motor gestures and marks are meaningless, but we have been able to learn conventions which imbue them with meaning. Cognitive plasticity means that language can be produced and interpreted as transient sounds, transient motor gestures, more durable disembodied marks, or combinations of the three.

However, multimodality is not just about language use in different modalities, it also involves different modalities combining together within a single signal (Kendon, 2009). Birds, for instance, use combinations of sound and gesture in their mating displays, integrating the two modes so they augment each other's message (Cooper & Goller, 2004). This multimodality, however, is not quite the same as human multimodality: humans can negotiate the same meaning in different modes, whereas the multimodality of birds seems to encode different parts of the signal meaning to each mode. Male bowerbirds must successfully integrate their song, dance, and the richness of their bower if they wish to be successful in their mating display.

In contrast, captive chimpanzees and bonobos create and use a wide range of gestural and vocal signals with overlapping meanings, allowing either or both modes to represent meaning; and they negotiate these signals, using cultural norms to agree meanings (Roffman *et al*., 2015). This is similar to human multimodality, which is negotiated rather being innate; but *Pan* species signalling seems to be more iconic, with vocal signs relying heavily on innate emotive calls (*e.g.*, for fear, rage, enjoyment, surprise, *etc*.), and gestural signs being more pantomimic and less representational.

Gillespie-Lynch *et al*. (2014) see similarities between the gestures of pre-adult *Pan* species and human children as evidence of a common set of non-verbal capacities. They propose an ontogenetic developmental sequence from pantomimic gesture to symbolic gesture as evidence that human symbolic gesture emerged from iconic gesture via indexical gesture. These gestural modes were simultaneously available to the vocal channel – in other words, human language evolved in a multimodal way. Zlatev *et al*. (2017) take a similar approach, but they see pantomime as a separate from iconic gesture: it was not initially available to the vocal mode, and only became available when pre-humans began to see other bodies as intentionally communicative. Pantomime can tell you what to do, but intentional iconic gesture can tell you why to do it.

For Levinson & Holler (2014), the development of human multimodal communication is a layered process. Ritualised gestures in the last common *Pan-Homo* ancestor led to pointing and iconic gestures in early *Homo*, accompanied by systems for interpreting facial gestures, joint attention and communicative intent. Voluntary vocal utterances were added by the last common ancestor of modern humans and Neanderthals, leading to full modern language capacities in *Homo sapiens*, and possibly Neanderthals and other species descended from their last common ancestor. Each new layer relied on, and added to, the previously existing structure. Gesture did not replace speech at any stage, nor did speech replace gesture; they existed together, and each provided scaffolding to enable the other to develop.

In contrast, Fröhlich *et al*. (2019) see the vocal and gestural modes of communication as active in all the other Great Apes as well as modern humans, but they notice a difference in emphasis: where gesture seems to be the main information mode for other Great Apes in proximate communication, in humans it is the vocal mode. For Fröhlich *et al*. this indicates that there has been an evolutionary shift in how information is shared; but it may instead indicate a shift in what information is shared. Humans tell each other about past events, speculate together about future events, and entertain each other with fanciful events; and all these irrealities are easier shared vocally than gesturally (Zdrazilova *et al.*, 2018).

Kang & Tversky (2016) show that, in current, multimodal language communication, gesture adds facility which is difficult to achieve in speech. Where speech is essentially stative, it tells you about how things are, gesture is processive, it tells you how things move and how actions change things. When I was young, an enlightened English teacher held a special lesson: telling a Pacific islander how to tie a tie. The scenario was that the class and the tie-tying islander were in contact using old-fashioned voice-only phones (or, as we used to call them, phones), and we had to describe the actions needed to produce a reasonable knotted tie. The teacher stood at the front of the classroom and carried out the instructions we gave. The resulting mess was both hilarious and frustrating for the class, and no ties got tied that day. The lesson, however, was well-learned: a few simple gestures can be worth any number of words.

Working with actors at a sign language theatre in Israel, Sandler (2022) showed that multimodality can occur in a single modality, by asking a simple question: where does gesture go if the gestural mode is busy producing sign language? The answer, unsurprisingly, is that it remains in the gestural modality; but the way that gesture and language work together in a single modality is rather surprising. First, all language is both linguistic and gestural, and both vocal and signed languages have linguistic and gestural modes, although both modalities may not necessarily be available. For instance, gesture would appear to be absent when using a voice-only phone; but tone, loudness, emotional markers, accent, lexis, and even silence all serve as gestures supporting language. Similarly, gesture merges into sign language simply because signs are an open set of lexical items; so any gesture can be treated as a new sign and assigned a semantic role. Treating signs and gestures as qualitatively different is to miss the point of communicative language: to negotiate toward meaning. As Sandler puts it,

> The physical transmission system is not irrelevant or unimportant. It is not a mere secondary 'externalisation' of inherent structural organisation in the brain (contra Chomsky, 2007). Instead, defining linguistic and gestural modes in each language type now makes it abundantly clear how the physical transmission system contributes to the form of each mode of expression in each type of language. The resulting model of language is dynamic, flexible, and extraordinarily creative. (Sandler, 2022, 16).

Language is, and from an early stage was, multimodal. This does not mean that it cannot be disembodied – that would make reading, writing, and this book impossible. Before writing was invented, negotiation toward meaning happened when a signal was made, and the bodies of the speaker and listener were immediately involved in it. When writing appeared, this immediate negotiation disappeared; signals no longer attenuated as soon as they were complete, and things written could even outlive their authors. Yet writing remains a negotiation toward meaning: the writer must anticipate the needs of the reader and try to satisfy those needs within the text; metalinguistic content and illustrations act in the same way as a gestural mode. As Lemke (1998) shows, even writing is multimodal.

Cognition as a Source of Language Grammar

Language inevitably involves cognition: it is a product of cognition and, once in place, allows people to think in new ways. Language provides a system for sharing complex thoughts and ideas, and this distinguishes language from other signalling systems in nature (Kendon, 1991). Since the 1960s, the relationship between language and cognition has been explored within a research framework known as Cognitive Linguistics, which Dąbrowska (2016) describes as:

> … an approach to language study based on three central premises: that the function of language is to convey meaning, that linguistic description must rely on constructs that are psychologically real, and that grammar emerges from usage. (479)

Cognitive Linguistics views language as the outcome of a capacity to map thought to meaning, and a need to share that meaning: language provides schemata which enable us to frame thoughts into communicable forms, to express those communicable forms using a shared but arbitrary communication system, and to interpret those shared forms back into thought. The enterprise of Cognitive Linguistics has been described as an archipelago of ideas rather than a peninsula of knowledge (Geeraerts, 2006, 1); but it now becoming an interlinked network of cognitive functions (some primarily linguistic, some less so) working together to let senders and receivers of information negotiate toward shared meaning. Agreed meanings rely on our shared physical experience of the World, which in turn relies on the fact we have physical bodies with complex sensory input and sophisticated expressive output. Central to Cognitive Linguistics is the embodiment of language.

For instance, our body knowledge of *up* and *down* is a product of the external force of gravity, and this is reflected in our use of the terms: *up* is skyward even if you are standing on your head. In contrast, body knowledge about *left* and *right* is view-specific: if you make a half-turn then your external orientation changes, and what was on your left is now on your right, and vice versa. This is why the terms *left* and *right* are often accompanied by possessives (*my left, your right*), or by reference to relatively immobile external objects (*stage left, starboard side*), or by gesture to indicate actual direction. The negotiability of meaning allows body knowledge references to be used in other ways: as metaphors (*e.g., the political left and right*), or as cultural signifiers (*e.g., dextrous, adroit, sinister, gauche, cack-handed*). This seems to imply that, semantically at least, language is both the product of general cognition and a driver for it: some meanings are projected onto thought (and thus onto language) by external imperatives; some are generated by thought and projected onto language and external actuality; and some are generated by language use and thus constrain thought and cultural reality (Kövecses, 2002, ch3).

Johansson (2005) gives a comprehensive and largely uncontroversial analysis of current Cognitive Linguistics thinking about the sources of language grammar. He describes syntax as consisting of four levels of elaboration, which apply to individual development and seem to apply to language evolution, too (*ibid.*, 230-235): first comes structure, allowing simple two-word forms; then there is hierarchy, allowing phrase structure; then comes recursion, allowing potentially infinite forms to be generated; and with recursion comes flexibility, which allows the same thought to be expressed in different ways. Johansson tentatively suggests that, in evolutionary terms, recursion came first; but he sees the development of recursion and flexibility in children to be effectively simultaneous. He also suggests that the drivers for linguistic forms are twofold: social scripts, which allow the communication of interpersonal relationships; and image schemata, which allow the description of events.

This description of grammar both explains how grammar evolved and gives reasons why that evolutionary process began and continued. What it does not do is directly address the question of fitness: why is it a Good Thing to have a grammar in language and what advantage does it give to those possessing it? For Cognitive linguists, this question is part of the evolutionary conundrum of language as a whole: why should senders be willing to give away valuable information; and why should receivers be happy to accept information when fact, opinion and lies are difficult to differentiate? Compared to these communication dilemmas, the sources of grammar seem simple: once the dilemmas has been overcome then the drive to share increasingly complex mental constructs inevitably creates the need for more complex language structures to make sharing possible. The problem for the sources of language grammar is not how complexity arose, but why those particular complexities arose (Cronin, 2005).

However, Cognitive Linguistics is not the only theoretical stance available in linguistics. For Generative linguists, language is a product of a language-specific module (or system of modules) in the brain. This theoretical difference has considerable implications for the sources of language grammar: the Cognitive linguist looks for grammar in general cognition and attempts to describe grammar in terms of a need to communicate cognitive constructs; the Generative linguist is interested in how the key component defining language, the specialised language module, originates, and how it is instantiated in the human brain. Where Cognitive Linguistics gives a naturalistic view of language and cognition, open to a gradualist Darwinian explanation for the sources of language grammar

(Palmer, 2006), Generative Linguistics offers a modular view in which language can be isolated from general cognition, and the sources of language grammar can therefore be given a non-gradual, or macromutational, explanation (Chomsky, 2002, 84-91).

The Generative approach to grammar is described more fully in chapter 3, and the Cognitive Linguistics view of grammar is discussed in Chapter 5. In summary, though, Generative Linguistics sees language as primarily for thinking, with its communicative and social roles as incidental; the Cognitive view of language as a cooperative activity inevitably defines it as a social phenomenon. For Cognitivists, language exists because humans cooperate in stable social groups: how they cooperate drives what they need to communicate, and why they cooperate drives when and where they communicate and who they communicate with.

Social Construction as a Source of Language Grammar

For Generativists, the role of language in social construction is incidental: language is for thinking. For other linguistic approaches, however, the role of language as social lubricant (Dunbar, 1996) is uncontroversial. Much of our everyday language is involved in constructing social relations, and viewing language as a social lubricant seems apposite in many ways: we use language to work together in groups to achieve common ends; we use language to build social institutions; we use language to pass on ideas, allowing our ideas to live on after our death and creating cultural continuity; and we use language to negotiate, to include and exclude people, to entertain … all of which are socialising activities. It is difficult to imagine a simple human society without language, and impossible to imagine a complex civilisation without it. As Mufwene (2016) says,

> Cultures are not anterior to the particular ways in which languages evolved; rather, the latter contributed to shaping the former, assuming that the term culture is used in reference to the particular ways in which members of a population converge in their beliefs, in their social behaviors, and in the ways they do things. Cultures are not static; they evolve in ways that reflect changes in the beliefs, behaviors, and activities of the populations that produce them in the process. (156)

The relationship between language and social construction is not just a matter of one enabling the other; human societies and the grammar used in language are structurally similar enough to map one to the other. For instance, both share three important functions. First, they are both segmented (*e.g.*, Bosch *et al.*, 2013). Human societies consist of individuals who are nonetheless capable of working together to produce solutions where individuals cannot; and language consists of semantic units which can stand alone or work together to produce composite meanings. Second, they are both differentiated (*e.g.*, Lehmann, 2015, 27-128): individuals in human societies take complementary roles to solve problems; and, while some roles remain casually *ad hoc* (*e.g.*, "you push, I'll pull"), some become formalised specialisms which define individuals both to others and to themselves (*e.g.*, "I am a potter"). In language, words take different roles to resolve meaning; some have formalised roles in language structure (*e.g.*, nouns, verbs, *etc*.), but some are more *ad hoc* and take semantic roles based on the context of their use (*e.g.*, holistic utterances like *Yes* and *No*). Finally, societies are hierarchical, with some individuals deferring to others in ritualised ways which sometimes do not promote the deferring individual's reproductive fitness. Languages are also hierarchical, with some words governing the meanings and roles of other words (*e.g.*, Asano & Boeckx, 2015). This coincidence of structure between human societies and human language does not tell us whether language emerged from social structure or social structure emerged from language; but it does tell us that language and social structure map to the same cognitive mechanisms.

The social aspects of language have been explored in detail within linguistics. A branch of linguistics, pragmatics, is devoted to it (*e.g.*, Thomas, 1995); and a major grammatical theory, Systemic Functionalism, is based around the fact that language is exchange (Halliday, 1994, ch4). In terms of language origins, the role of language in the construction of human socialisation, and vice versa, have both been examined. For instance, Dessalles (2007) sees conversation and narration as key features in the development of both language and human societies. For Dunbar (2004), the exchange of social information was the original function of language. Searle (1999), building on the work of St. Clair (1985), identified institutional reality (believing things are real because we agree they are real) as the reason why human society is unique – and language is the way it is because of the need to negotiate institutional reality. For Worden (1998), language evolved out of primate social intelligence (the capacity to know yourself and others) via Theory of Mind (the capacity to model the mental states of others); and language therefore relies on social awareness. While Locke (1998) sees social sound-making as the source of language: a capacity to communally create phonological segments (attaching agreed meanings to the individual sounds in a signal) provides a route to syntactic segmentation (creating grammar) via semantic segmentation (creating systems of meanings).

More recently, Dor (2017a) sees language as a technology, the mechanism behind an increasingly sophisticated social communication system. As culture became more complex in the early *Homo* species, the need to share that cultural complexity required increasingly complex communication processes. At a certain point the communication system became so complex that it ceased to be protolanguage – or language-like – and became recognisable as language. This “invention” of language was a collective response to a need for communicative complexity, and came about because social complexity requires imagination, and communicative complexity requires ways of expressing and sharing imagination. The invention of language happened long enough ago that communicative fitness was able to evolve, allowing the more social, more communicative, more cooperative, more imaginative and more vocally adept to get more genes into the future. It is a good story, but it rather kicks the can of language down the road: something must have made an increasingly complex culture advantageous for early *Homo* yet not for other hominids; how did that happen? And why did it become a runaway process? Also, what does full language have that protolanguage lacks? If the difference between them is imaginative complexity, is it a quantitative or a qualitative difference? And if the second, what is the quality? Dor’s model shows how socialisation and language are linked by complexity, but it does not fully show how they work together as a species-making evolutionary system.

There are several ways that socialisation and language can work together, and they are not necessarily mutually exclusive. In a study of over 2,000 languages, Lupyan & Dale (2010) found that the demography of a language is correlated with the morphological complexity of that language: the larger the group of speakers, the simpler the morphology. They suggest that each new generation introduces simplifications to the morphology of a language, and a language with a greater number of speakers accumulates simplifications quicker, while languages with small groups of speakers retain idiosyncratic morphology to increase redundancy, which makes first language learning easier. They call this the Language Niche Hypothesis: languages vary because the variations improve negotiation toward meaning in different environments. However, Atkinson *et al*. (2018) found no evidence for group size effects on language complexity in extended use; but their group sizes (two people is a small group, three is large) were not comparable with those of Lupyan & Dale. It can be argued that the two experiments, which used very different methodologies, are not studying the same thing; the results obtained from one study do not necessarily prove or disprove the other; both may be right.

Looking at language as a marker of membership, Kinzler *et al*. (2007) found a bias in young children toward people who spoke the same language. Infants preferentially watched, and accepted more toys from, a person who spoke their language; and pre-school children preferred to make friends with people who spoke their language. Kinzler *et al*. offer three reasons for this preference: first, a shared language likely indicates a shared culture, making negotiation toward meaning simpler; second, a shared language is a useful shortcut for identifying potential allies; and third, a shared language saves time – the child does not have to build a new socio-cultural and linguistic environment to negotiate toward meaning. The one thing this study shows clearly is that language is treated from an early age as a social marker of in-group and out-group membership.

The social construction of language has considerable implications for grammar. If language is a conduit for social information then it must be capable of expressing the social complexities of the group, to allow members to discuss the group. It should be capable of signalling relationships between others, and between objects and others, as events; and it should also be able to locate relationships both in space and time, because the events being communicated need not be current. This important source of language grammar, communicating social constructs, is explored in more detail in chapters 6 and 7.

The Magnificent Seven

Language seems to be central to several effects that define us as human. As play, language allows us to communicate ideas in an environment where the pressure of reproductive fitness has somehow been switched off, or considerably dampened. As a tool-like mechanism it enables us to do things we could not otherwise do: it lets us formulate a problem as a segmented series of subproblems, and also lets us recruit others to assist in solving problems that are beyond any one individual. As a signal of fitness, it enables us to demonstrate our individual fitness, and show that we can freely give away our knowledge and still be fitter than our conspecific rivals. As an embodied gestural system, language behaves like other gestures in that it is both a way to do things and a way to communicate things. As a multimodal system, it can be produced vocally and gesturally, either separately or simultaneously; and it can be perceived aurally and visually, with writing letting us store information outside of human heads. As a cognitive process, language lets us generate complex models of the relationships around us, and then manipulate those models to test different outcomes. And as a communicative act, language lets us cooperate with others to solve problems and undertake large projects. Language does all these things; and a theory of language origins that does not address all these issues is likely to miss something important.

3 Generativism and Sources of Language

Grammar

> ***Rational discussion is useful only when there is a significant base of shared assumptions.***
> *(Noam Chomsky)*

Chomsky has been so important to modern linguistics that its history can be divided into two eras, pre-Chomskyan and Chomskyan. Before *Syntactic Structures* was published (1957), linguistics was largely an anthropological enterprise: linguists lived with the locals of other countries, learning their cultures and languages simultaneously (Boas, 1938). Language was described in terms of social theory, so a gap in language functionality indicated a corresponding gap in culture (Sapir, 1921), and even in cognition (Whorf, 1956). The approach was largely behavioural – language was seen as culturally learned, with children being trained into language – and the possibility of language being an innate human capacity based on genes was largely unconsidered. In the 1920s and 1930s, Bloomfield adopted a mathematical approach to linguistics (Tomalin, 2004); and the Prague Linguistic Circle (notably Jakobson, publ. 1987) developed a structural approach based on Saussure's earlier work (publ. 1972); but the approach to language remained conspicuously anthropological.

When Chomsky's Generative Grammar programme appeared, it comprehensively altered the way linguistics was done. Generative Grammar views language as more than just learned, it has an innate, genetic nature with its own rules and forms. Human languages are similar because they all do similar jobs, but behind that similarity of role is a similarity of form: human languages all do the same jobs because they are all generated by the same genetic language system, and those are the only jobs that languages can do. After 1957, linguistics and anthropology largely parted company: culture no longer dictated language, it could only impose a light dusting of difference onto a solid core of genetic sameness. This chapter gives a short overview of the development of Generative Grammar since 1957, hopefully providing an insight into a half-century that redefined linguistics.

Uncovering the Structure of Language Grammar

In *Syntactic Structures* (1957), Chomsky showed the grammar models then available were inflexible, incomplete, and incapable of analysing complex grammatical utterances. He proposed a new methodology, Transformational Grammar, to encompass the grammars of all possible human languages, past, present and future – effectively, a Universal Grammar. The transformational rules in this grammar would not just be descriptive, they would be generative, able to explain any utterance in any language: from a limited set of utterances, Transformational Grammar would comprehensively predict the grammar of any language. Being computational, the generative template for languages would work equally well on a computer or in a human brain. However, while Chomsky gave examples of how the transformational template could work, he did not provide the complete Transformational Grammar. *Syntactic Structures* was a promissory note for a solution still to come.

In *Aspects of the Theory of Syntax* (1965), Chomsky elaborated what has become known as the Standard Theory, proposing two structural levels in language, deep and surface. Deep structure consists of the underlying "rules" which define the nature of language as a phenomenon, and which rely on innate, genetically-controlled language capacities common to all humans – the Universal Grammar. The surface structure consists of rules specific to individual languages which are not part of the Universal Grammar, and therefore not innate. Surface structure grammar rules must be learned rather than activated.

Standard Theory proposed transformational rules interpreting between deep structure forms and surface structure utterances, with a simple coding process converting surface structure forms into phonological signals. The signal itself is just the tip of the generative language iceberg. Chomsky identified three resources in deep structure: the lexicon (words themselves), phrase structure rules (how words work together), and semantic values (what the words mean). These resources are not language-specific in Standard Theory, they are part of the innate universal resources of humans.

Although Chomsky initially intended Standard Theory to fully describe Universal Grammar, he left the mechanisms for others to discover. Without a governing theoretical description, however, different people found different solutions. Fillmore (1971) proposed *Case Grammar,* where deep structure was the relationship between the verb and other sentence components; *Relational Grammar* (Postal, 1968, among others) was almost the mirror-image, concentrating on the hierarchical relations of nouns as subjects, objects and indirect objects; and Lakoff's *Generative Semantics* (1971) concentrated on semantic content to discover deep structure.

Soon after the publication of *Aspects*, it became clear that there were problems with Standard Theory (Chomsky, 1977 [1998], 151-152); in particular, semantic interpretation seemed more variable than a deep structure resource should be. The role of semantic interpretation had to change, therefore, to moderate surface structure as well as deep structure. However, this left surface structure without a simple one-to-one relationship with the phonological interpretation; so new mechanisms were needed to convert between surface structure and both phonological and semantic interpretations.

Yet this amended model remained problematical: lexicon was divorced from semantic and phonological interpretations; but then, what is a mental lexicon if not an encyclopaedia of meanings and sounds? Standard Theory provided no satisfactory answer: the only feature that could occupy the role of lexicon was word class (noun, verb, *etc*.), but Chomsky placed this firmly in phrase structure. The importance of the lexicon in moderating deep structure was downplayed, as was the semantic interpretation; but this left deep structure even further impoverished.

Elaborating the Structure of Language Grammar

To solve this, Jackendoff (1972) revised the Standard Theory into Extended Standard Theory (EST): he inserted semantic interpretation parallel to the transformations between deep and surface structures and, like the transformations, gave access to and from both structures. Semantic interpretation thus became a resource for other processes rather than a process itself. The EST model remained robust for over a decade, during which time the theory was elaborated in several ways. For a while, it seemed that Generative Linguistics finally had a stable base.

One aspect of EST was X-bar theory, the idea that language phrases consist of an X-value (the phrase-defining word) and an optional specifier. So, the phrase *simply happy man* consists of a noun phrase (X-value *man*, specifier *simply happy*), and an adjectival phrase (*simply happy*), which consists of an X-value (*happy*) and a specifier (*simply*). This hierarchical binary relationship occurs throughout an utterance, creating a tree structure with meanings combining at each level to eventually create sentential meaning; and every language uses this tree structure, making it a language universal. X-bar theory contributed significantly to our understanding of language structure, and it is extensively used in linguistic analysis today.

Chomsky, however, wanted EST to place greater emphasis on language universals, and during the middle 1970s he worked on what became known as the Revised Extended Standard Theory (REST) (Chomsky, 1975 [1998]). This theory emphasised Movement in grammatical construction, an effect usually illustrated by the English interrogative form. In a question like *Who did you see?*, we take the statement form *you saw X*, move the object to the front and insert a neutral verb, such as *do* or *be*. The movement seen in the English interrogative is also used in other grammatical forms: for instance, it can be seen in the English passive (*you saw the man → the man was seen by you*) and the English noun phrase (*you saw the man → the man you saw [was happy]*). The emphasis on Movement meant that REST moved the semantic interpretation from between surface structure and deep structure to between surface structure and semantic representation, leaving the transformations to manage Movement.

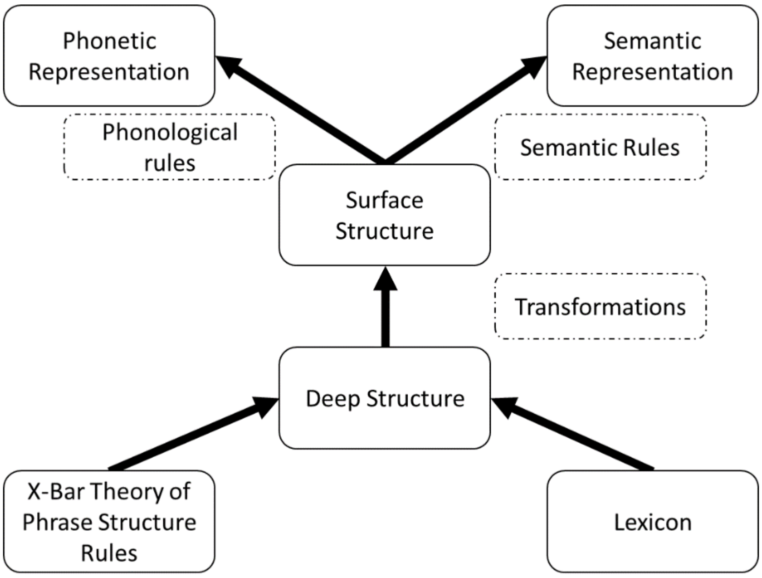

Figure 3.1: Revised Extended Standard Theory

Movement is important in Generative Linguistics because it shows how the deep form of an action between two objects, an *actor* (or doer) and a *patient* (or done-to), is independent of the surface utterance; and Movement transformation is needed to let the deep structure thought become the surface structure utterance (Radford, 2009). There are, however, issues with Movement, including whether it is needed at all. If the deep thought has a natural order, why subvert this to an unnatural order in surface utterance? For instance, the English interrogative can be, and often is, not subject to Movement, with *you saw X* becoming *you saw whom?* or *you saw what?* Generative analysis treats this as a marked form, asking for confirmation of something already identified by the original speaker. Movement also occurs only when the object clause in the sentence is being questioned; the subject interrogative form is not subject to Movement (*X saw you* → *who saw you?*), and the indirect object interrogative form treats the moved and unmoved forms as exchangeable (*you saw her yesterday* → *when did you see her?* or *you saw her when?*), implying that only a subset of interrogatives use Movement. So, what function does the moved interrogative serve that is not met by the unmoved interrogative?

There are also problems where constructs which appear to have been move-transformed do not seem to work in their unmoved state (*Aren't I* *right?* ← * *I aren't right*[4]), a problem known as the Movement Paradox; and, while adverbial Movement is used mostly for emphasis, in some cases it changes meaning (*she built it quietly / she quietly built it* both have the same meaning, but *she built it simply / she simply built it* do not). We can either assume that Movement works differently in deep structure and surface structure, or we can assume it is a feature of surface structure only; or, as Sag *et al*. (2020) do, we can propose a Movement-free formalistic explanation. This, however, seems to be cognitively quite burdensome compared to an *ad hoc* case-by-case treatment – especially as the case-by-case treatment predicts what we actually see: a wide variation (dialectal, contextual and idiosyncratic) in what is considered grammatically acceptable.

One grammatical concept required by the Move function is the linguistic *trace*, an unexpressed element of structure. This phonologically zero element marks the pre-Move position of a Moved element, so it represents deep structure within surface structure. Chomsky saw traces as fully expressed in deep structure, with the phonological rules zeroing them and then passing the silenced traces to the phonetic representation. So, for example, the surface structure of *who did you see **t**?* (where ***t*** represents the position of the *X* in *you saw X*) is rendered in the phonological rules to the surface form *who did you see?* The interrogative phonetic form comes from a transitive deep form (*you saw X*), so the trace is a marker of transitivity for the verb *see*. The trace, like Move, is important to Generative Linguistics because it emphasises underlying cognitive forms in the construction of utterances.

Traces, however, bring their own headaches, one of which is their role in holistic utterances. Take, for instance, the utterance *Yes*: it is used as a complete construct, but it also has a recognisable X-bar form: it involves an action (*e.g.*, agreement or acquiescence) between two people, one as actor (or agreer) and the other as patient (or agreed-with). It is both transactional and contractual, establishing a relationship about something between the actor and patient. Semantically we can describe it as [I] [agree/acquiesce-with] [You] [about-X], which Austin (1962) would describe as a

[4] In linguistics, a preceding asterisk is used to indicate an ungrammatical construction, a preceding question mark is used to indicate disputed or uncertain grammaticality.

performative. However, where in the word *Yes* are the actor, action, patient and target? They would all seem to be traces. This is not a trivial problem, holistic utterances are common in everyday language use: *yes*, *no*, *perhaps*, *maybe*, *OK*, *drat*, *why*, *but*…; the list is long. Either we must exclude them from linguistic analysis or accept them as grammatical anomalies.

Yes			
Actor	Action	Patient	Target
[I]	[agree-with]	[you]	[about-X]
S			
NP	VP		
	V	NP	
		NP	PP
[trace]	*[trace]*	*[trace]*	*[trace]*

Figure 3.2: The Holistic Utterance *Yes* and its Traces.
NP=noun phrase; PP=adpositional phrase; S=sentence; V=verb; VP=verb phrase

Exceeding the Theoretical Baggage Allowance

By 1982, the zero element was a major preoccupation for Generative Linguistics, and the theoretical structure of Principles and Parameters (P&P) was being assembled to cope. In P&P, certain features of deep structure are cognitively present in the final utterance even if they are physically absent. They pass through the transformations of the REST without excessive change because they are “pre-parameterised” by language-learning. The Universal Grammar available at birth permits several different language structures; but the first language learned switches off some of those options, which reduces the range of permissible language forms and creates a standard utterance template. If any element of the standard template is not present in a final utterance, it can be interpolated as a trace. At this stage, Chomsky saw Universal Grammar as a potentiator in newborn humans, generating one of a large but finite set of real languages (Chomsky, 1986, 38-39).

P&P introduced new terminology to the Revised Extended Standard Theory. The initial state of language, identifiable with Universal Grammar, became S_0; and the attainable state, identifiable with an individual's grammar, became S_L. There was also I-language, language internal to an individual's mind, and E-language, social or external language. This reflected Humboldt's (1836 [1999], 74-76) earlier distinction between *Language* as a capacity and *Languages* as communication.

Vital to P&P were the principles of Government and Binding. Government extended X-bar theory, setting out how X-values governed their specifiers (Cook & Newson, 1996, 51); basically, individual languages show preferences for specifier-first or X-value-first constructs. So, if a language puts the governing verb before the governed object, it is also likely to put the governing noun before the governed adjective. However, examples of non-compliance with this rule, such as English (verb before object but noun after adjective), weakened its universality. It is now mainly used to illustrate the relationships between verb inflection and subject, between verb and object and between adposition and indirect object. For instance, in the sentence *Alf saw a house with chimneys*, *saw* governs *Alf* as a past tense form, *see* governs *a house with chimneys* as a verb-object form, and *with* governs *chimneys* as an adpositional form.

Binding is a semantic process allowing lexical items to reference previously defined items, sometimes at some distance. Pronouns are an obvious example, but adjectives like *same* (as in *it's the same idea*) or adverbs like *again* (*it's happened again*) can also have binding properties (Chomsky, 1988, 52).

Contrary to the original intention of simplicity, Universal Grammar was becoming complicated. Chomsky himself included X-bar theory, Government theory, Binding theory, θ-theory (Theta theory), Case theory, Bounding theory and Control theory into P&P (Chomsky, 1982, 6). θ-theory states that all components required to formally define a sentence are expressed in deep structure even if their surface form is zeroed. Case theory is concerned with case assignment (*e.g.*, for agreement or tense) as an abstract feature of deep structure. Bounding theory addresses the conditions of separation and linkage operating on items subject to binding. Finally, control theory examines phonologically zero elements, both the old *t* and the new PRO, used in languages like Spanish to explain dropped subject pronouns, (hence PRO-drop languages). Yet, despite decades of investigation by a legion of linguists, this extensive theoretical framework still did not offer a description of Universal Grammar.

Back to Basics

Chomsky realised that Universal Grammar had to return to first principles, so he produced *The Minimalist Program* (1995). In this book, Chomsky reiterates that the P&P approach of REST provides vital tools for linguistic analysis; but a single, rather enigmatic paragraph in chapter 4 renounced years of Standard Theory research:

> A linguistic expression of L is at least a pair (π, λ) meeting this condition [capable of Full Interpretation] – and under minimalist assumptions, at most such a pair, meaning that there are no levels of linguistic structure apart from the two interface levels PF and LF [Phonetic Form and Logical Form]; specifically, no levels of D-Structure or S-Structure. (Chomsky, 1995, 219.)

Chomsky (2000, 10) restated this radical position, sweeping away differentiation between deep and surface structures. Zero elements also disappeared, if they were ever there (Chomsky, 2005). This radical theoretical revision has not endeared him to many older grammaticians who spent much of their lives identifying deep and surface components and mapping the transformations between them. Each new version of Generative Grammar left behind theorists who continued working on an older model, although many later upgraded. However, the Minimalist Program carried forward fewer adherents than previous Generative Grammar versions. As Newmeyer said:

> If I were to write this book several years from now, I would opt for the MP [Minimalist Program]. However, at the present time, I find the concrete claims of the MP so vague and the total set of mechanisms that it requires (where I have been able to understand them) so *unminimalist* that I see no

> reason to encumber the exposition with my interpretation of how the phenomenon in question might be dealt with within that approach. It is also worth pointing out that even leading developers of the MP typically appeal to strictly GB [Government and Binding] principles in presentations to general audiences of linguists. (Newmeyer, 2000, 12-13.)

The Minimalist Program has inspired a new generation of linguists, however. Hornstein (2001) proposed that the seven theories of REST (Binding, Bounding, Case, Control, Government, X-bar, and θ theories) are all types of Movement, simplifying Chomsky's (1995, 297-312) theory of attraction, or A*ttract/Move*, which itself replaced the *Merge and Move* doctrine. Hornstein's approach was simpler than Chomsky's, and therefore more minimalist; but it has since been superseded by a greater concentration on recursion (referred to as *Merge* in the Minimalist approach). In the tradition of Humboldt (1836 [1999]), linguistic recursion "takes a finite set of elements and yields a potentially infinite array of discrete expressions" (Hauser *et al*., 2002, 1,571).

Recursion in language is often illustrated using the following story:

> 'Twas a dark and fearsome night. Brigands great and brigands small were gathered around the campfire. 'Come, Antonio,' they called to the terrible chief, 'tell us one of your famous stories.' And Antonio arose and said:
>
> "'Twas a dark and fearsome night. Brigands great and brigands small were gathered around the campfire. 'Come, Antonio,' they called to the terrible chief, 'Tell us one of your famous stories.' And Antonio arose and said:
>
> "'Twas a dark and fearsome night. Brigands great and brigands small …".[5]

The most exciting promise of *Syntactic Structures* for most linguists was a mechanism to translate any utterance from one language to another – a Universal Grammar. However, the Minimalist program views Universal Grammar as probably absent from adult minds. It is active in child minds, guiding their first language acquisition; but, when the grammatical switches

[5] First attested use was in the Buffalo Times of New York, March 1900 (https://quoteinvestigator.com/2019/11/10/antonio/).

of their first language have been set, their Universal Grammar becomes inaccessible or ceases existence. Just as a dress pattern can be cut to a particular size, so Universal Grammar can adapt to a particular language; but, once cut or adapted, neither can be restored to their "universal size". The universal is destroyed by the production of the necessary specific.

In the Minimalist Program, Chomsky insists that language is recursive, relying on *discrete infinity*:

> Human language is based on an elementary property that also seems to be biologically isolated: the property of discrete infinity, which is exhibited in its purest form by the natural numbers 1, 2, 3,... (Chomsky, 2000, 3.)

Infinity, however, is slippery, comprehensible only by reference to examples. For Chomsky, it refers to the mathematical concept of open sets: for instance, no matter how large a number, you can always add 1 to make another number; and no matter how long a sentence, you can always add a phrase to make a longer sentence. However, while it is trivially true that language can produce infinite sentences, there is no conceivable way that this could be genetically advantageous. It is also true that the range of possible sentences in a language is larger than the number of sentences that will ever be uttered, but this is also trivial. Language is dialogic, it is not about what can be done but what should be done – and what is done – to communicate; and infinity is neither needed nor produced in actual language.

Chomsky sees discrete infinity as key to our capacity for recursion, allowing established language forms to be used at new levels of construction: the noun phrase stands in place of a noun and contains a noun; but the noun in the noun phrase can itself be a noun phrase and contain its own noun or noun phrase. This hierarchical reuse seems to requires sophisticated mental modelling available only to humans, emphasising the importance of recursion in language (Hauser *et al.*, 2002). However, the assumptions behind this approach have been comprehensively challenged (*e.g.*, Pinker & Jackendoff, 2005; Evans & Levinson, 2009; Pullum & Scholz, 2010; Progovac, 2019); and, as discussed in chapter 9, the role of recursion as the key to language grammar is far from settled.

Universal Language Grammar?

By treating grammar and language parameters as genetically innate, Generativists can claim that language is universal. Yet they remain unable to describe the evolutionary process that created a universal grammar system, and to date they have been unable to definitively describe the structure of Universal Grammar. This may well be due to insufficient time studying the system: linguistic science may be too young to provide all the answers (Baker, 2001, ch7). However, the generative view of language is minimalist: the universal parameters of language are few and powerful. Universal Grammar must therefore have a finite, and probably highly compact, structure. It is not a moving target, and it does not involve elusive components. True, it must be studied through instances of actual language, which can only show Universal Grammar "through a glass, darkly"; but the number of actual language utterances is enormous, much greater than any single linguist could analyse in a lifetime – if universal grammar remains mysterious, it is not for want of data. Continuing difficulties in discovering the universals of grammar and how they work must raise questions about their nature and even about their existence.

At the heart of Generative Linguistics is the view that language is optimised for cognition, not communication, and cognitively it is a near-perfect system (Chomsky, 2002, 105-109). Kinsella has questioned this approach, particularly in relation to the Minimalist Program (2006 as Parker, 2009). She has shown that some of the constraints on which the principles of Universal Grammar have so far been built are less simple and less universal than is necessary in a near-perfect system. Human brains are individual and idiosyncratic; so any universal constraints on grammar would have to be compromises, and therefore cannot be near-perfect; while if they are near-perfect they must be complex and contingent, so probably not truly universal. There are contradictions at the heart of the Minimalist Program which need to be addressed if it is to be compatible with a Darwinian approach to language evolution.

A major problem for Generative Linguistics is that language changes: somehow the rules that transform thought into utterance in one generation or group are subverted and redefined by other groups and following generations. This process is known in linguistics as grammaticalization (Hopper & Traugott, 1993; Heine & Kuteva, 2012). According to Generativist theory, changes to one part of the system should create cascade effects in other parts of the system, so that the parameterisation of the new form of the language is consistent with the dictates of Universal Grammar. Yet the historical evidence seems to show no such heavy parameterisation; instead, we see in many languages a richness of valid forms, as each grammaticalization adds to, rather than replaces, acceptable grammatical forms.

For instance, *at this point we diverged* is similar in meaning to *we diverged at this point*, and *this is the point at which we diverged*, and *our divergence occurred at this point*, and even *we diverged here*. All these forms express the same idea and have a semantic relationship which cannot be explained in grammatical terms. While other linguistic theories seek to explain this semantic relationship, Principles & Parameters and Minimalist theory are largely silent. Indeed, the traditional Generativist interpretation is that the grammatical differences mean that they are different constructs, and therefore cannot be co-analysed. Yet it seems reasonable to expect a theory of linguistics to address this: language is largely about meanings, whether being manipulated in a single mind or exchanged between minds, and it would appear reasonable for semantics to be central to any linguistic theory. Chomsky disagrees with this view (2002, 110-111), but he recognises that he is ploughing a lonely furrow.

Another question that Generativism sidesteps is whether full grammar is always needed for language utterances. Words themselves encapsulate meaning, and they can work without a grammatical overlay. Indeed, as we have seen, some of our commonest utterances, like *yes* and *no*, live in a strange grammatical limbo; and an utterance like *You! Here! Now!* demonstrates that messages can be clear without grammatical complexity. Idioms must also be treated with care, so constructs like *that's something I wot not of* (I have no knowledge of that) cannot be analysed in a standard Generativist way. Local linguistic variation also creates problems. For instance, *I'm gonna go see he's there* (I am going to go and see if he is there) can be viewed either as aberrant standard English grammar or as correct dialectal English grammar – or even as someone using the minimal amount of grammar needed to convey meaning. The speaker and the listener are both part of the semantic and grammatical context of the utterance. It could be argued that these utterances all indicate that surface structure relies on deep structure, but that sounds like the argument for phlogiston in early physics: things that burn must contain something to make them burn. Deep structure could be the phlogiston of linguistics, a position not incompatible with Chomsky's Minimalist Program.

However, we must also recognise that there do seem to be some true universals in human languages: nouns and verbs are one (however we name them), and the Subject-Verb-Object construction is another (although not necessarily in that order). These generalised language features point toward some form of universal basic structure behind language. There is also evidence from child studies that some bootstrapping mechanism or natural sensitivity to language is already present in a child's mind from birth. Generative Grammar does therefore have something to offer the search for language origins, and this book does not ignore Generativism as a description of grammar.

Generative Linguistics on the Sources of Language Grammar

Language origins poses a major problem for Generative Grammar theories. Chomsky has largely side-stepped the issue, maintaining that language cannot have emerged from pre-language communication: language is the product of a single, specialised and biologically integrated Universal Grammar whose sole function is making language; it has no pre-language precursor, and without this faculty there cannot be language. With it, however, language must map to the universal principles dictated by the faculty (Chomsky, 2007). Universal Grammar is seen as the product of an unexpected and unpredictable mutational event, and this has led to the widespread belief that Chomsky supports a sudden and catastrophic appearance of language in humans, often called a macromutation (*e.g.*, Evans, 2014). However, while "sudden and catastrophic" best typifies the

Generativist approach to language genesis, Chomsky himself displays only a cursory curiosity about the subject. By assuming sudden appearance he can treat Universal Grammar as an ideal system, ideal because it is the only game in town. There is no reason to look for variation in this universal system because it was the result of a relatively recent, single mutation in a single human, which propagated swiftly through the whole *Homo sapiens* species; by definition, there is no significant variation to be found (Chomsky, 2006, 106-113).

The Generativists who do take an interest in language genesis must deal with this Chomskyan assumption of sudden and catastrophic evolution. This is problematic because, in evolutionary terms, catastrophe is usually just that: highly disadvantageous to the phenotype. This creates a paradox for Generative evolutionists: Universal Grammar is posited as a unique faculty with no precursors, which must therefore have appeared suddenly as a single macromutation; it did not evolve incrementally from pre-existing faculties because it has no evolutionary precedent. This makes it, unlike any other human faculty, impossible to explain using evolutionary theory – because evolutionary theory is based on the Darwinian dictum of descent with modification (Darwin, 1859, 132); and, as Darwin himself said, *Natura non facit saltus*, Nature makes no jumps (Darwin, 1859, 194). Darwin's theory has been redefined several times since it was originally published (*e.g.*, see Klein, 2009; Dennett, 2009; de Waal, 2009; Szathmáry, 2006; Wang *et al*., 2006); but rewriting the very basis of the theory to accommodate the peculiarities of human communication and socialisation seems somewhat of an overkill. Rewriting may help us with some apparent peculiarities of human cognition; but, as our understanding of nonhuman cognition remains rudimentary, we cannot yet know whether human cognition really is exceptional. Nonetheless, attempts have been made to differentiate human and nonhuman cognition, either because nonhumans cannot handle "higher-order, systematic, relational capabilities of a physical symbol system" (Penn *et al*., 2008), or because language and awareness of self are attributes that only humans have (Malik, 1998). It is all rather circular: humans are different because we have X and other species do not; and we know that having X is important because it makes us different.

The Generative evolutionary paradox can be summarised as follows. First, Universal Grammar is innate: it started with a genetic mutation in a single individual and then propagated very quickly through the population, becoming not just dominant but universal in very few generations. This, in turn, means it must have satisfied an urgent pre-existing fitness need, solving a serious fitness deficiency in the population: something made the individuals with the mutation spectacularly more fit than those without. However, a fitness solution can only work on a pre-existing fitness need, a mutation cannot simultaneously generate both a fitness solution and the need for it; and Universal Grammar is discontinuous from the rest of nature, so there cannot have been a pre-existing fitness need. This makes Universal Grammar a solution seeking a problem, a *hopeful monster* mutation seeking justification; and, as Pinker & Bloom (1990) show, successful hopeful monsters are rare.

Three attempts to solve the Generative evolutionary paradox are outlined below. They are far from the only solutions that have been proposed, but they do give a flavour of the problem that faces Generativists: how, in a Darwinian universe, can a complex, innate faculty like language come into being *de novo* as a unique but integrated system?

The first approach to the Generative evolutionary paradox is that of the core Generativists around Chomsky. It consists of three steps: accept that most of the discontinuities between human language and nonhuman communication are actually continuities and argue for only one discontinuity, recursion or Merge. Next, downplay the genetic difference that Merge entails: it fundamentally reconfigures cognition, communication and socialisation, and it improves fitness so completely that all individuals without it disappear within a few generations; but somehow it requires only a tiny genetic change. Finally, appeal to mystery: with no precursors and no way of dating the catastrophic mutation, we have no way of knowing how or why language evolved (Berwick & Chomsky, 2016).

The second approach is to avoid the catastrophic change, and instead formulate language origins as an extended evolutionary transition from pre-linguistic communication to language throughout hominin development.

Each small increment in language capacity is selected for because language in any form is an advantageous thing to have (Pinker, 1994, 333). The only problem with this approach is that the increments should be described and justified: how does each one enhance fitness? As a hypothesis this second approach works well; but converting it to a theory requires the details to be spelled out – because, as Charles Eames said about the chairs he made, “The details are not the details. They make the design”.

The third approach to the Generative evolutionary paradox takes a middle road. Instead of one giant step or a large number of small steps, it proposes a small number of large steps between pre-linguistic communication and full language. Bickerton (1998) sees only one intermediate signalling stage, which he refers to as protolanguage. He provides evidence for this single stage from child language and aphasic language, and from the development of unstructured pidgin languages, which occur spontaneously where two different linguistic communities interact (Bickerton, 1990). Pinker, however, considers the gap between protolanguage and full language in Bickerton’s model to be too great for a reasonable evolutionary explanation, and refers to Bickerton’s approach as “reminiscent of hurricanes assembling jetliners” (Pinker, 1994, 366). Between the two positions, Mufwene (2008) envisages more than one route from pre-language communication to complex languages. The conceptual gaps between the routes were closed by a constant stream of interpersonal negotiation: just as pidgins and creoles are generated by interpersonal negotiation toward both meaning and structure, so increments of complexity were negotiated between individuals as human communication became more language-like. It was a long, intergenerational process driven by idiosyncratic differences in individual models of communication, with genetic variation embedding shared complex innovations into the species rather than generating them.

If Generativism is right about the existence of Universal Grammar then Mufwene’s position on its origins is more tenable than those of Pinker or Bickerton – even though Mufwene does not identify as a Generativist. For Generativists, Universal Grammar is an “organ” (Chomsky, 2000, 4), present in humans and absent from all other animals. It is species-invariant and does not assist language – it generates it. It is also indivisible: the rules do not define specific language forms, they determine the structure of language completely (Bickerton, 2000). For Chomsky and Bickerton, there is no halfway with Universal Grammar: if it is present, full language can be generated; if it is absent, full language is impossible. As Bickerton says, "a common code is an all or nothing thing – you either have one or you don't" (Calvin & Bickerton, 2000, 96).

For Bickerton, the change from protolanguage to full language is catastrophic (Bickerton, 1998), because Universal Grammar is a single, indivisible system. Bickerton acknowledges no social factors in the genesis of protolanguage; instead, he sees it as a product of interactions between individuals concerning the environment: protolanguage contains information about foraging and other survival matters only. The truth of this environmental knowledge is immediately testable, so deception is not an issue (Bickerton, 2002). However, the model assumes a pre-existing level of socialisation and cooperation which makes information sharing worthwhile for both sender and receiver. Bickerton places his speakers in a language-friendly environment to overcome the problem of language genesis, but he does not explain why a language-friendly environment came about without language.

While Bickerton's protolanguage remains problematic, it should not be dismissed. It is indeed likely that full language was preceded by a functional communication system which had aspects of segmentation (a signal contains subunits of meaning which can be reliably identified), differentiation (different segments mean different things) and hierarchy (segments can be combined into signals, and signals can be combined into discourses); and it is the nature of this system that Wray (2000, 2002a, 2002b) examines.

For Wray, the problem of protolanguage is continuity: it must form an intermediate state between holistic primate communication, where the whole signal represents a single idea, and analytical human language; yet protolanguage in Bickerton's model seems to have continuity with neither. Wray solves this problem by showing that modern human languages, like primate signals, contain holistic utterances. This is evident in little words like *yes*, *no* and *thanks*, and even in apparently segmented utterances. Listeners can analyse idioms like *how are you?* as segmented or holistic constructs, although their reactions are likely to be quite different. Wray describes holistic utterances as “performance without competence” and shows they are more common than we think. The segmentation of utterances into words is sometimes illusory.

The sources of language grammar involve not only building separate meaning-units into utterances, they can also arbitrarily divide holistic utterances into separate meaning-units. Wray proposes two possible routes from holistic utterance to segmented language. The first is a slow increase in the number, range and use of analytic constructs throughout the history of protolanguage; the second is a slow evolution in cognition of the features which allowed analytic language to emerge, but with the actual emergence into speech being a single event. As sharing analytical language imposes limitations on the analytical structures possible, Wray favours the second solution, letting analytical language evolve in cognition unrestrained by the limits imposed by communication itself.

There remain some issues with Wray’s analysis, such as her reasons why grammar appeared at all. She also does not address why language is so over-engineered for communication, or why it can so easily produce lies. But her theory, like that of Mufwene, does offer a middle path between Chomsky’s and Bickerton’s unlikely catastrophic events and Pinker’s optimistic incremental evolution. And significantly, like Mufwene, she does not identify as a Generativist.

4 Structuralism and Sources of Language Grammar

> ***Everyone, left to their own devices, forms an idea about what goes on in language which is very far from the truth.***
> *(Ferdinand de Saussure)*

One of Chomsky's aims with Generativism was to give linguistics the same scientific credibility as other natural sciences (Chomsky, 2002, 56-60). There is a prejudice common among scientists that their discipline involves doing actual science, while other disciplines work with data of questionable validity that do not produce credible, replicable or applicable knowledge. The divide between the natural, or "hard", sciences and the social sciences is particularly marked, and even today some physicists, chemists and biologists refuse to accept economics, anthropology, sociology and psychology as real sciences. If science is about external physicalities then this is a valid viewpoint; if it is about the practical value of your research and your approach then it is not.

Chomsky believed linguistics should have the credibility of a full science to progress beyond mere description. The Generative approach led to new types of meta-analysis: linguistics was no longer just about language differences, it was about language as a human capacity – the language of languages. However, by emphasising language as computation, Generativism inevitably produces a computational theory of forms and rules.

This chapter looks at alternative contributions to the language grammar debate from Systemic Functional Linguistics, other Functionalist theories, and linear approaches to grammar.

A Short Introduction to Systemic Functional Linguistics

Language exists to carry out a range of functions: it enhances individual cognitive processing; it provides a reliable information channel between people, enabling negotiation toward meaning; it expands an individual's cognitive capacity by co-opting other brains; and, with writing, it co-opts external resources for offline storage and communication. These functions are not discrete, they are interrelated parts of a system which makes language a fit strategy in Darwinian terms. *Systemic Functional Linguistics* (SFL) aims to identify the full range of language functions and the way they work together as a system.

SFL linguistics developed out of Halliday's research in the 1960s (*e.g.*, 1969 [2003]). However, while the philosophy of Generative Linguistics has changed radically several times, abandoning old theories and the developing new ones, the SFL doctrine has developed incrementally. The seminal text for SFL remains *An Introduction to Functional Grammar*, first published in 1985 (Halliday) and now in its fourth edition (Halliday & Matthiessen, 2014).

The SFL grammatical ideology can be summarised as, "semantic (concerned with meaning) and functional (concerned with how the language is used)" (Bloor & Bloor, 1995, 2). This description somewhat resembles the doctrinal position of Lakoff's (1971) *Generative Semantics* (GS); but GS emphasises the role of universal and deep semantic structures, which rely on innate meanings. GS is concerned with "nonsyntactic semantic regularities" (Ziff, 1960, 42) which generate meaning independently of the social and grammatical contexts in which they are produced.

Unlike GS, the SFL approach to meaning is inspired by the semantic theories of Sapir and Whorf. Sapir was especially interested in the effect of culture on language, believing culture to be so pervasive that it generated different grammar structures in different languages, a hypothesis now known as linguistic relativity (Sapir, 1921, 119). Whorf took a similar view, saying "linguistics is essentially the quest for MEANING" (Whorf, 1956, 73). More dogmatic than Sapir, Whorf believed that culture affected not just language structure but the thought processes behind that structure. Basically, culture affected language by changing how a person thought. This stronger form of linguistic relativity is now known as linguistic determinism.

Meaning for SFL is both cognitive and communicative, offering the speaker/writer choices in the meanings they express. For instance, in the requests *give me a sandwich please*, *please may I have a sandwich*, and *sandwich please*, the speaker wants the same outcome. Cognitively, their meanings are similar, but communicatively they are quite different. The speaker is not transforming internal language into external language, they are communicating intentional meaning and suppressing unintended meanings. This is a major difference: Generative Grammar sees language as mostly subconsciously produced, with only a thin structural overlay of conscious choice; SFL grammar sees language as mostly about choice.

Another difference is that the mode of communication is trivial for Generative Grammar but can be significant for SFL. The sandwich requests above share an immediacy that makes them unlikely written forms, except as reported speech: the dialogue between writer and reader usually has a time-lag, so immediate requests cannot work. Thompson (1996, 6) says that Transformational-Generative Grammar "does not reflect how the users themselves view language. They respond above all to the meanings that are expressed and the ways in which those meanings are expressed".

The Systemic Functional Approach

In SFL, component meanings can be carried by individual words, or idiomatic word combinations, or novel word combinations; but holistic meaning is effected at the clause level, using both words (lexis) and grammar: utterance production occurs on a lexicogrammar continuum (Morley, 2000, 21). For instance, the word *cold* implies a temperature which is less than ideal; but in the construct *this isn't a cold fish* there are several semantic effects at work. First, the term *this* indicates to the receiver that the utterance is about something recently discussed or indicated by gesture. Second, the verb construct *isn't* implies that the actual thing that follows is not the thing being discussed. Third, the construct is ambiguous, with meaning determined by whatever is believed to be the comparator: is the *cold fish* significant because it should be cold, or a fish, and is not? Does the semantic relationship between *cold* and *fish* render them a single entity for this utterance? Or is *cold fish* being used in its idiomatic meaning of an unemotional person? Usually, the receiver uses context to interpret meaning, with the sender's choice of form indicating they believe the receiver has sufficient context to make a congruent interpretation. The sender may be wrong but, because language is a dialogue, the receiver can check their interpretation. Utterance context affects its meaning and therefore its analysis; and it also shows that congruent conversations require constant negotiation toward meaning.

SFL grammar interprets language clauses using modes of meaning grouped into three metafunctions: Interpersonal, Ideational and Textual. The Interpersonal metafunction evaluates the clause as an exchange, analysing the social transfer of information between minds. The Ideational metafunction evaluates the clause as a representation, analysing the way thought is represented in utterance and *vice versa*. The Textual metafunction evaluates the clause as a message, analysing the words and forms chosen for the utterance (Halliday, 1994, 34). Each metafunction carries a complementary thread of meaning, with clauses carrying all three metafunctions simultaneously.

The Ideational metafunction subdivides into two further metafunctions: Experiential, analysing representation inside the clause; and Logical, analysing representation between clauses. The Logical metafunction is a key difference between SFL and GL: the other metafunctions, like GL, analyse the clause; but the Logical metafunction analyses discourse – how clauses work together. The logical metafunction is perhaps Halliday's greatest contribution to linguistics; his descriptions of the other three metafunctions was prefigured by Harman (1968 [1971], p68):

> Theories of meaning may attempt to do any of three different things. One theory might attempt to explain what it is for a thought to be the thought that so-and-so, *etc*. Another might attempt to explain what it takes to communicate certain information. A third might offer an account of speech acts. As theories of language, the first would offer an account of the use of

> language in thinking; the second, an account of the use of language in communication; the third, an account of the use of language in certain institutions, rituals, or practices of a group of speakers.

Harman’s analysis closely, although not completely, corresponds to the Interpersonal, Ideational and Textual metafunctions.

<table>
<tr><th>Interpersonal</th><th colspan="2">Ideational</th><th>Textual</th></tr>
<tr><td>Socialisation</td><td colspan="2">Cognition</td><td>Communication</td></tr>
<tr><td rowspan="3">Constructing our relationships with others in our World</td><td colspan="2">Constructing our human experience of the World</td><td rowspan="3">Constructing our communication systems</td></tr>
<tr><td>Experiential</td><td>Logical</td></tr>
<tr><td>Constructing a model of our current experience</td><td>Merging our current experience into our known experience</td></tr>
</table>

Figure 4.1: Halliday’s Systemic Functional Metafunctions

To show SFL grammar at work, let us consider ways of asking for a drink. The textual metafunction works in terms of theme (the key item in the clause, usually the first word or phrase) and rheme (the rest of the clause). In the clause *give me a drink, please,* the theme is giving, establishing the clause as task-related. This compares with *please may I have a drink*, which asks the receiver to meet a need of the sender, rather than just do something. *Drink*, *please* emphasises the object itself rather than performing an action or meeting the sender’s needs. Interpersonally, the first clause establishes a service role for the receiver, the second requests that sender and receiver work together in a joint venture, and the last diminishes the sender and receiver roles, emphasising outcome instead. Experientially, all three clauses have an actor (you), a process (supplying a drink) and a circumstance (satisfying my thirst), and all should result in the same actual outcome. Experiential context is important, however: using a clause in the wrong context may result in a drink being thrown at me rather than offered to me.

Halliday’s multimodal analysis may appear unwieldy, but it is effective: prosodics and melodics (stress, intonation, tone and rhythm), largely ignored by Generativists, were brought back into linguistics; it emphasised discourse over sentences and utterances; and idiosyncratic language was once again central to linguistics. For instance, SFL grammar can identify differences between parataxis and hypotaxis using the Logical metafunction. Parataxis involves linking clausal elements at the same level by simple connection (*e.g., I went upstairs and got my hat, and then went to the market*); hypotaxis involves linking clausal elements hierarchically (*e.g., I went upstairs to get my hat so that I could go to the market*). Tannen (1994, chs3-5) has shown that this difference is identifiable in cross-cultural and cross-gender misunderstandings: men tend to use and be more comfortable with hypotaxis than parataxis, and women *vice versa*. This, though, is a trend, not a rule, which means Generativist approaches are unlikely to identfy it.

One problem for all grammaticians is cognitive dissonance, where the message understood by the receiver does not match the sender's intentions. This is not because the sender wishes to deceive the receiver, or because the signal was misspoken or misheard, this happens because the communicative meaning system is inadequate. The only way dissonance can occur in a formal language system, where internal representations are converted to external signals, is where the internal languages of sender and receiver are different. In a functional language system, however, dissonance can occur in the cultural assumptions of the sender or receiver. Functional conflicts between the different metafunctions of the utterance can mean the message received is not what the sender intended; or the different contexts of the sender and receiver can create different interpretations. SFL grammar gives us mechanisms to analyse the causes and outcomes of dissonances; Generative Grammar largely does not.

Other Functionalist Grammar Descriptions

Systemic Functional Grammar, although slightly younger than Transformational-Generative Grammar, continues a history of Functional approaches to language going back to the 1930s Prague Linguistic Circle. This remarkable group established the basics of phonology, signal theory, semantics and semiotics, and were the first researchers to consider linguistics as a rule-driven science (Mackenzie, 2016). The principles of the Prague Linguistic Circle have inspired several interrelated research programmes, and these in turn have produced their own theory structures. SFL grammar is arguably the main theoretical base in Functional Linguistics, but not the only one.

An approach initiated by Mulder in the late 1960s is Axiomatic Functionalism or AF (Hervey, 1979; Mulder & Hervey, 1980), which attempts to reconcile Functionalism with the axiomatic approach of Generativism, to give Functionalist analyses access to many of the tools of Generativism, particularly hierarchical tree structures. Like Generativism, AF is rule-driven and looks for regularity and commonalities in utterances. Also like Generativism, it is introspective, using judgements of well-formedness to identify rules. Unlike Generativism, however, it attempts to describe rule systems not just for syntax, not just for language, but for the entire semiotic process. It is an ambitious project.

Analysing *the quick brown fox jumped over the lazy dog* in AF produces the following. First, the nominal [*fox*] has dependencies of determiner [*the*] and adjective [*quick*] and [*brown*]; similarly the nominal [*dog*] has dependencies of determiner [*the*] and adjective [*lazy*]; the verb [*jump*] has dependencies of past tense [-*ed*] and adposition [*over*]; and it also has nominal dependencies of [*fox*] and [*dog*], forming the semiotic syntagm of the sentence. Like Generative analyses, this is essentially hierarchical, with syntagms at lower levels combining into higher syntagms; but where Generativist analysis is concerned with syntactic form, AF analysis is interested in how meaning is generated using signs.

The main difficulty with the AF approach is the project size (large) versus the research team (small). Where Generativism and SFL grammar quickly pass theoretical issues to teams of postgraduates, Axiomatic Functionalism has fewer human resources and therefore longer solution times. The large project and the small team conspire to make progress in Axiomatic Linguistics slow, although there is still progress (*e.g.*, Bican & Rastall, 2014; Dickins, 2009).

Functional Discourse Grammar (FDG) (Hengeveld & Mackenzie, 2008) is another approach with potential but few resources. Primarily explored in the 1980s (Dik, 1981), it proposes that the metafunctions of SFL grammar work together in a strict hierarchy to produce meaning. The interpersonal level represents the speaker's intentions, and this governs the representational level where meaning is introduced. This in turn governs the structural level, where grammar occurs; and the structural level governs the phonological level where sound or writing happens. This language production hierarchy differs significantly from Generativism in that grammar is a product of, and governed by, intention and meaning: meaning structure is more important than rule structure in FDG.

As the name suggests, FDG is discourse-bound and not word- or rule-bound. This emphasis produces analyses different from both Generative models and SFL grammar: pragmatics, the study of the contexts and assumptions in language, comes first; then semantics, then grammar, then phonology. Analysis is top-down: the sources of language production and the intention to mean must be established before grammar structures and sign patterns can be applied.

Lack of academic popularity means FDG makes progress slowly. However, despite its relative newness, it is already offering insights that other Functional theories do not, and is beginning to establish a firm and faithful following (*e.g.*, Giomi, 2020; Keizer, 2015).

Danish Functional Linguistics (DFL), established in 1989, is based on the work of Hjelmslev (Harder, 1996), who, in the 1930s, helped found the Copenhagen Linguistic Circle, a group inspired by the Prague Linguistic Circle. DFL takes an inclusive, rather than dogmatic, approach, and therefore includes a divergent set of theories. Hjelmslev (1961) emphasised phonological form, which he called Glossematics; but this has been supplemented by semantic and pragmatic threads, producing novel analyses across multiple languages (Fudge, 1995). The theoretical base may not be as strong as for other linguistic traditions, but DFL provides a non-judgemental haven where good research can and does happen, providing an exemplar for other academic initiatives.

Functional-Typological Linguistics (FTL) is the last methodology discussed here (Noonan, 1999). This is mostly a North American enterprise, although it is now used by linguists globally. A non-structuralist theory, it has strong links to Cognitive Linguistics (see chapter 5), which puts it in direct opposition to both Generativism and the structural aspects of Functionalist grammars. It is therefore similar to Integrationist Linguistics, which sees no need for universal rule systems in linguistic analysis (Toolan, 1996).

FTL is about linguistic processes. Using the traditional Generativist differentiation between linguistic competence (the capacity to produce language) and performance (the actual language produced), FTL is performance driven. Language is not a self-contained mechanism to transfer knowledge between minds, it is a form of knowledge itself and works like other forms of knowledge. The approach has proved to be of particular interest in second language acquisition studies (*e.g.*, Ramat, 1999).

FTL inevitably has a particular view on grammar: it is not a system dictating language production, it instead emerges from language use. Language grammar does have universals, but they come from sharing universal types of knowledge. Negotiating toward a common cognitive map imposes its own regularities on sharing, generating the illusion that regularities are part of language itself and not emergent from general cognition. According to FTL, this illusion has allowed linguists who want rule systems to find what is not actually there – more linguistic phlogiston. FTL reminds us that all language theories rely on assumptions, and these should not go unchallenged.

Speech is Linear, So Why Not Grammar?

Many more functional and structural linguistic traditions exist than are discussed here (*e.g.*, Lockwood's Stratificational-Cognitive Linguistics, 2002, or Hoey's Lexical Priming, 2005); but the traditions described above do share a common principle: variation is not just tolerated, it is encouraged and celebrated. Functional Linguistic models also share three other principles. First, language is about meaning, so meaning must be primary when identifying lexicogrammar rules. Second, language is about signalling, so the sender and receiver are intrinsic to the language model. Third, language is about signs, so manipulation of signs in language utterances must be a paramount consideration in establishing lexicogrammar rules.

Functional Linguistics can be as hierarchical as Generativist analyses, but the hierarchies are different. Generativism describes utterances in terms of dependencies, and it uses tree structures to describe the hierarchy of dependencies; Functionalism describes discourses, rather than utterances, in terms of threads of meaning simultaneously expressed throughout the discourse – and in the case of FDG, the threads are also hierarchically ordered. Functionalist analyses therefore replace the Generativist tiers of structure with simultaneous threads, but there is hierarchy between and within the threads.

One important difference remains, however: Generative analytical hierarchy is inherently binary, utterances are deconstructed two-dimensionally; yet speech, by nature, is a one-dimensional vector of sound; and writing, although expressed on a two-dimensional surface, retains this one-dimensionality. Functionalism identifies language as multi-vectored, but it analyses the vectors as separate one-dimensional threads – essentially, grammar, or grammars, without tiers.

The one-dimensionality of language utterances has, however, inspired several attempts to represent the linearity of the final state – speech or writing – as part of the language production process. Linear grammars challenge the assumption that language is essentially multidimensional, with linear form being imposed by the communication medium; instead, they see linearity as inherent throughout the language production process, making the final linear state a natural outcome and not a compromise. Linear grammars recognise the one-dimensional functionality of language that other grammars mostly ignore; a review of functional grammars would therefore be incomplete without some examples of linear grammar.

Linear Approaches to Language Grammar

Generative Grammar is a hierarchical description of nested constructs, definitely not linear. Functionalist grammars analyse utterances as a series of levels, or metafunctions; so, although each metafunction is essentially linear, they together describe utterances multidimensionally as hierarchical phrase structures. New evidence from child language acquisition is challenging this need for hierarchy, however (*e.g.*, Perfors *et al*., 2011); and stone toolmaking processes (*e.g.*, Stout *et al*., 2021), indicate that hierarchy, and not linearity, is learned, while linearity, and not hierarchy, is innate. Linear grammars therefore look at the relationships between neighbouring segments of an utterance first; and the relationships between non-contiguous linear segments only after this. Analysis thus differs between languages: for instance, the linear solution to *Alf hat Bette im Zug gesehen* is different from *Alf has seen Beth on the train*. Linear grammars tend to be less didactic about universal or standard templates, and they do not try to resolve language complexity in a single model (Croft, 2001, 202). Instead, they address aspects of language production with different models.

Kathol (2000), working in the Generativist tradition, builds a linear adjunct to binary hierarchy. He analyses sentences as both Generativist tree structures and linear Functionalist domains of meaning, showing that a hybrid approach informs and enriches analysis. His model drills down through the tree structure to the semantic units, and then analyses those units linearly. It also applies a linear analysis at the clause level, reflecting the SFL division of ideation into experiential and logical metafunctions. Sentential coherence is particularly important in this analysis.

Kathol tests his hybrid approach with Germanic languages – English, German, Swedish, Yiddish, Dutch, Icelandic, Danish and Norwegian – finding the linear analyses of each language to be similar, but not identical. Apparent similarities between the tree structure analyses of these languages rely on word order being treated as a surface expression of a standard deep form, rather than as a creator of meaning, arguably its actual role. It also does not accommodate the type of language: whether a language is mainly fusional (incorporating grammatical markers like tense and case into words) like German, mainly agglutinative (incorporating several meaning-units in one word) like Finnish, or mainly analytic (one marker or meaning per word) like English, affects the basic assumptions of tree structuring. When seeking an exemplar language for universal features, English (a rather extreme analytic language) may be a wrong choice, even though it is easy to analyse into trees.

Bod (1998) compares English and Dutch grammars to argue for a linear model he calls Stochastic Context-Free Grammar. He is concerned only with constructs likely to occur in real language usage; and he uses Generative transformations to convert non-terminal verb phrases (which can be further analysed) into non-terminal plus terminal verb phrases (which cannot be further analysed). These transformations change the theoretical sentence, *S*, into its produced form, *R*. Some constructs are common and mostly acceptable, some are rarer but still mostly acceptable, and some are seldom used and often unacceptable. However, it is impossible to define a construct as always acceptable or unacceptable because language grammar is probabilistic.

Using Data Oriented Parsing (DOP) techniques, Bod creates an analysis model with three levels: DOP1, where common, acceptable generative transformations occur; DOP2, using known constructs to predict the structure of new constructs encountered; and DOP3, using probability analysis to predict meaning from structure. Bod's approach is therefore semantic as well as syntactic, with the semantic interpretation produced linearly. Although Generative tree structures are used for DOP1 and DOP2, segments of meaning at DOP3 are analysed linearly.

Bod recognises two problems with computational analyses: the rule structure used by one individual may be different from that used by another; and the rule structures may differ between contexts even for a single individual. These differences need not cause communication between individuals to break down: the receiver's tolerances for comprehension can encompass improbable and even unacceptable forms. Because Stochastic Context-Free Grammar is probabilistic, grammar can be different for every user; Universal Grammar of the type Generativists propose is unnecessary (Bod, 1998, 145).

Perhaps the most effective linear grammar is Word Grammar (Hudson, 1998). This views language constructs as a series of heads controlling dependents. For instance, adjectives seem to be dependents of nouns, so nouns are therefore heads in noun phrases; and nouns seem to be dependents in verb and adpositional constructs, establishing a hierarchy of sequential word-bound constructs. This, however, is actually a single dimension of suspended expectations, not a two-dimensional cognitive hierarchy expressed in a one-dimensional medium.

For instance, when we hear the partial construct, *this is a matter of…*, our expectation is for a noun, not a noun phrase as in the Generativist model. The fact that we encounter an adjective, *...great…*, does not remove the expectation of a noun, merely suspends it: the adjective is a dependent of a noun so it is acceptable in this position. If we had encountered an adverb, *…really…*, then our expectation would then be for an adjective, further suspending our noun expectation. Grammaticality is not produced by well-formed grammatical trees, but by fulfilling our expectations of what should happen next. We see the construct above as ungrammatical or incomplete until we meet the noun we expect: *this is a matter of really great interest*. Word Grammar sees the single dimension of speech as a single dimension of meaning; it is not a two- (or multi-) dimensional grammar being squeezed into a single dimension.

While most work on Word Grammar has been in English, it has been successfully applied to other languages (*e.g.*, Creider, 2000); but, as Word Grammar is case-driven and explanatory rather than predictive, this does not require a search for universal principles. Emphasis on linearity as a first resort for analysing language provides an appealing alternative to the multi-dimensional complexities of functional analyses and Generative Grammar.

The linear grammars reviewed here are all data-driven: theory must emerge from evidence, not *vice versa*. They all accept that detailed grammatical structure is likely to be *ad hoc*, not subject to universal significance, although universal features can be imposed on language by non-linguistic cognition. Linear grammars also remind us that sometimes the obvious structure is all the structural explanation we need.

A theory which is not strictly linear, but which draws on the same linguistic approaches as linear grammars, has been proposed by Lamb (2016). This theory sees language as a network, both in terms of meanings and in terms of grammatical structure. While language is commonly seen as a network of meanings – represented by morphemes, words, and phrases – the idea of grammar as a network integrated into the semantic network is unusual. The standard approach, that language is a computational rule system inside the brain is, in neurological terms, inexplicable: the brain itself is a network, so any rule system must be contained within a network; to see it as a von Neumann-like machine is therefore counter-intuitive – especially as we have found no other system in the human brain which uses von Neumann architecture[6].

Lamb shows that the semantic network in the brain has both local representation (there is a neuronal unit dedicated to each concept) and distributed representation (there are links to other meanings throughout the brain which define the context of the concept). So, for instance, the concept CUP is instantiated in a single neuronal unit in the brain; but that neuronal unit has active links to other concepts (*e.g.*, DRINK, TEA, THIRST, *etc*.) which define the context of CUP, and which may link to other concepts (*e.g.*, MUG, TEAPOT) which can reinforce or suggest alternatives to the original concept. This certainly feels like the way that humans do their thinking, and it is supported by neurological studies.

Lamb's view, that the semantic network in the brain can simultaneously have local and distributed representation, challenges the traditional belief that nodes on a network either hold meaningful concepts themselves, or they

[6] Von Neumann architecture is a design principle used in most of our digital computers. It has the following components: a processing unit with an arithmetic logic unit and processor registers to hold intermediate computations; a control unit with an instruction register; rewriteable memory to store data and instructions; external mass storage; and input and output mechanisms to communicate with the humans.

are merely the way that meaningful connectors combine to make concepts. In the first traditional model, the concepts exist of themselves; in the second, concepts are always constructed from primitives. Lamb's approach models what we know of the structure of a neuron: it consists of a soma (a node) and dendrites which connect with dendrites of other neurons (connectors); it is not one or the other, it is both.

Functional Linguistics on the Sources of Language Grammar

Functionalism is providing answers for the sources of language grammar where Generativism is silent. This is partly because the inclusive nature of Functionalism encourages speculation and discovery: it accepts a wide range of ideas, there are no heresies. In comparison, Generativism is quite dogmatic: each new hypothesis has generated a range of para-hypotheses, many of which have been condemned as non-canonical, or forced into apostasy by a change in the main hypothesis (Harris, 1993). As a result, many of the controversial issues in Generativism have not been fully debated to a satisfactory conclusion.

For instance, phrase structure grammar is a good description of language structure, but a poor description of language production. If production followed Transformational Grammar methods we would expect chunks lowest in the construct hierarchy to be evaluated first, so they can be slotted into the higher levels. This seems counterintuitive to what happens. For instance, the sentence, *This is the dog that chased the cat that caught the rat that ate the malt that lay in the house that Jack built*, would have to be reverse-engineered: we must evaluate the nature of the house (*that Jack built*) before we can evaluate the malt (*that lay in the house that Jack built*), and so on. This begs the question why it is ordered as it is; why not start with the part needed first? Halliday's theme and rheme structure in the Textual metafunction gives a better explanation of how this construct works: each theme except the first – the cat, the rat, the malt, the house – is also the rheme of the previous part. The regular structure, using the logical linker *that*, creates an expectation of semantic nestedness. Visualising concepts containing concepts in a top-down way is easier than constructing a bottom-up structure of grammatical forms contained by grammatical forms.

However, one major problem with Functional grammars is that they make few predictions about language innateness or the sources of language grammar: Functional Linguistics does not require an explanation of these matters, unlike the Universal Grammar of Generativism. Considerable work has been done in describing child language acquisition in an SFL model (Peters, 1995; Craig, 1995; Gaylard, 1995; Torr, 1997, among others), and SFL grammar itself came out of Halliday's analysis of his own son Nigel's early linguistic acquisition (Halliday, 2004). Other researchers of child language acquisition, while not officially working in a Functionalist framework, adopt a pragmatic approach based around evidence (*e.g.*, Hirsh-Pasek & Golinkoff, 1996; Chiat, 2000); and they find the openness of Functional theories fits their evidential approach better than the constrained theories of Generativism.

Language origin theories also tend to be pragmatic and evidential (although the evidence is indirect), so they fit better with Functionalism than Generativism. There is, however, remarkably little literature that takes a strictly Functionalist perspective on language origins. A volume of four papers (Benson & Greaves, 2005) does analyse the effectiveness of some nonhumans using a gestural form of human language, particularly the bonobo Kanzi and others at the Yerkes Primate Centre. Their gestural communication uses several Functionalist methods, including interpersonal and ideational metafunctions, the syntactic element of the textual metafunction, and structured phonology. The papers show that human language metafunctions have precursors in nonhuman signalling; and, for human-language-trained primates, the precursors provide an effective communication bridge between them and humans. Primates may not debate the relative merits of different linguistic theories, but they can reliably receive information from us and transmit their own information back.

Rose (2006) proposes some principles for a Functionalist approach to language evolution – although he is more concerned with evolution from the first language (grammaticalization) than with evolution to the first language.

On the evolution to language, he proposes four stages. First, complex cultural behaviours must be able to develop over time, transferring between generations through teaching and learning. Second, there must be an identifiable way for primate vocalisations to evolve into wordlike units (satisfying the Textual Metafunction). Third, the wordlike units must be able to progress from socially transferring information between minds (satisfying the Interpersonal Metafunction), to representing the speaker's experience (satisfying the Experiential Metafunction). Finally, complex patterns of discourse must emerge, allowing stable representation between utterances as well as within them (satisfying the Logical Metafunction). The order in which the metafunctions are activated is significant. I have elsewhere argued that the only metafunction active in signalling by other animals is the Experiential Metafunction (Edwardes, 2005), which Rose places last in the development of metafunctions. This proves neither Rose nor me right, but it does show that alternative viewpoints are arguable.

Functionalist lack of interest in the sources of language grammar is, however, more apparent than real. Many researchers using Functional methodology to analyse discourse use a related methodology, Cognitive Linguistics, to work with language more abstractly. Not all Functionalists are also Cognitivists, and not all Cognitivists are also Functionalists; but, because the theories start from similar assumptions and reach similar conclusions, the overlap is wide. As we see next, Cognitive Linguistics does offer a lot to linguists interested in the sources of language grammar.

5 Cognitivism and Sources of Language Grammar

We know that someone who has channelled his anger into something constructive has not had a cow. How do we know these things?
(George Lakoff)

Generativism and Functionalism seem to be only superficially interested in language origins. However, as linguistics traditionally studies current language use, and historical linguistics needs written evidence of a language, this is not unreasonable. Speech leaves no record, so languages spoken before sound recording became available are somewhat conjectural; and before writing they are highly conjectural. Linguistics probably cannot access the early history of language (Diamond, 2011).

One branch of linguistics directly interested in the sources of language grammar is Cognitive Linguistics. Younger than Generativism or Systemic Functionalism, it has nonetheless survived and thrived for over 50 years, becoming a mature discipline. It is interested in language as a social, cognitive, and communicative instrument, evidencing its theories with examples of actual language usage; and it has much to offer in the study of language grammar causation.

Cognitive Linguistics is different from Cognitive Psychology; although they share a common adjective and a common origin, they have taken different directions. They both use scientific method to check theory against reality, and they both study cognitive processes in terms of inputs and outputs. However, Cognitive Psychology, like Generative Linguistics, emphasises the mind as a computational mechanism (Eysenck & Keane, 1995, 1), while Cognitive Linguistics emphasises the mind as a meaning mechanism (Geeraerts, 2006, 3). This difference is not trivial, it produces different models of cognition. For Cognitive Psychology, the mind's processes are largely modular and reactive, interpreting inputs and producing outputs in a predictable way; and language is a separate cognitive module so can be examined separately. For Cognitive Linguistics, language is distributed throughout the mind, and individual minds are idiosyncratic and unpredictable; language must therefore be examined concurrently with other cognitive processes.

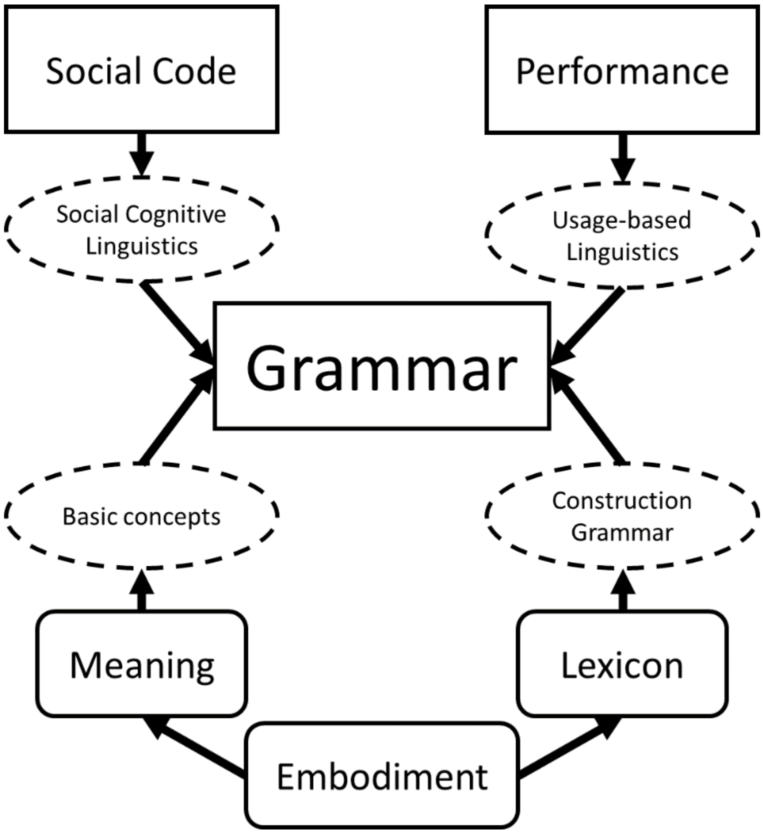

Figure 5.1: A Possible Interpretation of the Components of Cognitive Linguistics (Based on Geeraerts, 2010)

A Short Introduction to Cognitive Linguistics

Although Cognitive Linguistics was established in the 1970s, cognitive modelling was only recognised as relevant linguistic tool in the late 1980s. Before then, Chafe (1970), Fillmore (1976) and Lakoff & Johnson (1980) had worked on semantic systems in language; and Langacker (1987 & 1991) had worked on syntax. These research initiatives were pursued in parallel with each other and alongside Functionalist approaches. However, the early cognitive initiatives all shared the view that language is a feature of general cognition, and opposed the view that language is a unique and isolated phenomenon. Cognitive Linguistics and Generative Linguistics are therefore founded on markedly antithetical principles.

The theories of Cognitive Linguistics began to be codified in the 1980s: Fillmore worked with Lakoff on the foundations of Construction Grammar, and Langacker published the first volume of his two-volume work on Cognitive Grammar (1987). Lakoff also published his key work on metaphor and embodiment in language, *Women, Fire and Dangerous Things* (1987). In 1989, the Cognitive Linguistics community held its first large conference in Duisberg, Germany, and the International Cognitive Linguistic Association was founded. The first community journal, *Cognitive Linguistics*, began publication in 1990.

Cognitive Linguistics continued to grow throughout the 1990s, supported by important technological developments, such as functional Magnetic Resonance Imaging (fMRI). fMRI uses magnetic fields to measure the flow of blood into and out of brain areas, and it works by detecting changes in blood oxygenation caused by neural activity. Oxygenated blood is directed to a brain area to fuel neural activity in that area, and deoxygenated blood (containing carbon dioxide instead of oxygen) is removed from the area, creating what is known as a Blood Oxygen-Level Dependent (BOLD) signal. This process can be observed magnetically because the two gasses have different magnetic signatures, and in this way the brain activity that corresponds to a physical or cognitive activity can be mapped. It is relatively easy to put subjects into an fMRI scanner, ask them to complete particular physical or mental exercises, and see which areas of the brain activate (Loued-Khenissi *et al*., 2018). With an image capture rate of up to 60 images a minute, fMRI in the 1990s allowed fast, simple and relatively cheap imaging of the human brain, although detail was limited. Since then, the size of a voxel (the smallest distinguishable volume scannable) has reduced from $3mm^3$ to $0.5mm^3$, reducing voxel volume from 27,000 nanolitres to 125 nanolitres. This has greatly increased sensitivity and allows for more nuanced analyses (Bollmann & Barth, 2021).

With fMRI, Cognitive Linguistic theories could be tested against what happened in actual brains, and the results did not always support the theories. The first theory to be questioned was modularity, the idea that certain parts of the brain are dedicated to language: scans showed large areas of the brain involved in even simple linguistic processing, and in both hemispheres. While there is evidence in many people that some brain areas are used preferentially for language, they do not correspond closely – or at all – between individuals (Thulborn *et al*., 1999). Language cognition seemed to be less independent of general cognition, and less universally organised, than Generativism had predicted.

In a more recent study, Gagliardi *et al*. (2018) have shown that Williams Syndrome does not involve a virtually intact language centre with aberrant social cognition, as some Generativists used to believe; instead, the Williams Syndrome brain is extensively differently wired. Similarly, the atypical brains of some left-handers or ambidextrous people have been shown as quite different from neurotypical right-handed brains. It was believed that handedness could create a mirrored brain, where left-hand dominance caused systems in the two hemispheres to swap, thus enabling left-handed people to have language in the right hemisphere. However, Biduła *et al*. (2017) showed that language activated the left hemispheres of atypical brains regardless of handedness, but atypical individuals used brain areas in both hemispheres and in unique ways. There seems to be more than one way for a human brain to process language.

The fMRI evidence began to challenge some tenets of Generative Linguistics and led indirectly to the development of Chomsky’s Minimalist Program. This, as we saw in chapter 3, attempted to redefine the nature of language, making it a conceptual universal of the mind rather than a physical universal of the brain. The effect of Minimalism on Generativism has already been covered, but it also had a significant effect on other linguistic traditions. It changed Generativism’s view that language is largely hard-wired and physically modular, and it added plausibility to some Cognitive Linguistics theories. For instance, The Generativist principle that linguistic rules and systems must be essentially abstract (*e.g.*, Lupyan & Winter, 2018) has been challenged by Behrens (2021). Comparing constructs in Dutch, German and English, Behrens argues that:

> Constructivist approaches question the underlying motivations of structuralist and generativist linguistics, that linguistic generalisations have to be as powerful as possible. If children learn from local clusters, and manage to generalise on the basis of quite concrete form-function associations, are "abstract" generalisations needed and attested at all, and what is their added value […] ? In other words: Are the generalisations that linguists derive the same as those made by speakers, and is processing always based on the most abstract available representation? (977-978)

Cognitive Linguistics continues to gain adherents in linguistics and associated disciplines. Tomasello's work on nonhuman signalling (*e.g.*, Tomasello & Call, 1997) and human child communication (*e.g.*, Tomasello, 2003a) has provided significant support for Cognitive Linguistics, as has the work of Steels on robotic communication (*e.g.*, Steels, 1998). Croft (2001) has continued to codify Cognitive Grammar, and Evans & Green (2006) have unified several research initiatives in Cognitive Linguistics into a coherent enterprise. Cognitive Linguistics now provides a viable meaning-based alternative to the rule-governed doctrine of Generative Linguistics.

The Cognitive Approach to Linguistics

Cognitive Linguistics has many threads, each representing a cognitive approach to a particular aspect of language. Cognitive linguists are involved in all the traditional divisions of linguistics – semantics, pragmatics, grammar and phonology – which means that Cognitive Linguistics research is wide-ranging, from niche usages of a particular language to broad investigations into the phenomenon of language. Cognitive Linguistics is not interested in a single macro-theory to explain everything about language, so there is no need for hypotheses and theories to conform to a single ideology. Instead, Cognitive Linguistics often entertains and works with multiple alternative solutions (*e.g.*, Wilson, 2011); it is not interested in how language should work but how it does work.

Because of this, Cognitive Linguistics has often produced novel approaches to the structure or usage of languages, and to the relationships between language and other cognition. For instance, Palmer (2006) shows that cognitive models of language acquisition and usage have informed and enriched anthropological models of cultural acquisition, and *vice versa*; and a Cognitive Linguistic approach to our expression of temporality shows how we think about and understand time (Evans, 2005). For Cognitive Linguistics, language is only part of the issue; and the explanations found tend to explain more than just how language works. Cognitive Linguistics aims to describe language in terms of the mind and the brain, not relying on unevidenced hypotheses; and to analyse linguistic evidence in terms of general cognition, rather that language-specific features (Evans *et al*., 2007). These two aims give the Cognitive Linguistics enterprise a distinctive trajectory.

Language features of particular interest to Cognitivists include metaphor and metonymy. Metaphor identifies systemic features from one model of reality and applies them to another, thus letting us understand the second reality differently. For instance, interactions between individuals can also be used to describe interactions between groups; so, nations having relationships with each other – such as "special friendships" and "understandings" – emerges naturally from the idea that groups are themselves entities as well as being composed of entities. This metaphor, treating the group as an entity, is explored in more detail in chapter 8.

Metonymy is a specific form of metaphor, allowing objects to be arbitrarily associated with labels that are only indirectly referential (Lakoff & Johnson, 1980, 35-40). Metaphor and metonymy are products of arbitrary associations in language between intention, concept and utterance, with two cognitive mechanisms between the three components. The first mechanism is relational, between the intention and the concept; and the second is translational, between the concept and the utterance. For instance, when I use the idiom *shut your cakehole* I am relying on the context and culture I share with the listener; and I am also relying on the listener's recognition of my intended meaning (cakehole equals mouth). My direct intention would appear to be identifying a mouth in its role as a conduit for eating, because of my food reference. The phrase *shut your cakehole*, however, is actually intended to refer to the mouth in its role as a conduit for speech, making this particular usage metonymous as well as metaphorical: I am indirectly suggesting to the listener that they should use their mouth primarily for eating and not for communication.

Metaphor has an important role in language, it is not just a poetic trope. For instance, in the previous sentence, the word *metaphor* is metonymous: I am using it as an entity with agency ("Metaphor plays…"). This usage, however, is just a cognitive construct, it has no reality beyond what the speaker and listener agree it should have. Some physical relationships in the actual world are so pervasive that they provide a rich source of metaphor; Lakoff and Johnson refer to these as *experiential metaphors*, marking them with capitalised text. For instance, the metaphorical equivalence that MORE IS UP comes out of the observation that bigger piles are also higher: the metaphor is governed by a universal human experience of gravity. This primary metaphor generates a second, weaker, metaphor that MORE IS BETTER, because having too much is better than having too little. MORE IS BETTER can then merge with MORE IS UP to give UP IS BETTER, and this is represented in cognitively generated relationships such as *heaven* and *hell*. However, UP IS BETTER does not have an external actuality like gravity to support it, so MORE IS UP can take precedence. For instance, *a rise in crime* or *a growing fear* are examples of MORE IS UP, but not of UP IS BETTER. Metaphor produces overlays of new meanings on top of our semantic network.

Metaphor also affects how we view grammar. Verbs, for example, can be treated as a single word-type because their structural relationships with other words in a sentence is often identical. For instance, the sentences, *Alf is the man in the hat*, *Alf saw the man in the hat*, *Alf greeted the man in the hat*, *Alf chose the man in the hat*, and *Alf moved away from the man in the hat*, can be viewed as sharing a common grammatical form of *Alf Xs the man in the hat* – even though the semantic relationship they establish between *Alf* and *the man* is, in each case, different. Yet even here, the reflexive nature of *is* in the first sentence, making *Alf* and *the man* the same entity, is markedly different from the non-reflexivity of the other sentences. So, are all verbs truly functionally similar? This question is of interest to Cognitive linguists.

The metaphorical nature of word classes often goes unrecognised. For instance, the adjectives in *happy teacher* and *drama teacher* seem to be in the same class; yet we say *the teacher is happy*, but not **the teacher is drama*, and we say *teacher of drama* but not **teacher of happy*. Should we class drama as a descriptive noun? But then why is a *drama teacher* also a *teacher of drama*, but a *university teacher* is not a **teacher of university*? It seems that there are more word-class rules that there are word classes – labels like *adjective* and *noun* do not tell the whole story. Metaphorical grammar does not compress words into universal classes which ignore what they actually do, it instead accommodates the way actual languages work (Dąbrowska, 2004, ch8).

Although Cognitive Linguistics is primarily concerned with language as cognition, it does not ignore interpersonal aspects: if language is about human cognition, then human attention to social interaction must be accommodated; Cognitive Linguistics recognises that meanings include social information. For instance, the Western timeline with the future in front and the past behind is not universal: the Aymara of South America put the past in front (where it is immediately visible) and the future behind (which can be seen only by physically changing viewpoint) (Núñez & Sweetser, 2006). This mapping difference shows that moving forward into the future is not universal, and the future can be imagined into any physical place except the current location of the speaker: ahead, behind, above, below, to left or right, all can be treated as concrete locations for an abstract position in time. Yet there is still a universal: wherever the future is, the past is always in the opposite direction.

Cognitive Linguistics is also interested in grammaticalization, how the grammar of a language changes over time (Hopper & Traugott, 1993; Lehmann, 2015). For Cognitivists, grammar changes reflect changes in word meanings and usages. For instance, the auxiliary *will* as an indicator of future tense seems to come from its Middle English meaning of *intend to*; while the other future tense indicator, *be going to*, extends the spatial meaning of *go* into temporality, and seems to come from early modern English usage. As simple future tenses, they have come to have slightly different meanings. *I am going to do it* seems to have greater imminence than *I will do it*, but when used together they have come to represent a future intention to act: *I will be going to do it*. The form **I am going to will do it* is never used, although it is relatively easy to work out what it should mean. The way language grammar changes, and therefore why it changes, requires cognitive, semantic and cultural explanations.

Cognitive Linguistics also uses frames and schemata extensively, a metacognitive approach which originated in psychology (Eysenck, 1993, 33-38). In a frames and schemata analysis, every utterance is comprehended contextually (the frame), which includes the sender's intentions, the receiver's interpretations and their shared pre-existing knowledge. The information in the utterance provides the schema, which can be innate embodied knowledge, innate or acquired environmental knowledge, and acquired social or cultural knowledge. So, the phrase *I'm left-handed* does not just inform you about one of my hands, it gives information about my hand preference for writing, the possibility of mental and physical traits statistically linked to left-handedness, my likely social stance about handedness, my social conformity … that short phrase can contain a lot of information. The frame can indicate solidarity (I'm in that minority, too), differentiation (I'm in that minority and you are not), capacity (I have appropriate dexterity to do the task), incapacity (all left-handers are naturally awkward), identity (this is one of my key traits), and so on. A phrase has only trivial meaning in isolation; but in context it can contain multiple levels of information. Frames and schemata emphasise the communicative nature of language, subordinating linguistic structure to the needs of meaning.

Handedness can be important culturally: right-handed dominance has been appropriated in some cultures to make minority handedness unacceptable (McManus, 2009). The genetics of handedness can also be significant: while individual handedness has been observed in other primates, there is little evidence for species-wide preferences, except in humans (Uomini, 2009); and, while right-handedness often puts manipulation and language in the same left hemisphere of the brain, there is no strong correlation between handedness location, language location, and manipulative or linguistic facility (Papadatou-Pastou et al., 2020).

Human brains commonly lateralise some cognitive specialisations, such as language, into just one of the two hemispheres; but why this should be so is not clear. It seems that the most heavily lateralised cognitive skills are both learned rather than innate, and modern rather than traditional – for instance, modelling, planning, Theory of Mind, awareness of self, and socialisation (Hartwigsen *et al*., 2021). This is probably because learned skills gravitate to processing areas suited to their needs, and not because certain brain areas are evolutionarily designed to handle particular skills. Olulade *et al*. (2020) provide evidence that we are not born with a left hemisphere specialisation for language, it develops as we develop. Children aged 4-6 years use both hemispheres equally in linguistic tasks; but right hemisphere activity often reduces as we age, with 60% of adults having no significant right hemisphere activation for language. However, there also seems to be a slow but steady reversal of language lateralisation after age 25, indicating that hemispheric lateralisation is both less fixed and less necessary for effective language use (Szaflarski *et al*., 2006); our hemispheric specialisation is not developmentally preordained, not wholly reliant on our genes, and not fixed during adulthood. It seems there is a particular developmental path that most human brains take, and other paths that only some take.

A considerable literature has been generated on the topic of hemispheric lateralisation. Gotts *et al*. (2013) identify a communicative lateralisation in humans between the left and right cerebral hemispheres: the left hemisphere showing a preference to interact more exclusively with itself, particularly when dealing with modelling and fine motor coordination, while the right hemisphere interacts more equally with both hemispheres. However, this maps more to decision-making than to language. If most decisions are made in the left hemisphere (Vallar *et al.,* 1988) then the left hemisphere is the terminal point for most cognition: interactions that start in the right hemisphere tend to end in the left, and interactions that start in the left hemisphere tend to terminate there, too. The homogenous interactions of the left hemisphere and the heterogenous interactions of the right may not be reliable signifiers of processing differences between the two hemispheres.

Karolis *et al*. (2019), mapping hemispheric asymmetries in the human brain, identified four asymmetrical axes: symbolic communication, perception/action, emotion, and decision-making. They also found that asymmetrical cortical regions (those that do different work in different hemispheres) have fewer connections with the opposite hemisphere. They suggest this has the effect of reducing noise in the brain and mitigating conduction delays between the hemispheres.

Güntürkün *et al*. (2020) see cognitive lateralisation as typical throughout the animal kingdom. It starts with asymmetrical genetic expression, where copies of the same gene produce different effects depending on their context. This asymmetry generates cells with different roles and different activation in different hemispheres. Most of the human neural organisation of perception, cognition, emotion and action is lateralised differentially, so if we are to understand human cognition then we must study the asymmetries of our human brains; but we should consider those asymmetries as typical, and not exceptional, in nature.

In a recent study, Rogers (2021) also identified individual lateralisation as common in nonhumans: simultaneous semi-independent cognition streams allow an individual to attend to two things at once – such as feeding and predator avoidance. However, while it is common to find undirected lateralisation (the lateralised system can be in either hemisphere) in nonhuman individuals, genetically moderated directional lateralisation at the species level (where the lateralised system tends to occur in only one hemisphere) is rare. Rogers takes the view that this is because directional lateralisation provides no advantages over individual lateralisation – except in social interactions, where similarly skewed brains work together better. For Rogers, genetically moderated directional lateralisation is explicable only as a system of socialisation.

These different approaches to cognitive lateralisation illustrate the undogmatic nature of Cognitive Linguistics: it is flexible, in terms of both theory and accepted adherents. Unlike other schools of linguistics, it includes many scientists whose interests are not primarily linguistic, so it can encompass any topic which has elements of language and cognition.

Cognitive Linguistics is fundamentally a cognitive theory about language, where Generative Linguistics is a linguistic theory about cognition.

The Body in Cognition and Language

Embodiment, as mentioned in chapter 2, is significantly linked to cognition, and therefore important in Cognitive Linguistics, which views the brain primarily as a regulatory node for the body. The brain generates emotions via hormonal flows, physical action via the nervous system, consciousness via the capacity to model, and, in humans at least, metacognition via models of thinking itself – thinking about thinking. The brain performs most of its activity autonomously, governing unconscious or subconscious activity with innate models of the body in its environment; but it also uses conscious iterative modelling, creating an awareness of self as both model and modeller, a process examined further in chapter 7.

For Cognitive Linguistics, the body is the reason for cognition, and where it happens. Our capacity for abstract thinking relies on our physical relationship with the world because our cognition is grounded in actuality (Johnson, 1987). While embodied cognition seems obvious, some other theories see human cognition as an adjunct to the physical phenotype – interpreting, but not part of, physical existence. Even Empiricism, which sees all cognition as experience-based, identifies cognition itself as different from things experienced, leading Bishop Berkeley to propose that the external world has actuality only when being thought about (Berkeley, 1710 [2003]). Fortunately, for Berkeley, God is available to do all the thinking needed to keep everything in existence.

This assumed dichotomy between mind and body is insidious: while recognising that we are clearly holistic objects, governed wholly by our DNA program, we nonetheless see our “selves” riding around in our heads, controlling our internal machine in the same way as external machines. Cognitive Linguistics views this dichotomy as spurious: thinking and thinking-about-thinking have the same nature and use the same systems. Feeling they are different does not make them so.

In evolutionary terms, cognition involves supervising and improving the body's surviving and thriving: it improved the fitness of our forebears and thus enhanced their reproduction. Cognition is embodied because it is primarily about our physical selves and our body-related processes, as we saw in the metaphors of gravity and the concrete spatial directions abstracted to describe time passing. When faced with non-physical abstractions we rely on physical metaphors to allow us to share meaning: we *have* and *exchange* thoughts, we *produce* ideas, we *seek* agreement: it is all about what we would do if the abstractions were physical objects in a physical landscape (Gibbs & Perlman, 2006).

Cognitive Linguistics relies on embodiment: it is the doctrine that separates Cognitive Linguistics from other language philosophies. When studying language as a cognitive outcome there is no value in seeing language as a set of rules and regularities; behind (and containing) language is cognition, and behind (and containing) cognition is embodiment. This is where research into language must begin.

Is the Brain Modular?

One issue that divides linguists – and cognitive scientists of all kinds – is whether the brain has discrete, specialised modules, or a more *ad hoc* organisation. Do some types of cognition need dedicated brain areas or do they gravitate toward particular areas on a functional basis? Few people nowadays believe in the fully free-form brain of Locke's (1689 [1998]) *tabula rasa* – and those that do must refute good evidence supporting localised areas for specialised thinking in the brain (Pinker, 2002). Brain scanning studies, using systems such as fMRI, show that, for a majority of people, a significant proportion of language processing occurs in the left hemisphere of the brain (Jung-Beeman, 2005), in cortex regions known as Broca's area (Embick *et al*., 2000), Wernicke's area (Spocter *et al*., 2010) and the arcuate fasciculus (Rilling *et al*., 2008).

However, the terms *majority* and *significant proportion* indicate this evidence is statistical. As discussed earlier (Olulade *et al*., 2020), the brain areas used in communication seem to be widely dispersed across the brain at birth; yet for most adults there is no detectable cognitive activity for language in the right hemisphere. Left-side lateralisation for language is more common in right-handers than in left-, and in males than in females (Carter, 1998, 46-47), making language seem more like a functional subsystem of general cognition than an innate, independent brain module.

Large-scale brain modularity was first suggested by Fodor (1983). He proposed that the modularity of physical control systems, such as for movement, vision and hearing, was part of a general cognitive modularity for things like language, memory, selfhood and Theory of Mind. There was already some evidence for language modularity: in the 19th century, Pierre Paul Broca and Carl Wernicke identified correlations between language difficulties and left-hemisphere brain lesions in the pars triangularis and superior temporal gyrus, and other language issues were later correlated with the left-hemisphere angular gyrus (Bozic *et al.*, 2010) and the arcuate fasciculus (López-Barroso *et al.*, 2013). However, treating this as a human-only, language-only phenomenon is misleading: a study of captive chimpanzee oro-facial motor control for vocalising indicates that innate calls are mostly lateralised to the right hemisphere; in contrast, learned sounds are mostly lateralised to the left hemisphere, which is the location for language in most humans. The preferential lateralisation of human language may have a deeper, pre-human cause (Losin *et al.*, 2008).

At the statistical level, Fodor's weak modularity has proved correct: selfhood, Theory of Mind, and planning have all been correlated with different parts of the cortex. There do indeed seem to be functions that largely congregate in the same places in different human brains, making the cortex less generalised and more modular than a blank slate approach can explain; but, at the individual level, plasticity of brain development remains significant.

Fodor (2008) has more recently revisited his approach, moving away from weak modularity toward a more Nativist view of linguistic conceptualisation and a more Generativist model of language. He took the view that most of our everyday concepts (*e.g.*, HORSE, TREE, CARBURETTOR) are innate (we are born with them), and they must therefore be unlearned if we do not need them. Concepts cannot be learnable because to learn a concept requires pre-existing knowledge of the conditions in which the concept is needed; and knowing when something is needed but not what is needed is counter-intuitive. It is a philosophically defensible position, but it produces some weird effects outside of philosophy. For instance, concepts that we need today (like carburettor) must have been innate for hundreds of generations before they became useful. All those concepts must have been unlearned by previous generations, which means that, as human culture seems to have a continually-increasing need for new concepts, earlier generations had a much bigger burden of unlearning to do than modern humans. In fact, the earlier the generation, the bigger was the task of unlearning, which also seems counter-intuitive.

A stronger modular approach has been adopted by evolutionary psychologists. Cosmides & Tooby (1994) describe the brain as like a Swiss Army knife, consisting of a range of specialised components “each of which is well designed for solving a different problem” (p60). Natural selection should favour individuals with brains specialised for the fitness challenges they encounter: everything in the brain has a purpose, and we need to view the brain in terms of the problems it solves. Problems are locks, and solutions are keys; and each key must be different because each lock is different. The human brain, however, has a feature not covered by this ironware metaphor: plasticity (Willis *et al.*, 2009).

Explaining cognition in terms of evolution is commendable; but the strong modular approach brings its own problems, especially in relation to language. First is exaptation: the current role of a brain area may not be why it evolved. For instance, dyslexia is a recognised brain-related issue, with the size of the left *planum temporale* implicated in its severity – although without good evidence of its implication in the cause of dyslexia (Bishop & Snowling, 2004). However, reading and writing have been widespread human activities for 500 years at most, and the *planum temporale* cannot have evolved in the hope that one day humans would have widespread literacy. The discovery of humanlike asymmetry in the *planum temporale* of other Great Apes indicates that it probably has other functions which make it suited to exaptation for literacy (Hopkins *et al*., 1998); and exaptation into its new role is more likely that prospective evolution. Stroke recovery provides more evidence against strongly modularity: stroke damage to defined language areas can severely diminish language capacity in the short term, and often in the long-term, too. Yet recovery of capacity without recovery of the affected language areas is quite common: some brain “rewiring” seems to occur naturally when needed (Mosch *et al*., 2005).

Perhaps the most telling argument against Evolutionary Psychology’s strong modularity is the complexity required by a fully modular system. Humans, like all mammals, come in a wide variety of cognitive models, with natural variations creating notable differences in brain function. These include handedness, sexuality, manipulative dexterity and imaging accuracy. Additionally, some people quickly learn cognitive skills like mathematics, while others do not; yet we know these skills can be learned throughout life, they do not have a critical period for acquisition (FitzSimons, 2019). If all human capacities are prefigured in our brains then we would expect extensive redundancy in pre-adult brains and large variations between adult brains – but we do not see it. Instead, we see a high level of cognitive specialisation made possible by Metaplasticity, the increased neuroplasticity that seems to be available only to modern humans (Roberts, 2015). The argument for strong modularity remains, at best, unproven.

Cognitive modularity has many critics. Samuels (2006) considered how a massively modular human brain would work: how would the separate modules for different types of thinking work together to produce the appearance of a single mind? He concluded that weak modularity is supported by the modular behaviour of peripheral areas of cognition, such as the receptive senses and the productive motor systems; but the central systems of cognition, those involved in reasoning and decision-making, do not work in a modular way, acting instead as a highly integrated network.

Luppi *et al*. (2022) look at the synergy and redundancy in interactions between regions of the human brain. They identify an important difference between the two types of interaction: redundant interactions (those where the brain regions reinforce each other by containing the same information) are predominantly associated with modular sensorimotor processing, while synergistic interactions (those where the brain regions contain different information, so that bringing them together generates new information) are associated with integrative processes and complex cognition – thus mapping redundant interactions to Samuels' peripheral areas, and synergistic interactions to Samuels' central systems. Once again, the evidence supports a mixed methods model: strong modularity is present, but only in peripheral, redundant processes; the central, synergistic processes display, at best, only weak modularity. The advantage of having both types of modularity is that redundancy makes cognition robust, while synergy makes cognition integrated. This also answers Samuels' question: a massively modular mind is unlikely to be innovative, while a massively unmodular mind is unlikely to be consistent.

Looking at toolmaking by early humans, Stout *et al*. (2021) identify what they describe as grammars of tool manufacture. They suggest that making a particular type of tool was not the product of a single module dedicated to the task, it required the integration of a range of skills each generated by their own cognitive module. Modularity is present, but trivial: the tool is produced by the integration of skill-schemata into a frame, and then the execution of the frame – a form of weak modularity which is also hierarchical. Based on the work of Stout *et al*., Pain (2022) argues that the overlap between grammars of tool manufacture and grammars of language is sufficient to indicate a correlation between the two processes; and that, in turn, indicates that grammar probably developed from both improvements to the complexity of communicative procedures and increases in their number.

The Cognitive Linguistics approach to modularity is that, while some modules need strong modularity and weak interaction, the human brain's typical modules are functional – dependent and interactive – rather than structural – independent and self-contained. Language modules in particular must be integrated and leaky because other cognitive systems need access to them: language is a subset of, and highly dependent on, general cognition. Language cannot be studied as a self-contained entity, it is in a network of cognitive functions which cannot be properly understood without each other. Language did not evolve independently as a Generative solution to a cognitive need or a functional solution to a communicative need, it is a function of general cognition exapted for communication. Exchanging ideas improved the survivability of the group, and therefore improved the survivability of the individuals in the group. As a successful strategy, exchanging ideas became subject to natural selection, evolving under its own selection pressures.

The Cognitive Approach to Grammar

Cognitive Linguistics views grammar differently from other linguistics methodologies: as a tool of communication, not as an organising principle of language. While it certainly does organise language, this is just a secondary function of grammar; the primary function is to produce and comprehend meanings – to negotiate toward meaning.

Cognitive Linguistics has several different approaches to grammar, but one methodology is becoming the tool of choice for linguistic analysis. Croft's Radical Construction Grammar (RCG) (2001) is a direct descendant of Fillmore's Construction Grammar, but also incorporates features of Langacker's Cognitive Grammar. It offers a consistent and comprehensive grammar framework, and is used here to show the key features of Cognitive Linguistics grammars.

Croft presents RCG as a radical alternative to Generativist Minimalism. Instead of universal syntactic structures, Croft sees only diversity. Grammar comes from *ad hoc* production choices made in forming an utterance, which are dictated by the needs of meaning. Any regularities in the system are produced by structural metaphor and common framing: a structure useful in relating one event is extended to cover events with similar descriptive needs; and less frequent events are shoehorned into pre-existing frames for more frequent events. Universal rules are not needed in RCG, common framing is sufficient for meanings to be negotiated between sender and receiver. In this, RCG is ideologically closer to Systemic Functional grammar, and further from Generativism.

Croft differentiates between component and construction approaches to grammar because they analyse the syntactic and semantic representations of a sentence differently: the component approach links syntactic and semantic components via a lexicon, while the construction approach embeds semantic and syntactic features into the lexicon. The component approach is deconstructive, treating meaning, grammar and lexis separately; while the construction approach treats meaning, grammar and lexis as a coordinated description of language. Croft shows that the component approach arbitrarily attaches words to word classes, even where they have features of two or more classes; and he suggests that a flexible approach, creating new word classes as needed, better explains what is actually happening. He maintains that grammatical universals need not rely on universal "atomic" primitives, and instead proposes statistical primitives in language interrelationships (Croft, 2001, 47-61). This means the primitives proposed are complex, and their atomic features need not be the same in every language. For example, a car can be seen as a "primitive" of transport, because it is a mechanism for transferring people between places; but the components of different cars need not be similar – one could be petrol-driven, the other electric. The components themselves are only indirectly involved in the process of transport: the engine – petrol or electric –gets nobody anywhere by itself, so trying to describe the process (transport) in terms of the components (engines) of its primitives (cars) is pointless.

The idea that primitives can be complex liberates syntactic analysis. The linguist no longer identifies the basic particles of language while ignoring the functions performed; instead, function is paramount, because language relies on functions and not components. There is no deep grammatical significance if a language has no adjectival word class, unless the language cannot express adjectival qualities in any way. Similarly, if a language (like English) expresses adverbial qualities in several ways (*apply again, reapply, apply once more*) this also has no deep grammatical significance; the several ways can be attributed to new word classes if this proves useful, or dealt with as complex primitives.

With RCG, Croft views grammar as a conceptual syntactic space onto which individual language grammars are mapped; looking for universals among languages is unproductive because each language occupies only part of the syntactic space. RCG defines a language grammar in terms of exclusivities (where having one feature automatically prevents a language from having another) as well as inclusivities (or universals). Croft admits the syntactic space is enormous and he has only begun to map it; but, compared to Generative Minimalism, RCG seems likely to produce better maps of how languages actually work.

Cognitive Linguistics on the Sources of Language Grammar

Of all types of linguistics, Cognitive Linguistics has the greatest interest in language origins. Language, for Generativists, is a discrete entity generated by a single evolutionary event; intermediate states are unlikely because language only works in its final, complete state. Although some Generativists do propose routes to full language, such as protolanguage, any pre-language stages cannot have hierarchy or recursion (which would make them full language), so they have limited segmentation and differentiation (*e.g.*, Bickerton, 1990). For Functionalists, language is just one communication system on a multispecies continuum, and everything that makes language work is present in some form in pre-linguistic animals. For example, primate grooming has interpersonal significance: grooming establishes social relationships, so it is given preferentially to specific individuals, it is not indiscriminate (Lehmann *et al*., 2007). Primate warning calls have ideation: they are about specific threats and seem to signify similar things to sender and receiver (Cäsar *et al*., 2013). And Diana monkey (*Cercopithecus diana*) conditional warnings have textuality: adding a "boom" vocalisation to a warning call moderates its meaning from "threat" to "probably threat", and the receivers alter their behaviour appropriately (Zuberbühler, 2000).

In contrast, Cognitive Linguistics emphasises the evolutionary nature of language. Apparent discontinuities in language evolution occur because the evolution story is not just about language; and the presence of language precursors in other animals does not change the fact that human language needs its own evolutionary explanation.

Cognitivists currently working on the sources of language grammar use a range of methodologies. For instance, Johansson (2005, 236-240) is interested in primate social structure as a model for grammatical structure in language; baboon societies in particular have structure and hierarchy, with hierarchies of individuals within a hierarchy of families providing a limited form of recursion. Hurford (2007, 20-64) has looked at segmentation, differentiation and hierarchy in primate cognition, showing how they form grammatical stepping-stones between cognition and communicated human language. Heine & Kuteva (2007) use reverse grammaticalization to investigate early language forms; they see language developing from unsegmented noun phrases to two-segment noun-verb constructs, and then into two-argument (or propositional, or A-Relationship-B) forms, and then three-argument and increasingly complex forms, thus introducing complexity to language incrementally rather than by macromutation. Tomasello (2003a) and his team conduct extensive investigations of primates and children, showing that apes cannot understand certain aspects of Theory of Mind, such as pointing and attention, which human children grasp from an early age. Their work indicates that humans have a more sophisticated Theory of Mind than other primates. Brighton *et al*. (2005) used computer modelling to show that language itself is affected by the learnability process: it adapts to prevailing culture by becoming easier to learn. These are just a few examples of researchers using cognitive models of language to improve our knowledge of the sources of language grammar.

More recently, Schulkin & Raglan (2014) looked at the role of music as a source of language grammar. They see music, particularly group singing, as a source of both social cohesion and semantic signalling. Working together to produce music requires cooperation and coordination, and subordination of the self to the group; and the emotional content of music is evidence of meaning in the music, providing a basis on which a more conventional semantic system can build. Ettlinger *et al*. (2011) show that music, like language, has rule systems which can be learned explicitly or acquired implicitly. While nowadays we tend to notice the explicit learning and memorisation of music, this is a relatively recent phenomenon; music is essentially social, and things like rhythms, time signatures and scales are acquired subliminally as part of the local culture, just as local languages are acquired. In addition, Koelsch *et al*. (2013) show that music activates areas of the left hemisphere which are traditionally associated with language processing, implying that "processing of hierarchical structure with nested nonlocal dependencies is not just a key component of human language, but a multidomain capacity of human cognition" (15,443). Music seems to be a growing field in human communicative research.

Within linguistics, the cognitive approach currently offers the most complete analysis of the sources of language grammar. To properly understand how grammar works, the reason for its existence must be addressed: what does language do that cannot be done in another way? What features of language are specifically implicated in being human? What functions does it use, and where did they come from? And what is language for? The answers to these questions link grammar origins to general cognition, which means they are amenable to a cognitive approach. They also identify differences of emphasis between the cognitive approaches and other analytical methods: the question for the Cognitivist researcher is not, *how does grammatical language define us as humans?*, it is, *how did becoming human generate a need for grammatical language?* This is the question we explore in chapter 6.

6 Becoming Language Users

My humanity is bound up in yours, for we can only be human together.
(Desmond Tutu)

Modern humans are members of a single species, *Homo sapiens*; and, as the word *species* signifies, our species is comprised of individuals who are genetically close copies of each other. In biology the definition of species is even tighter: individuals must be similar enough to successfully interbreed. If they can interbreed, they are the same species; if they cannot, they probably are not.

Nature, however, does not always recognise the interbreeding rule, allowing closely related species, and sometimes quite distantly related species, to cross-breed. For instance, the gull species who live around the North Pole have a complex genetic relationship: the European herring gull (*Larus argentatus*) interbreeds readily with its neighbour, the American herring gull (*L. smithsonianus*), and this in turn interbreeds with its other neighbour, the East Siberian herring gull (*L. vegae*), which interbreeds with its subspecies, Birula's gull (*L. v. birulae*), which can interbreed with Heuglin's gull (*L. heuglini*). Finally, Heuglin's gull interbreeds with the lesser black-backed gull (*L. fuscus*). Yet, despite both living in Europe, European herring gulls do not interbreed with lesser black-backed gulls (Liebers *et al*., 2004).

Another example is the European white-headed duck (*Oxyura leucocephala*), which has become endangered, partly because the ruddy duck (*O. jamaicensis*) was introduced from North America. Not only do the two species interbreed (despite being geographically separated for 200,000 years), white-headed female ducks seem to mate preferentially with ruddy duck males (Henderson & Robertson, 2007). This created a major conservation issue, with ruddy ducks being subject to extermination programmes in Spain and Britain to prevent interbreeding (Muñoz-Fuentes *et al.*, 2007). The speciation of the *Larus* gulls is disputed, with between two and eight species being recognised by different authorities; and the *Oxyura* ducks, while recognised as different species, are distressingly cross-fertile. So where does a subspecies end and a new species begin?

Comparing genomes is an unreliable way to identify different species. For example, the genetic variation between humans and chimpanzees involves only a small fraction of our DNA (Marks, 2002), while the genetic variation within a single species of yeast (*Saccharomyces paradoxus*) is greater than the genetic divergence between chimpanzees and humans (Liti *et al*., 2009). It can be difficult to characterise two individuals as same or different species based on similarities and differences in their DNA; which is why the sequencing of the human genome, while a magnificent project, was only the start of our ongoing genetic exploration of what it means to be human (Hood & Rowen, 2013). It is the physical and behavioural differences in individuals, differences which are only indirectly produced by the genotype, that explain both the variation between individual humans and the uniqueness of the human species.

Nonetheless, we are beginning to unpick the history of the various species in the hominin clade. We know that, until about 50kya, there were at least four other species of *Homo* living on this planet alongside *Homo sapiens* – *Homo denisova* (unofficial designation), *Homo erectus*, *Homo floresiensis* and *Homo neanderthalensis*; and, like the white-headed and ruddy ducks, there is good genetic evidence for interbreeding between those other humans and *Homo sapiens* (*e.g.*, Prüfer *et al.*, 2021). Recent studies comparing the *Homo sapiens* genome with that of other *Homo* species indicates that our genetic admixture with Neanderthals is between 3.4% and 7.3% in populations outside of Africa, and our admixture with Denisovans is between 4% and 6% in East Asian and Melanesian populations. There is also a small unidentified admixture from another species, assumed to be *Homo erectus* (Petraglia & Groucutt, 2017).

If there was so much interbreeding, though, why do we class *Homo sapiens* as a full species? Simply because we are the last group standing. Sympatric evolution happens when two groups of the same species start evolving in different directions. Initially there may be intermediate populations which can ensure that the two groups do not fully speciate; but if the intermediate populations die out, then the two groups can become genetically distinct populations which speciate fully (Dieckmann & Doebeli, 1999). In our case, all the other *Homo* species have died out, leaving us as a full species with no co-subspecies. Being the last subspecies standing may well be more a matter of luck than evolutionary fitness; but, as E.O. Wilson said, “It's always been a great survival value for people to believe they belong to a superior tribe” (Wilson & Tyson, 2008).

So, what are the capacities that make the *Homo sapiens* species different enough from other humans to culminate in us becoming the last humans standing? We have already looked at several of these capacities in earlier chapters; but it is now time to bring them together to tell the story of becoming human.

Manual Dexterity

Bipedalism is the most obvious external physical feature that differentiates us from our closest surviving co-species, the chimpanzee. It gave us increased height, allowing us to see further; it set us on a path toward increased speed and endurance, allowing us to run faster and further; and it changed our hands from dual-purpose tools of mobility and dexterity to tools with the single role of dexterity.

Manual dexterity facilitated the development of many other capacities: tool use, tool manufacture, accurate throwing, precise gesturing, and a whole range of other useful skills. The original primary function for manual dexterity is, for our purposes, unimportant: what is important is the fact that it permits a skill-set that is peculiarly human, and which our surviving close genetic relatives can only palely imitate. However, it turns out that one important component of our dexterity is enabled not by evolution but by the lack of it: while orangutans, chimpanzees and bonobos developed hands more suited to life in the trees, the genetic lineage from *Proconsul* (the likely predecessor for all the ape clade) to modern *Homo sapiens* retained shorter fingers and a shorter palm, which meant they also retained the precision thumb-and-finger grip of the early apes (Almécija *et al*., 2015).

Manual dexterity is, however, mostly a product of cognitive control over the muscles of the hand and arm, and it requires increases in brain capacity to accommodate this control. The part of the brain involved in motor control of the hand is known as the handknob gyrus, and in most human brains it is located about 5cm above the top of the ear, which places it close to the usual language areas. There are two handknob gyri, one in each hemisphere; but, unlike the language areas, which are located in the left hemisphere for most people regardless of handedness, the two handknob gyri are less lateralised. In most right-handed people there is greater activity in the left hemisphere than in the right when performing manual tasks; but for left-handers the activity is more equally divided (Johnstone *et al.*, 2021). The handknob gyrus is, unsurprisingly, bigger in humans than in other primates, but it is easily identifiable in all primates and in many monkeys (Hopkins & Pilcher, 2001).

Frey (2008) speculates that asymmetry between the handknob gyri and their proximity to language areas may indicate that language is a by-product of dexterity, especially if pre-human signalling went through a gestural stage (Corballis, 2002). Several parallels can be drawn between dexterity and language: manual activity requires manipulation of discrete objects, just as language requires syntactic ordering of discrete morphemes; manual activity requires attention to objects and to the ways they interact, while language requires attention to objects and actions (nouns and verbs) and the meanings they generate together; manual activity requires manufactured objects to be recognised as outcomes of planned actions, while language requires meanings at the sentence and discourse level to be recognised as constructions of morphemes; and, perhaps most tellingly, language relies heavily on gesture, with a gestural phonology as informative as vocal phonology (Goldin-Meadow & Alibali, 2013).

However, the need for a specific manual gestural stage in language development remains unproven, and evidence of continuities between chimpanzee and human vocalisations continues to accumulate (*e.g.*, Watson *et al.*, 2015). It was formerly claimed that, while humans have highly expressive volitional vocal control, the volitional vocal control of chimpanzees is limited; but this is now disputed (Slocombe & Zuberbühler, 2007). Ape vocalisations are managed by brain areas that correspond to those used by humans (Ackermann *et al.*, 2014): the lower half of the face is controlled by the primary motor cortex, premotor cortex ventrolateral division and the caudal face area of the midcingulate cortex; the upper half of the face is controlled by the supplementary motor area and the anterior face area of the midcingulate cortex (Gothard, 2014); and vocalisation is controlled in the cortex by the supplementary and pre-supplementary motor areas, Broca's area, Wernicke's area, the orbital cortex, the angular gyrus, the superior temporal gyrus and the dorsal laryngeal motor cortex (Jürgens, 2002). The areas active for vocalisation are usually in the left hemisphere of the brain, although there is more lateralisation in humans than in other apes. Apes also seem to rely on learning and local cultural significance to acquire their signals, the same things humans rely on to learn languages (Marshall *et al.*, 1999); and, in bonobos at least, vocal calls are used multimodally with gesture (Pollick & de Waal, 2007). Gesture remains an expressive mode of signalling in both apes and humans; but the proposal by Corballis (2003), that vocalisation was initially subordinate to gesture in pre-humans and became predominant only with later *Homo sapiens*, remains speculative. It is more likely that both vocalisation and gesture are subservient to a cognitive need to communicate, and each is pressed into service as appropriate. After all, writing is just another example of speech by gesture, and one that has not had sufficient time to develop a genetic basis. Its widespread global adoption in the last 500 years has been possible because writing is a simple code which fairly faithfully maps sounds to symbols (Barton & Papen, 2010).

Dexterity and Working Together

As well as tools and signalling, manual dexterity affects socialisation and cooperation. First, and probably least, it enhances grooming: the finer finger control allows a range of new physical interactions. While chimpanzees tickle each other (Johnson, 2003), they do not massage each other, and caressing is not really part of their repertoire. This enhanced ability to "be nice" to each other is, in a small way, likely to have encouraged interpersonal cooperation in the *Homo* clade.

Second, manual dexterity allows complex manufacture: humans can tie things together, binding one thing to another to make a third thing. Lithic technology needs dexterity, but stone-knapping is essentially about taking bits away to leave a final object; binding is about making separate objects work together as a new object, which creates dramatic technological improvements – for instance, a spear is about ten times more effective at getting lunch than a sharp stone (Massey, 2002). Manufacture by binding may even have affected our understanding of how social groups work: just as sticks and stones can be bound together to make a spear, so individuals bind together to make hunting parties. Making sticks and stones work together is similar to getting individuals to cooperate within groups – a metaphorical exaptation from one domain to another. However, as metaphors do not leave archaeological traces, this remains an untestable speculation.

Another social function made possible by manual dexterity is shared manufacture: humans can work together to achieve a goal unachievable by individual effort. They share this capacity with other primates (Wolkenten *et al*., 2007); and, as with other primates, being able to detect and deal with cheats makes it possible (Fehr & Gächter, 2002). Yet the ventures humans undertake together are unlike any in the rest of Nature. This is partly because manual dexterity makes cooperation more reliable: others in a joint venture can be relied upon to give useful support. For instance, holding a post while someone else hits it only works if the holder knows the hitter can reliably hit the post, and not their hand. Joint ventures need not just cooperative intent, the cooperation must be effective (Gibson, 2002); and there must be trust (Wacewicz & Żywiczyński, 2018).

Notable differences between humans and other primates are the sizes of our social groups and our capacity for cooperation (Dunbar, 1993). While all apes (except, possibly orangutans) live in social hierarchies, the relative sizes of our social groups indicate that human social relationships are more extensive and more complex than for other apes; which gives us new ways for an individual to gain status.

Chimpanzees live in tribes of around 50 individuals, divided primarily by gender: the males, who are mostly related, form close alliances of two or three individuals, and looser alliances of five or six; the females, unrelated and more solitary, have a social hierarchy based mainly on age and personality. Females seldom challenge their place in the hierarchy, waiting instead for the death of rivals to move up. Chimpanzees are patriarchal (the group is directed by the males) and mostly patrilocal (the males stay in their birth territory, the females move to new territories) (Goodall, 1990, ch2), making male hierarchies more earnest than those of females. Nonetheless, the hierarchical position of a female does seem to affect their foraging success, with higher-ranked females appropriating the better feeding areas (Murray *et al.*, 2007).

Bonobos, although closely related to chimpanzees, have a different social system. There is one important distinction – bonobos are matriarchal, not patriarchal – but it changes their social organisation significantly: the unrelated females jointly intervene to suppress male aggression. This female solidarity against isolated males makes females, and their agenda, central to bonobo society (de Waal & Lanting, 1997, ch4). Chimpanzee society is governed by the reproductive imperatives of males: fertile sex is emphasised, while child-rearing is secondary. Bonobo society, based around female reproductive imperatives, emphasises child-rearing, while fertile sex, and males, are secondary. This means that recreational sex is more common among bonobos: reducing the importance of reproductive sex allows a more permissive attitude to sex in general. Where chimpanzees are only interested in sex when females are fertile, bonobos use heterosexual and homosexual sex for social grooming and bonding. Sex defuses conflicts and establishes alliances – and is even used as a greeting (de Waal, 2006a).

While humans may use sex more like bonobos than chimpanzees, bonobo social systems are more like those of chimpanzees than those of humans. Like chimpanzees, bonobos live in groups of about 50; both species are patrilocal, eat the same things, have similar lifespans, and have similar body shape. Humans, in contrast, live in complex groups numbering into millions, with each group containing iterations of groups within groups; and this group-living requires the development of specifically human strategies (David-Barrett, 2023). Our groups can be patriarchal, matriarchal or egalitarian; and territorially we can be patrilocal, matrilocal, nonlocal (where most people move away from their birth group), or even local (where few people move into or out of the group, leading to genetic inbreeding and often social collapse – Alvarez *et al.*, 2011). Our diet is omnivorous and varied, depending on locality and social grouping; our bodies – bipedal, gracile, largely hairless – are quite different from our genetically close primate cousins; and, even though menopause occurs in all three species around age fifty, our lifespan is on average twenty years longer than chimpanzees and bonobos, giving females an extended period of infertility in later years (Hawkes *et al*., 1998).

The divergence between human and *Pan* psychologies makes our social systems fundamentally different from those of our closest relatives. Chimpanzees and bonobos live in an individualistic social environment where Darwinian imperatives of genetic survival and personal promotion limit social interaction (more so for chimpanzees than bonobos: Diogo *et al*., 2017); while humans live in societies where a cultural overlay mitigates the pursuit of genetic survival. We are still subject to natural selection, and many human characteristics can be traced back to genetic imperatives: sexual selection, interpersonal competition, social ranking, the existence of schadenfreude (joy at the misfortune of others), all indicate natural selection at work in humans. Yet we also have group strategies for peacefully resolving conflicts, addressing inequities, and maintaining social cohesion; and, as individuals, we seem mostly happy to follow the codes imposed on us by our group strategies, or cultures (Currie *et al*., 2021).

Human culture is not outside of evolution, however; we evolved to cooperate because cooperation advantages individuals. Although human culture seems to contradict Spencer's *survival of the fittest* maxim (Spencer, 1898, 444), it exists because cooperative individuals got more genes into future generations than selfish individuals. We cooperate because our genes make us cooperative; and we have cooperative genes because they make humans reproductively more successful than selfish genes.

The social and cultural outcomes of manual dexterity described here all seem to involve enhanced cooperation. They are not, however, proximate causes for cooperation, or for the signalling system needed to sustain cooperation. What is advantageous to the group is not automatically selected for at the species level; the capacity must give a reproductive advantage to the individual using it. In other words, the capacity must be evolutionarily effective. This poses a dilemma for any theory of human origins, whether social, linguistic or grammatical: how can cooperative mechanisms needed for human socialisation overcome the natural advantage of uncooperative selfishness?

Cooperation is problematic because of the freerider effect (Boyd *et al.*, 2003). Any cooperative behaviour works by providing advantages to all cooperating parties: everyone puts in a little and gains a lot. Unfortunately, this is open to exploitation by cheats, those who do less than their fair share but still profit from the venture. These freeriding individuals get a larger net gain than cooperators: they increase the burden of investment made by non-cheats and reduce their gain. If cheats prosper sufficiently then, over time, they replace the cooperators and cooperation collapses.

Cooperators need a strategy to protect cooperation and disadvantage freeriding; and one possible solution is altruistic punishment (Egas & Riedl, 2005). If freeriders can be excluded from group cooperation, or punished for not doing their share, then cooperation works despite cheats. Unfortunately, the mechanisms for detecting and punishing cheats do not come cost-free. Identifying cheats requires capacities to recognise individuals and remember their reputations so that punishment can be meted out (Baumeister *et al.*, 2016) – and these costs generate a new cheating strategy: individuals who cooperate in the original venture but not in punishing freeriders have lower costs than individuals who both cooperate and altruistically punish. This selects against altruistic punishment and, when it becomes too low, freeriders go unpunished. Once again, the whole cooperative mechanism collapses (Fehr & Renninger, 2004).

One final mechanism, however, makes the whole system work. Individuals have different cooperative capacities; so, the cost they pay to cooperate, even though largely equal in absolute terms, varies with the individual's capacity. Those individuals who cooperate in both the original venture and in the punishment of freeriders are fitter, and therefore tend to be selectively chosen as mates – if the mates can identify them. Altruistic punishment is an opportunity to signal fitness: not only can I cooperate but I am also fit enough to punish those who do not cooperate. Selective mating with these fit individuals spreads the punisher's genes through the population and allows cooperation with altruistic punishment to become the default individual disposition for individuals. Fowler (2005) shows that this happens if the individual gains from cooperation are greater than from non-cooperation, and the individual cost of punishment is less than the cost paid by punished cheats.

Human cooperative culture therefore comes down to three questions: what makes cooperation more valuable than non-cooperation; what makes cheating relatively costly; and what makes punishing freeriders relatively cheap? To answer these questions we need to look at the nature of human reproduction: what is costly (and cheap) in the way we produce the next generation?

Why Human Reproduction is Weird

Human reproduction is based around child developmental features exclusive to *Homo sapiens*: our infants are born notably less capable than other Great Apes, they are dependent for longer, they have longer childhoods before sexual maturity, and there is a peculiarly long limbo state of adolescence before physical, emotional and cognitive maturity is achieved.

This extended period of helplessness is known as altriciality. Contrary to popular belief, biological puberty is quite a short event, limited to weeks or even just days; it is during adolescence that most of the biological changes caused by puberty happen, and in humans this takes years (Bogin, 2015). Altogether this amounts to a significantly greater reliance by the child on attentive parenting; and for this increased dependency on the parents to have evolved, increased adult nurturing must have already been available. From where, though, could this have come?

In most species, the costs of pregnancy and nurturing are paid by females; so, involving males in nurturing is one way for females to reduce their costs. For instance, males who provide food for the female or offspring (or both) reduce female foraging needs. In Darwinian terms, however, males willingly commit to an individual female only if she is part of their genetic future – that is, her offspring are also the male's offspring. If a male is certain enough of his paternity then he should be willing to assist his genes getting into the future; and males who support their own offspring tend to be more successful reproductively because their offspring have an extra source of nurturing (Charpentier *et al*., 2008).

There are two ways to involve males in nurturing: paternity certainty and paternity uncertainty. With paternity certainty, the female ensures the male knows it is his child by truthfully signalling fertility and only mating with that one male while fertile. This is the strategy of some gorillas (*Gorilla gorilla*), a species where both single-male and multiple male groups have been recorded (Nsubuga *et al*., 2008). When single males take harems, nurturing by the single male is spread over several females and offspring; which means it is largely limited to play and protection, with foraging support being inconsistent. However, in multimale groups quite strong relationships do form between males and immatures, mostly based on the rank of the male and not the male's kin relationship with the immature (Rosenbaum *et al*., 2015).

The alternative strategy of paternity uncertainty means females may not signal their fertility (humans do not); or, if fertility is signalled, the females mate with as many males as possible to confuse the males about paternity. If every male is possibly the father of all children in a group then it makes sense for all the males to give every child some level of support. Paternity uncertainty seems to be the strategy of chimpanzees: females show a high level of sexual activity with multiple males during their fertile periods (Pusey & Schroepfer-Walker, 2013). It does not, however, lead to better general provisioning for the females or children – mainly because male chimpanzees can identify their own offspring by other means, allowing the males to preferentially care for them (Murray *et al.*, 2016). If paternity uncertainty was originally the female strategy, it has been evolutionarily subverted by the males.

Bonobos also use paternity uncertainty, but with the important difference that a female coalition is in charge; this gives then access to a second type of shared nurturing: female-to-female. Bonobo females cooperate in both food-sharing – even with other groups (Fruth & Hohmann, 2018) – and in alliances against male aggression (Parish, 1996). When other females are present, males defer to females; but when only one female is present, males may try to assert dominance (White & Wood, 2007). In inter-group conflicts, males often cooperate to attack out-group individual males, while females are more tolerant of out-group females (Tokuyama *et al.*, 2019); in fact, bonobo females have a higher tolerance than chimpanzees for all out-group conspecifics, grooming and sharing food with them (Tan & Hare, 2013). Female bonobos also tolerate the offspring of other females, although evidence of non-kin nurturing remains uncertain. There is some evidence of females adopting non-kin infants, but it is not sufficient to treat the behaviour as species-typical (Tokuyama *et al.*, 2021).

So how do humans cooperate to meet the costs of reproduction? In the past, some anthropologists viewed the Western monogamous female-male pair-bond as species-typical for human reproduction. They argued that a human female engages the close support of a single male by offering him absolute certainty that her children are also his (Deacon, 1997). This would be a good model if the evidence supported it, but in so many ways it does not (Clutton-Brock & Isvaran, 2006). Human females do not give paternal certainty at birth because only 10% of babies are delivered on the official due date – counting back from birth to conception does not work. They also do not give paternity certainty at conception because human ovulation is concealed. Some human cultures even build paternity uncertainty into their societies by encouraging copulation with multiple males, creating the myth of partible paternity: it takes several fathers and several contributions of sperm to make one child (Beckerman & Valentine, 2002). One study of male parenthood in the pair-bonded West (Baker & Bellis, 1995), estimates the proportion of children being raised by men who are unknowingly or knowingly not their biological fathers ranges from 9% to 30%. In fact, the desire for paternity certainty seems to be quite attenuated in human males, although how this evolved remains obscure. Recently available genetic paternity tests can provide better verification of paternity; but, in a species where fatherhood is culturally, and not genetically, defined, it is reasonable to argue that, “There is no justification for performing infidelity testing on a child. Because fatherhood is not contingent on genetic relatedness, suspicion of misattributed paternity is no justification for paternity testing.” (Draper & Ives, 2009, 407).

If humans do not need paternity certainty, what creates the high level of reproductive cooperation in our species? Cooperation between related females is one solution, available to human mothers from both their own offspring and their own mothers. Humans have extended childhoods, creating a gap between the end of their dependence on parental foraging and the start of their own breeding programme, giving them time to support their mothers in raising their siblings (Sear & Mace, 2008); and human females tend to become infertile when their own children are beginning to giving birth, letting them help their daughters raise their grandchildren (Hawkes, 2004). Chimpanzees also have a pre-fertile pubescence beginning at age 8-10 (6-10 for bonobos), but their patrilocality means that females move to new tribes away from their mothers at age 11-13 for chimpanzees (just before becoming fertile) and age 6-10 for bonobos (just after becoming pubescent). For both species, first conception probably occurs soon after becoming fertile (11-13 for chimpanzees, 8-10 for bonobos), although data on this is sparse (Lee *et al*., 2020). Both chimpanzees and bonobos also have a post-fertile stage, but it is unusual for a female to survive long enough to reach it. Human females would therefore appear to have greater opportunities to exploit their female kin for child-rearing than do chimpanzees and bonobos.

Opie (2004), however, shows that these forms of female-female cooperation have limited provisioning value and can only explain a part of human reproductive success – male cooperation must have been involved, too.

To raise our human, big-brained, helpless infants there must be a high level of cooperative support available to the mother – more than can be provided by a single pair-bond. How the need for this cooperation and the cooperation itself co-evolved is addressed in chapter 8. For now, we can state the problem as: raising human offspring requires complex social cooperation which involves other fertile females, pre-fertile and post-fertile females, and males of all ages. It truly takes a village to raise a human child; but, fortunately, a village can raise many human children simultaneously.

Cooperation, Cheating, and Countering Deception

If human reproduction requires complex social cooperation then it is clearly open to cheating, especially by males. While females can cheat in terms of paternity, getting males to help raise children which are not theirs, they can usually carry only one male's offspring at a time. Until relatively recently, they have been unable to safely abandon the costs of carrying the child to birth, and the costs already incurred up to the time of birth mean they are already committed to the long post-birth care of the child. In contrast, a male can philander, getting several females pregnant but not following through with childcare; he can freeride, relying on the efforts of others to help his children survive; and he can attempt to dominate, forcing others to support his reproductive effort at the expense of their own. All these cheating strategies are used by human males, so it is not sufficient to treat cooperation as a genetic given: there must be mechanisms at work to keep most males cooperative by identifying cheats and punishing them (Riehl & Frederickson, 2016).

The first anti-cheating mechanism humans use is memory: we remember when someone has offended us and reduce our trust in them. For this to work a state of trust must exist – we cannot withdraw what has not already been given. Evolution has therefore given us a hierarchy of trust mechanisms: we tend to trust strangers only enough to not ignore them; but, over time, we build a biography of their reliability based on our own experiences and those of others we trust, and this trust expands the range of joint ventures we undertake with them; eventually, we work together on ventures where failure is quite costly. Essentially, we use memory – our own and those of trusted others – to build a personal knowledge of an individual's reputation (Molleman *et al*., 2013).

As well as using both first-hand and second-hand knowledge of a person's reliability, we also share our knowledge. For the receiver, this knowledge is both second-hand (my experience of the person's reliability, and third-hand (what others have told me about that reliability) – or, to give it a more colloquial name, gossip (Dunbar, 1996). This information-sharing has its own costs, though – and therefore creates new ways to cheat: an individual can listen without talking, increasing their knowledge without sharing; or they can lie, creating false information about the reliability of others. How do we deal with this without distrusting everything we hear, rendering knowledge-sharing pointless?

The first thing we do is interpret no utterance, no information given, in isolation. We treat most of the social information we receive as neither true nor false but as expressing a position: Alf says Gemma is friendly, Beth says she is not; what does this tell me about relationships between Alf, Beth and Gemma? And, taking into account my own relationships with Alf and Beth, what does this tell me of Gemma's likely reaction to me? Every item of social information is a two- or three-argument form: it is about the relationship between two people, and is given by one of the people or by a third person; and each item is tagged with the reputations of the protagonists, built up from previous encounters and previous social information. Scott-Phillips (2008) sees reputation as a powerful tool to keep speakers honest; but in the treatment given above, honesty is not an issue. Even downright deception, when detected, gives me useful information about the deceiver and their relationship with me. It is never worthwhile not listening.

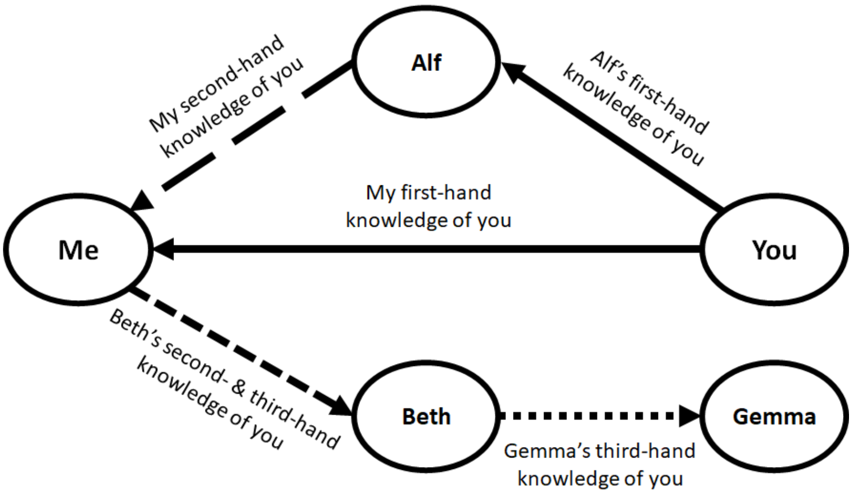

Figure 6.1: How gossip spreads reputation: my first- and second-hand knowledge becomes second- and third-hand knowledge when shared

So why is it worthwhile to share valuable social information about others as gossip? One possibility is that sharing social information is a costly signal, showing I can give away information without compromising my fitness. However, if sharing costly information is to advertise my fitness, it should be done loudly and publicly, and preferably in the presence of the other sex (Smith & Bliege Bird, 2000); this is not a good description of how we share social information. Another possibility is that giving away social information indirectly benefits my genes by giving genetic advantages to my relatives. Once again, this does not reflect how we use language: we talk to anyone, not just our kin (Beecher, 2021). A third possibility is that sharing social information is reciprocal, I give you information today in exchange for information you have already given or will give me (Dores Cruz *et al*., 2021). This seems more like the way we use language, but even this does not capture the full garrulousness of humans: gossiping about others is not just a reciprocal duty, it can be a positive joy.

Humans get pleasure out of talking and listening (Blyth, 2009), which indicates it is probably an evolutionarily fit activity: pleasure is Nature's way of telling you you're doing something which advantaged your ancestors, and which is likely to similarly advantage you. A pleasurable activity is more likely to be undertaken; and, if it helps get the individual's genes into the future, the activity is evolutionarily selected for. The near-universality among humans of "talking genes" and the relative success of the human species together indicate that talking is likely to have advantaged individuals. Our habit of talking to ourselves indicates that the advantages of talking and listening may be more than just interpersonal: hearing myself describe a relationship gives me a feel for how my description sounds to others. Self-talk lets us hear our ideas from both sides, as both speaker and listener: there are advantages to taking both sides of the conversation (Deamer, 2020); and treating others as having their own minds is what allows me to treat myself as having a mind (Edwardes, 2019).

Social information-sharing lets us coordinate our group efforts, creating cooperative ventures which achieve more than individuals can. Is this, by itself, sufficient for language to have evolved? Did physical cooperation and communicative coordination co-evolve, each feeding the other in a ratchet effect? We have certainly seen this process at work recently: cooperation in shared ventures of design and manufacture produced computers, which generated new ways of exchanging information, which produced the Web, which generated new ways of exchanging information, which produced the Internet, which generated new ways of exchanging information, which produced the blogosphere, which produced Facebook and Twitter … which broke the virtuous cycle. All good things come to an end eventually.

Yet, even though we are now in a "post-truth" phase, we still seek ways to cooperate. Gossip works as it always has: it is information, not knowledge, opinion heavily tainted by its source. It is never worthwhile not listening, but it is also never worthwhile to believe without scrutiny. So, despite the viruses, worms, trojans and other malware, and despite the conspiracists, Karens and Terrys, trolls, grifters and scoundrels, cooperation continues to grow through enhanced connectedness.

The Importance of Models and Model-making

Social information sharing involves knowledge from the past, and references people not usually present (Ganea, 2005); both sender and receiver must tokenise the shared representations of others and manipulate those tokens cognitively – essentially, we must make mental models of other individuals, events and relationships. Humans are gifted model-makers: we create entire worlds in our heads and share them with each other; and we do so from a very young age (Gredebäck *et al.*, 2018). Not only do we model things which are no longer there, we model things which are not yet there, and things which never will be there. Indeed, it can be argued that the whole of human culture is based upon a counter-reality where representation – the model – is more significant than the actuality (Knight, 1999). For instance, in the West we often count value in units of money, a fictitious and notoriously variable measure; but money lets us compare otherwise-incomparable relative values, and so has a significant effect on our views of the world and of other people.

Making models informs more than our social understanding, it is also involved in our toolmaking: envisioning the end-product tool and the steps needed to get there require modelling of the same nature as language modelling (Brozzoli *et al.*, 2019). We do, however, need to differentiate modelling from planning. For example, if we consider chimpanzees foraging then we can see that they start from a simple state of being hungry. Hunger is endemic in nature, unremarkable, and requires no modelling; whereas imagining a juicy pear in your mouth to satisfy that hunger is modelling. If, however, a chimpanzee decides to move to a place where it remembers recently seeing unripe fruit now likely to be ripe, that is planning; and there is evidence that chimpanzees can do this (Janmaat *et al.*, 2014). Planning embodies a purpose, and there is no need for the end-result of that purpose to be cognitively modelled: hunger is satisfied by foraging based on planning, there is no need for the outcome, eating, to be modelled. Chimpanzees use their local knowledge to maximise their foraging effort, and so meet their generalised needs; but humans do not just forage, we shop, both modelling and planning a future in which our wants have been satisfied in specific ways by specific items. What matters for humans is not just satisfaction of sensations, it is the actualisation of models, seizing control over the future. We are interested not just in having but in owning.

Model-making seems to be a type of cognition at which humans are particularly adept (Sloman, 2008). It is involved in complex toolmaking; in the mechanisms of human culture (Premo & Kuhn, 2010); and in the social calculus we use to navigate the complexities of our social groups (Fowler *et al.*, 2009). It uses the same system we find in language grammar, placing objects in relationships with each other. A mental model of a social event requires segmentation, to identify the individuals in a model; differentiation, to identify which individuals are doing and which being done unto; and hierarchy, both of time – to identify what happens first and last – and of space – to identify what happens closest and furthest. The mechanisms involved in mentally modelling social events are the same mechanisms we need to explain grammar.

7 Modelling Society and Sharing Models

Man is least himself when he talks in his own person.
Give him a mask, and he will tell you the truth.
(Oscar Wilde)

Humans have evolved to socialise; which means that humans need to cooperate extensively; which means that, if they are not to be exploited by others, humans need to maintain a complex social calculus of the society around them. Cooperation also means that humans need a complex communication system to share their complex social calculus, which means that the communication system is likely to faithfully reflect the complexities of the social calculus. Where Chomsky takes the position that language evolved for thinking (Chomsky, 2000, 27-29), the viewpoint taken here is that language developed to share thinking – specifically, social thinking. The structure of language did not dictate the structure of cognition; instead, the structure of social calculus generated the structure of language.

Social calculus is just a shorthand term for the modelling that humans undertake to navigate their social environment. It involves modelling relationships with others and between others, building a network of nodes (other individuals) and the links (relationships) between them. This modelling can be both subliminal (without awareness of the computation) or intentional (using deliberate computation); and humans seem to be particularly adept at intentional modelling. Jordan (2023) describes two types of intentional modelling used by humans in their social calculus: individuals cooperate to maintain their reputation for cooperativeness in the group; or Alf cooperates with Beth because they have seen Beth cooperating with others.

Modern human socialisation relies on models of the actual world endorsed and enforced by the group. Because of the models, we live in a reality of money, manners, rituals and conventions abstracted from the actualities of food, sex and survival. We also create shared virtualities (abstract concepts, art, nations, and suchlike), which become realities if accepted as such by enough people. Popper (1967 [1985], ch4) described these as World 1 (actuality), World 2 (virtuality) and World 3 (reality)[7], identifying World 3, reality, as the key to human cognition. No other animal on this planet has demonstrated the capacity to live in and share these three worlds simultaneously.

We often overestimate the authenticity of virtuality and reality, assuming beliefs to be actualities and therefore The Truth; and that alternative interpretations cannot be true. Yet these Truths are not actualities, and cannot therefore stand against scrutiny based on actualities. It is often people peddling a Truth who introduce ideas into the group reality which are unsupported by actuality. Galileo Galilei was tried by the Catholic Inquisition for the heresy of suggesting that the Earth orbits the Sun, which means it is not the fixed centre of the Universe. Forced to recant, he is reputed to have said, after the trial, “eppur si muove” – and yet it moves. The Catholic Church was able to insist on the “evidence” that the Earth does not feel like it is moving, and the Sun can be seen to move across the sky – evidence that Flat-Earthers insist on even today. Yet most of us – including the modern Catholic Church – now accept that the actualities of physics provide a better explanation than the realities of perceived evidence.

However, treating Actuality as the final arbiter of truth is not always useful. We do not use actuality to arbitrate right and wrong in everyday life; instead, we rely on our group morality, a set of shared beliefs about acceptable behaviour (Hoffmann *et al.*, 2014). Morality is often seen by the individual as universal, but it seldom is; it is based on locally agreed realities of behaviour, and often what is considered moral in one culture can be viewed as immoral in another. It is also partible: the morality of intragroup

[7] Popper describes them as Worlds 1, 2 and 3; Actuality, Virtuality and Reality are my terms.

relationships – how to behave with other members of the group – is often different from the morality of intergroup relationships, as almost any conflict between groups illustrates. Tomasello (2018) describes a three-stage process in the development of morality: the first stage is individual intentionality, the Machiavellian me-first attitude we see in chimpanzees; next is joint intentionality, where individual success becomes reliant on success within a group – a we-before-me strategy probably used by early humans; and the third stage is collective intentionality, where common practices establish a shared culture with a shared objective morality.

A shared culture and a shared morality imply a shared convention about what is right or wrong, which can act as a gauge for authenticity in communication. Human language can share virtual and real thoughts as well as actual facts, which means it is capable of expressing both truths and nontruths; but this raises the question of why a system apparently designed for communicating opinion, beliefs, negotiable meaning, what-ifs, fiction and downright deception should be so central to being human. It may be, as Marcus Aurelius wrote, "The universe is transformation: life is opinion"; or, as Nietzsche put it, "All things are subject to interpretation; whichever interpretation prevails at a given time is a function of power and not truth".

Goodman (2022) sees lying as central to the way humans communicate, not an exotic outlier use of language. He says that, "the history of language is also one of subtle lies, not clear truths. Recognising that our communication is a mix of such evolutionary influences can help us better understand our origins and broach big problems of our time, discerning truth from falsity and honesty from disinformation."

So the problem of listening to language is not the objective truthfulness of the utterances, it is the speaker's subjective intention to deceive. To be effective, a lie should not be identifiable as such – or, to put it another way, a good lie should be indistinguishable from the truth. The semantics of the lie must be believable, the structure of the lie must be the same as the structure of a true statement, and the phonology used to lie should not be distinctive. Detecting lies is not easy (Persaud, 2005); in fact, we usually detect dishonest utterances by "errors" in production; and these errors are seldom in the utterance itself, they are in the contextual emotional and physical presentation of the lie. This is why aphasics (people who have lost their language) detect lies better than the language-capable (Etcoff *et al.*, 2000). There are few lexicogrammatical forms in human languages which reliably point to dishonest intent or indicate that an utterance is a lie. In this, lying contrasts with storytelling, where the indicators of virtuality are easy to find (Benson, 1993). For instance, the classic structure of a fairy tale is "Once upon a time … and they all lived happily ever after." These two utterances bracket the story, telling the listener that the utterance is a recounting of virtuality (the conditional world of "what if…") and not reality (the world of accepted truths and facts).

As humans, we share our thoughts about reality and virtuality much more than our thoughts about actuality, and this is what makes the communication game so different for us. Truth is no longer *what aligns with the contextual facts*, it is *what aligns with my interpretation of the contextual facts*, and this can be different from your interpretation of the contextual facts. Tarski (1936, 153) says, "With respect to this [colloquial] language, not only does the definition of truth seem to be impossible, but even the consistent use of this concept in conformity with the laws of logic" – in other words, everyday language is not designed for exchanging truths.

Even where a particular interpretation of the contextual facts is counter-indicated by actuality, it does not necessarily alter the interpretation: actuality is always interpreted by humans through their reality (customs and practices) – and sometimes through their virtuality (beliefs). The dismissal of Samuel Johnson's "I refute it thus"[8] is an example of the triumph of virtuality over actuality: Johnson's appeal to the actuality of physical pain is not sufficient to overcome the sophisticated virtuality of Bishop Berkeley's theory of immateriality, and Johnson's physical argument is dismissed as the

[8] From James Boswell's The life of Samuel Johnson, 292: After we came out of the church, we stood talking for some time together of Bishop Berkeley's ingenious sophistry to prove the non-existence of matter, and that every thing in the universe is merely ideal. I observed, that though we are satisfied his doctrine is not true, it is impossible to refute it. I never shall forget the alacrity with which Johnson answered, striking his foot with mighty force against a large stone, till he rebounded from it, "I refute it *thus*".

appeal to the stone fallacy. Despite the necessary actuality of actuality, the virtual thought experiment, that non-actuality could be actual, trumps actuality. The truly weird thing is that experimental physics currently supports Berkeley's position, showing that what we think of as actuality only exists while it is being observed or measured (Manning *et al.*, 2015).

It is in Popper's worlds 2 and 3, Virtuality and Reality, that the mechanisms necessary for social modelling work. My capacity to model others as entities comes from my ability to treat virtual speculations about others as if they were actual knowledge; and my capacity to model those modelled entities in relationships with each other and with me comes from my ability to treat my real understanding of their relationships as if it were actual existential knowledge about my group. Virtuality and reality allow us to construct the myth of being human, which in turn allows us to explore our models of becoming human.

The Structure of Social Modelling

To understand human socio-cognitive modelling, we must identify how humans encode and manipulate social structures to navigate our social landscape. The first thing needed is recognition of individuals; and the second is flagging those individuals with the emotional relationships between us. This individual-plus-relationship system differentiates between objects and actions, mapping the noun-and-verb distinction and the one-argument grammar form (Relationship-A). Individual-plus-relationship social modelling reliably generates and expresses appropriate social reactions to the individuals modelled – reactions such as submission, dominance, irritability, tolerance, and so on – which lets us concentrate our efforts on relationships likely to lead to alliances. It is, however, a basic modelling capacity and not limited to humans. Hurford (2003) shows that the argument-predicate (action-object) distinction is neurologically based, available to apes and monkeys as well as humans.

For social calculus, the one-argument form needs support from the two-argument, or A-Relationship-B, form. This models two objects or individuals together into a single event involving them both. For instance, understanding that *Alf likes Beth* needs a mind able to simultaneously model two individuals and the relationship of liking between them. This more complex form of cognition seems less widespread in nature than the one-argument form. Some human-acculturated chimpanzees do seem able to use two-argument social calculus (Tomasello & Call, 1997, ch10); so, while more limited than one-argument forms, it is not exclusively human. However, where humans use two-argument social calculus to understand how the individuals model each other, chimpanzees seem to be aware only that there is a relationship: "Although we should not prejudge the issue given the small amount of systematic evidence available, at this point there is little convincing evidence that any nonhuman primate species understands the knowledge states of either humans or their conspecifics in the way that human children do from four years of age." (*ibid*., 330).

This chapter looks at four interrelated types of cognitive modelling used by humans to navigate relationships with others. First is modelling the intentions of others, understanding others as beings with agendas and not just objects to be manipulated. Second is modelling past and future events as current, allowing events to be compared and collated outside of time, and individual reputations to be constructed; feelings produced by past events and outcomes are remembered and projected onto current situations, but from a disengaged, third-person viewpoint of the self inside the modelled event. Third is modelling *what-ifs*: as well as future probabilities and past certainties, we can analyse events that did not or are unlikely to happen, while still knowing the difference between *what-ifs* and actual events. And fourth, and most significant, is modelling our own self.

Selfhood has long been recognised as vital for understanding what makes us human. Even before Darwin's theory of evolution by descent, Humboldt (1836 [1999], 4) stated that "… the appearance of a greater individuality in individuals and in peoples, practically inexplicable by any derivation, interferes suddenly and without warning with the course more obviously determined by cause and effect." Even earlier, Descartes (1641 [1998], 79-86) proposed a duality of the physical automaton self and an intangible spiritual controlling self: in his theory, selfhood is a product of spiritual, not physical, cognition. Freud (1923 [1961]) went further than Descartes, finding three types of self at work in human cognition: the id, concerned with immediate satisfaction of wants (subconscious, affective, physical, self-directed); the super-ego, concerned with rights, wrongs, and keeping on the right side of others (moral, intuitive, cognitive, other-directed); and the ego, trying to find a balance between the id and the super-ego (rational, logical, social, self-directed). I myself have identified seven types of self (Edwardes, 2019, 163-180), and some other views of selfhood have even more.

Lakoff and Johnson (1999, ch13) reverse Descartes' spiritual-over-physical dichotomy. They distinguish between the subject, a physical entity largely unconscious of itself, and the self or selves, the cognitive models used to evaluate the subject's physicality. The physical subject is the essential person in their description, not the cognitive constructions. The self-modelling described here tends toward the Lakoff and Johnson approach, so Descartes' famous dictum on the duality of self should be reconstructed as: *I (the subjective self) think that I (the modelled self) am, therefore I (the objective self) am*.

How to Plan and Make Models

The difference drawn in chapter 6 between planning and modelling is crucial to understanding the distinctiveness of human cognition. To summarise, planning is devising a way to achieve a particular end, but the end itself need not be envisaged. For instance, deciding to rest at the end of a hard day does not need us to imagine ourselves asleep in bed, our bodies resting, our brains re-organising. A sensation of tiredness starts a series of pre-defined activities, ending with us in bed, unconscious. In contrast, deciding to buy a bed does require us to imagine ourselves asleep; and, when we test beds, it involves comparing an idealised unconscious state with several consciously assessed examples. We cannot know ourselves asleep, but we can model ourselves asleep, and our plan (to buy a bed) can lead to that cognitively modelled endpoint. Planning does not require a modelled endpoint: we do not have to imagine ourselves asleep to decide to go to bed.

Planning is not limited to humans. Despite being more Machiavellian than us, chimpanzees and bonobos build long-term relationships with each other, and influence each other's actions. Like us, they seem to have subtle mental constructs of their reactions toward others which they use to navigate their social group. They would understand, for instance, that if Gemma fears Alf but has a good relationship with Beth, and Beth has a fair relationship with Alf, then Gemma should stay close to Beth. Alf is unlikely to attack her because Beth is likely to support her against Alf, and Alf therefore risks unfavourable odds. Being alone with Alf is stressful, so Gemma will avoid it. Being alone with Beth is pleasant, so Gemma will seek it. Being in the presence of both is less stressful than being in the presence of Alf alone, so if Alf is getting too close then Gemma will gravitate toward Beth. There is planning here, but no need for conscious modelling. Gemma does not need to imagine the outcome of being alone with Alf; the stress created by imagining is likely to resemble the stress created by the actual event, so why do it?

Can other animals make cognitive models? There are some studies of wild primates which seem to show them modelling outcomes and planning toward them; but most examples can be explained equally well by operant conditioning – pursuing rewarding activities and avoiding unrewarding activities; wild primates do not need deliberate modelling or even Theory of Mind (Byrne, 1995, ch9). However, there remain some isolated primate examples which may involve conscious evaluation of means to achieve envisioned ends (Tomasello & Call, 1997, ch7). There is also extensive evidence that, when trained in a human environment, chimpanzees (Premack & Premack, 1983), bonobos (Segerdahl *et al*., 2005) and gorillas (Tanner *et al*., 2006) can produce behaviours that seem to involve modelled outcomes. Pepperberg (2021) has identified similar capacities in grey parrots (*Psittacus erithacus*) exposed to a human environment and training; and Osuna-Mascaró *et al.* (2023) have shown that Goffin's cockatoos can select the tools they need for a task, take them to the task, and use them in the correct order to complete the task – they seem to be modelling the outcome of the task to plan their actions. So, while we cannot discount nonhuman modelling, the distance between human socio-cultural achievements and those of other primates indicates that human modelling is more frequent, more complex and more effective.

Using the Alf-Beth-Gemma example with humans, Gemma can model the three-way relationship between Alf, Beth and herself with two one-argument models of her relations with Alf and with Beth, and her two-argument model of the relationship between Alf and Beth. Both forms of modelling involve second order intentionality, or *aboutness* (Arbib, 2021), but of different types: the first models your thoughts about me, and the second models their thoughts about each other (Dunbar, 2004, 45-51). In the first, your thoughts and intentions are about me, so directly affect me; in the second, their thoughts do not involve me, so only indirectly affect me through my own intentions toward those others.

As humans, we are each aware of our own and others' individuality. I am aware not just of myself (my first-person model of me), and not just that you can model me (your second-person model of me), I am aware that others can have third-person models of the relationship between you and me; and, most importantly, I am aware that the viewpoints of those others can be incorporated into my first-person model of me and your second-person model of me. I must simultaneously be aware of my own self-image, and of the images of me maintained by you and others. I must also be aware of my reputation – the group-image, or third-person viewpoint, of me built by people who have only indirect knowledge of me from the models of me shared with them by others. *Homo sapiens* may be the only species that does this multi-level modelling (Edwardes 2019, 95-119).

Modelling and Sharing the Self

The first question to ask about the self in language is simply, what is a self? Or, more directly, what is me? This is not the same as the ubiquitous question, *who am I*, which mostly aims to identify an intimate self by introspection; instead, *what is me* describes an externalised model of the self – not a subjective self-interested self but an objective disinterested third-person self. Jaynes (1990, 59-65) describes the objective self as the metaphor *me*, in contrast to the analog *I*; while Greenfield answers the question *what is me* by referencing an always-present but subliminal self; we become aware of that self through linguistic self-recognition:

> … language gives us a symbol for something that normally does not make inroads into our senses, simply because it is always there: one's self. As soon as we have a simple word for ourselves then we can inter-relate the self in context. We can become self-conscious. This self-consciousness, combined with the ability to escape from the here and now, is surely what really distinguishes us from almost all other animals, as well as seemingly inhuman human infants. (Greenfield, 2000, 169.)

Greenfield formulated the self in context (the metaphor *me*) as a conscious reflection, or model, of the true physical self (the analog *I*). We can describe our modelled self because it is produced by conscious cognition; but we can only imperfectly describe our actual self – others can see and describe our actual selves more accurately than we can. Tomasello (1999, 70-77) sees this process the other way around: during childhood, our increasing self-awareness creates our knowledge about the capacities of others. This inversion of cause and effect implies, however, that we should know ourself better than we know others, which does not seem to be the case. Dunning *et al*. (2004) show that the correlation between our view of our intelligence and our actual intelligence is low, that we always take more time to complete tasks than we estimate, and that we are incurable optimists about our own health. Baumeister *et al*. (2005) show that our self-esteem does not match the esteem others give us, and it has a low correlation with our academic achievement. The consensus seems to be that the self we recognise in language (*I*, *me*, *myself*) does not map well to our actual self. Benjamin Franklin said, "there are three things extremely hard: steel, a diamond, and to know one's self"; the evidence indicates this is a reasonable view.

We understand others better than we understand ourself because most of our self-knowledge comes not from introspection but from modelling the minds and intentions of others. Introspection gives us a biased view of our external self – what we want to present to others rather than what we actually present – while our internal, or actual, self is opaque to us. We discover our external self by learning what others think about us, which requires attending to two types of unreliable information: first is what the speaker tells us they think about us, which is subject to flattery and sycophancy; and second is what the speaker tells us about a third party's view of us, which is subject to the speaker's biased perspective.

I can only build a model of my self as others see me by interpolating between the models of me offered by those others; and communicating those models requires both speaker and listener to use a system allowing faithful transmission of social calculus models between minds – a system which, if not human language, must be something similar.

From Selfishness to Awareness of Self

Cognition is costly: the human brain uses a massive 20% of the body's energy intake, a cost justifiable in evolutionary terms only by the value of what it does (Raichle & Gusnard, 2002). Any cognitive process must be viewed in terms of its likely energy cost, and the continued existence of a cognitive capacity in a species shows a clear fitness advantage for paying that cost. The cognitive processes of understanding self and others must bring fitness advantages to the individual; but, in considering those advantages, we must also remember their cost.

Awareness of self, as the word *awareness* implies, is a conscious cognitive act for humans: we know we have a self and can construct a model of it. Awareness of self is therefore not the same as subliminal self-interest, or selfishness, which underpins the evolutionary process (Dawkins, 1989, ch5): where genetic selfishness is concerned with the needs of the genotype, or species, awareness of self is concerned with phenotypic, or individual, needs. In the phenotype, genetic activities are expressed as autonomic responses to environmental stimuli: genes do not choose to be selfish, they are selfish because only selfish genes survive – genetically inspired actions favouring the phenotype and its reproduction lead directly to genotypic survival, while actions that favour others do not. There may be indirect fitness gains in favouring others; but, if this altruism is to become genetically established in the species, it must give the individual sufficient advantage, not just to outweigh non-cooperation, but to outweigh every intermediate strategy between non-cooperation and the altruistic trait. There is no intentionality in Dawkins' selfish gene model, any more than there is intentionality in a tornado: both are natural phenomena controlled by explicable rules external to the phenomenon.

Genetic selfishness is different from sense of self, which in turn is different from awareness of self. Genetic selfishness is a default state which does not require consciousness. It ensures survival and requires no more knowledge of the self than that the world is divided into self and not-self. What is inside the line, the self, is the ends; the rest, the other, is just means; and, in the world of genetic selfishness, what favours the self is good, while what favours another is not good. With a binary model such as this, only one of the pair of self and not-self needs to be identifiable; but, while the self seems the easiest to know, but it is also the least useful. The self is self-controlling, therefore trivial; more important is that rest of the universe, which must be negotiated and manipulated. A feature of genetic selfishness is therefore likely to be a lack of comprehension of the self. *Sense of other* gives immediate advantages, it allows an organism to subvert the survival of those others to its own purposes; *sense of self* gives no such immediate advantages (Edwardes, 2019, 29-51).

Yet, although sense of self gives no direct advantage to an individual, it does give indirect advantage: it lets the organism exercise choice between strategies. In most situations, an organism has more than one viable strategy, and the ability to choose the best maximises the organism's fitness. Adopting a single strategy for a recurring event relies on there being no effective response at the other end of the strategy; because, when an inflexible strategy is met with a variable response (which an effective choice allows), it ceases to work reliably, and can even become counter-productive. To make choices an organism must understand that there is a self to make choices – they must have a sense of self. This need not be conscious, it can simply be a recognition at the genetic level that the other half of the binary division into non-self and self exists.

If I have choices, however, it becomes advantageous for me to model those choices onto other organisms. If my species is successful because it can make choices then my immediate same-species rivals will have that ability, too; and consciously modelling their counter-strategy reactions before making my own strategy choice is the next logical step in gaining a fitness advantage. The choices I make for myself do not require conscious attention, I can use “gut feeling” to choose my best strategy affectively; but, with no access to your gut feelings I cannot treat your choices affectively. Instead, I must consciously model your choices and their outcomes to anticipate them effectively. It becomes advantageous to develop *awareness of other*: anticipating how the choices of others can affect my choices.

This gives us two types of knowledge: my own choices are subliminal, they are dictated by my intuitive reactions; in contrast, my analysis of your choices must be a conscious act – I must be aware of your choices to cognitively evaluate them. Analysing your choices lets me evaluate your intentions: which choice will you make? But evaluating my own intentions is unnecessary – they emerge automatically from my feelings about your intentions without conscious analysis.

With awareness of other, you become a modelled entity in my consciousness; but I remain unmodelled and not part of my conscious cognition. My model of you includes the fact that you have intentions: I see you, to use Tomasello’s definition (1999, 176-180), as a mental agent. My model does not, however, include conscious empathy for you: not having a model of me, I cannot put myself “in your shoes”. Awareness of other allows me to generate increasingly sophisticated models of your intentions, but it provides only a limited toolbox to manage those modelled intentions. It provides improved Machiavellian intelligence, but not the empathic social cooperation of human culture.

Modelling the choices and responses of other organisms is constrained by what I know about my own affective choices, so increases in the number of my subliminal choices lead to more sophisticated conscious models of others. However, choice itself is an evolutionary capacity and has its own cost-benefit calculation. Having too many choices can become bewildering, lead to dissatisfaction and regret for choices not made, and dysphoria over the need to choose (Schwartz, 2004). Keeping my own choices subliminal must help to reduce this *opportunity cost* unhappiness.

While my social calculus remains inside my head, I can treat all modelled entities as third persons – *they* and not *you*. However, the sharing of social calculus changes this drastically: I become able to offer you social relationship models which include you; which means that you become able to offer me social relationship models which include me. Everyone is faced with the problem of how to build a model of themself into their social calculus. The easiest solution is to create a third-person placeholder which is the same as all the other nodes in my calculus. It is not quite my model of me, it is instead my model of the possibility of me.

Sharing social calculus models gave a new type of self to be recognised. Hofstadter (2007) identifies it as a *self-symbol*; Tomasello & Call (1997, 337-338) describe it as *self as social agent*; Pinker (1997, 134-136) names it *self-knowledge*; and Bruner (1986, ch4) calls it *the transactional self*. Here, the stage is classified as *awareness of self*. We extrapolate from our Machiavellian models of others to making models of ourselves, conceptualising ourselves as if we are looking in from the outside. Our model of ourself is often inaccurate, but the ability to generate it at all is an evolutionary conundrum: how have we become able to take a disinterested third-person view of ourselves?

From Awareness of Self to Awareness of Selfness

Awareness of self is a byproduct of information-sharing: exclusive modelling of the intentional behaviours of others, *awareness of other*, becomes inclusive modelling of the intentions of both others and myself. My model of my self becomes identified with my models of others so completely that their roles in cognition and language become interchangeable. I can see myself simultaneously in three ways: as the actor undertaking an action, as the patient receiving an action, and as part of the context in which an action is performed – in an utterance I can occupy first argument, second argument, and third argument roles, the grammatical roles of subject, object, and indirect object. I can also see you in any of the three roles, and both of us can be replaced by third parties: yesterday's *you* becomes today's *they*, which I can tell another *you* about.

Viewpoints generated by self-modelling are ubiquitous in language, and they can be identified in specific English usages. For instance, the apparently synonymous reflexive forms, *I hate me* and *I hate myself*, seem to have different roles in identifying the nature of self (Edwardes, 2003), with *me* representing Jaynes' (1990) analog *I* (and therefore co-identifying with the *I* in the sentence) and *myself* representing the metaphor *me* (an entity different from the *I*). This self-modelling role-taking is also recognisable in idioms like *if I were you* and *from my point of view*.

The self of which we are aware is not the same as the affective self we sense when we experience pleasure or pain, repletion or hunger, stress or satisfaction. My affective self is closely aligned with my actual self; and this is cognitively vast and subliminal. There is a common belief that a person is conscious of about 5% of what their brain is thinking at any time, while the rest remains beyond conscious access (*e.g.*, Solms, 2017). This 5% figure is a ready reckoner rather than an established scientific fact, but it is reasonable to say that we are rarely aware of more than 10% of our cognition and, in periods of deep sleep, we can be completely unaware of the world either inside or outside our head. Dehaene (2014, 64-65), however, shows that just because we are unaware of inputs does not mean that we have not processed them. He calls this the cocktail party effect, describing how we can follow a single speaker in the hubbub of a cocktail party, appearing to ignore every other speaker; but when we hear another speaker say something triggering, such as our name, we can immediately switch attention – or, if asleep, wake up. Raichle & Gusnard (2002) propose a more assessable measure of consciousness by looking at the energy budget of the brain: at any time, at least 80% of the brain's energy is being used for baseline cognition – what is needed to keep alive while under anaesthesia. We are unable to access this baseline cognition consciously, but its outcomes can force their way into awareness, often as an experience of the affective self: I feel hungry, or tired, or happy…

When I consciously create a cognitive model of another person, I am calling upon three things: my memories of my experience of that person, my memories of what others have said about the person, and my memories of my previous models of that person. These memories must have been consciously processed because they have been assigned meaning; however, the processes of laying down and recalling memories is handled by unconscious baseline cognition, so I have little control over the biases involved in what I remember and forget. All that can be said is that my unconscious actual self can be relied on to provide my conscious model-making self with sufficient information to make the decisions needed to model that other person. It may feel like *I* am consciously gathering and assessing the information to make my modelling decisions, but a large part of the gathering and assessment has already been done subconsciously before my conscious cognition is engaged (Morsella *et al*., 2016).

When I consciously create a cognitive model of my self, I do exactly the same as for a cognitive model of another person – except that my memories of my experience of myself are mostly subliminal and affective. The modelled self is heavily tainted with subliminal feelings about my self, and awareness of self is therefore not fully disinterested. However, it is disinterested enough for me to project my self-model into the past and future, and into speculative scenarios. My modelled self is delimited by my embodied actual self, so it feels more personal than my models of others, even though it is probably less true to my actual self than my models of others are to those others. Awareness of self relies on the dynamic relationship between the self-as-other and the actual self; but it is actually awareness of a model.

For Bloom, (2002, ch3), awareness of self is an important feature not just of being human but of language learning itself: children do not learn words by a process of association, they learn them by inference from the intended meanings of others. This means that children, when they begin to utter their first associative words between ages three and four, already have sufficient modelling ability to understand that the word-sign is a negotiation between them and other people. They also understand enough about intentionality to know that the meaning of a communicated word-sign is in the intention of the speaker (or sender), and it is the role of the listener (or receiver) to try to apprehend that meaning. The receiver must know that others have intentions, a capacity we call Theory of Mind.

Children can attain a Theory of Mind because they are born with a theory of theory. They seem to implicitly understand the process of thesis-antithesis-synthesis which is central to human scientific method. They apprehend the world, make models of it, check those models against new realities as they arise, and modify their models appropriately. Gopnik *et al*. (1999, 155-162) call this “the scientist as child”, showing how childhood modelling builds adult competence, establishing the human ability to continue modelling into adult life. Humans continue to play in the “mental sandpit” of modelling throughout their lives.

To be human means being able to use second and higher orders of intentionality: it means being able to see others as mental agents, with their own cognitive life; it means being aware of my cognitive life as a metacognitive event – being able to think about my thinking; it means being aware that there is a *me* to be thought about, to be planned for, and to have unrealistic expectations about, and a modelled *I* to do this; and it means having the ability to create a virtual world inside my head which is as significant for me as the actual world outside my head.

Socialisation is an important component of awareness of self. Malik (2000) argues that self-awareness is intimately tied to both language and social living. Unless we can understand that others have intentionality, and what that means for them, we cannot begin to understand our own intentionality: “The existence of a community of beings possessing language allows us to make sense of our inner world, and hence to become self-conscious. At the same time, I am only conscious of myself insofar as I am a member of such a community” (Malik, 2000, 220). Or, as I express it elsewhere, *Cogitant ut sum, ergo sum* – they think I am, therefore I am (Edwardes, 2019, 7).

The socialised nature of intentionality makes other things possible. The first is *Anticipation*, or a second-guessing between intentions: I know both your options, and my options, so I should choose the option that gives me the best result in response to your best choice; but then you know my options and you know your options, so you may choose the option that gives you the best result in anticipation of my best response to your best choice, so I should choose the best response to that option… There is a recursion between your intentions and mine within both of our minds, and this recursion is one of the capacities that enable language, as Hauser *et al*. (2002) argue – although the recursion of intentionality described here is a product of pre-existing social modelling, a cognitive capacity which precedes and is not limited to language. It is not the same as the FLN (Faculty of Language Narrow) proposed by Hauser *et al*.

The second capacity made possible by awareness of self is *Speculation* on the intentions of others toward each other, which requires no reference to the self's own intentions. Other individuals are not just animate agents linked by relationships, they are mental agents with their own intentions. Modelling the intentions of others is not done to identify strategies which are directly useful to me, but simply to identify, from a third-person viewpoint, what is going on. It is this capacity to take a third-person disinterested view that enables and informs the insatiable curiosity of humans (Stewart & Cohen, 1997, 163-164).

The recursion involved in awareness of self also permits the grammatical tool of *Reflexion*: the analog *I* is different from the metaphor *me*, but interchangeable with it. The intentions of others can be modelled into both the actor and the receiver of the action, which means they can be modelled as both actors and receivers, and the same individual can occupy either role in a modelled event. Reflexion occurs when the same individual occupies both roles simultaneously: they are performing the action of the utterance on themself, as in *I like me/myself* (Edwardes, 2003). Reflexion works in different ways in different languages, but the need for a reflexive construct which allows the same individual to be both actor and receiver of the action seems to be common throughout human languages, possibly universal (Lehmann, 2015, 45-52).

Self-modelling also raises the issue of *Temporality*: humans see themselves as continuous with their past and future selves, but they are also able to see those past and future selves as if they were other people. It is unlikely that non-linguistic animals do this: they probably have a subliminal sense of their own continuity – survival is its own testament to continuity – but they do not have a consciously modelled intentional self, just a feeling that they have existed and will exist beyond the current moment. This non-linguistic sense of continuity is a trivial thing which only serves to subliminally inform the actual self, not model it. Nonhumans can see time as an ongoing process, but they are unlikely to be able to model their own self passing through time. We look at the human trick of seeing time as episodic through the eyes of modelled past and future selves in chapter 12.

All these features – anticipation, speculation, reflexion and temporality – rely on my capacity to model myself. This is an unusual talent, and problematic in Darwinian terms. To make models of myself I must step back from myself: I must attempt to view the "real" me from a third-person viewpoint. This means I must be disinterested about myself, to try to see myself as others see me; and this is a skill at which we are far from practiced. Our self-models are almost always wrong in material ways: we overestimate ourselves, deluding ourselves about our abilities. we need to be factually honest to our self about one person in our universe, and that is ourself; but, in large part, we cannot do it.

Awareness of Selfness

There is a linguistic paradox, the Liar's Paradox, which has exercised philosophers for millennia. It consists of three words which iterate between truth and falsehood, being one, both, and neither simultaneously: "I am lying". If the statement is true then I am lying about lying; but that means I am not lying, so saying I am lying is a lie; and so on, round and round. Fraser *et al*. (2021) show that humans have a causal feedback loop between conscious and subliminal processes which they call **the Strange Loop**, after Hofstadter's work on the iterative processes of the Human brain (Hofstadter, 2007). They show that this Strange Loop allows apparent contradictions to be presented to consciousness serially rather than simultaneously, meaning that there is no irresolvable contradiction in the statement, “I am lying”. Instead, we generate two different models of reality, neither of which need be “true” because they are both models of reality, not actual things. Computers are unable to do this because they have no consciousness so cannot model, despite being very capable model-resolvers. Tarski (1936) shows that this is a product of human language: we cannot assess the truth values of a bald statement (for instance, “it’s raining”) because the truth comes not from the bare semantics of the sentence but from a higher level of meaning. This higher level lets us say, “it is true that it’s raining”, and then assess the truth of that statement. In this way, not only do we understand that a statement is true, we can assess how it is true, allowing us to negotiate toward Buddy Holly’s meaning when he sang, “The weatherman says clear today; he doesn’t know that you’ve gone away, and it’s raining, raining in my heart”. We can do this because our self-modelling gives us more than just a third-person model of our self, it also gives us an understanding of what and how that model represents: we have not just an awareness of self, we have an awareness of what it means to have a self, an awareness of selfness.

Being meta-aware of my existence, knowing that I am me, is not trivial. The natures of the *I* doing the knowing and the *me* being known have exercised philosophers over thousands of years. Plato described thinking as “the talking of the soul with itself”, and Aristotle (350BCE, [2004] bk9, ch9) said, “being conscious that we are perceiving or thinking is being conscious of our own existence”. Regarding the trustworthiness of our self-modelling, Marcus Aurelius observed, “It never ceases to amaze me: we all love ourselves more than other people; but we care more about their opinions than our own”.

The multiplicity of modelled selves produced by speculation, or *what-if*, makes the nature of the “real self” uncertain: different modelled selves can represent contextually different – and contradictory – targets for the self to aim to become. Viewing my self as a second or third person creates the knowledge that, contrary to our basic evolutionary instincts, other people can be as valuable to me as me – or sometimes more valuable. Interpreting others' models of me to be as valid as my own self-models increases the number of targets for the self to aim for, and adds targets that are not necessarily in my evolutionary interests. If sufficient people challenge my own models of me then I can be led to believe that the best me conforms to the majority's stereotyping of me – and if enough people offer me a negative image of myself then I am more likely to build my own negative image of myself (Granberg *et al*., 2001).

So being meta-aware of ourselves may be empowering intellectually, but it can also have disturbingly unDarwinian effects on our personal survival. Yet acquiescence to the wills of others provides an important clue to human cooperation. To maintain a healthy self-image we must ensure that the images of ourselves given to us by others are positive; and to do that we must ensure they think well of us. Our self-image relies on the reputation we construct in the minds of others, so doing things for others helps us preserve our reputation and therefore the image of ourself we receive from others. Altruistic punishment ensures that sociopathy – retaining a positive self-image regardless of the opinions of others – is kept in check; but it remains a genetic capacity in us all, ensuring that our self-image is often not fully reflective of our reputation.

The constant interplay between what we know by introspection and what we know through received knowledge leaves humans in a constant state of doubt, about both the external world and our internal model of the world. Uncertainty in our self-models makes human social cooperation possible, and this, in turn, makes the social exchange of language useful and necessary. As Voltaire expressed it, "Doubt is not a pleasant condition, but certainty is absurd". Self-doubt may be stressful; but, for humans living in groups valuing altruistic punishment of antisocial behaviours, it is less stressful than unmitigated self-interest.

8 Punishment, Metaphor and Groups

Evolution was far more thrilling to me than the biblical account. Who would not rather be a rising ape than a falling angel? To my juvenile eyes, Darwin was proved true every day. It doesn't take much to make us flip back into monkeys again.
(Terry Pratchett)

in Africa, between 6.5mya (Stringer & Andrews, 2005, 114-117) and 4.5mya (Byrne, 2000), a species ancestral to chimpanzees, bonobos and humans diversified into two populations. One group, the *Pan* clade, took to or stayed in the forest; and the other, the hominin clade, stayed in or took to the open savannah. The *Pan* clade remained in Africa, eventually developing into chimpanzees and bonobos; while the hominin clade became nomadic, which allowed isolated populations to evolve separately through allopatric, or geographic, speciation (Jolly, 2009). This resulted in many more species and subspecies in the hominin clade than in the *Pan* clade, with at least four genera: *Ardipithecus*, *Australopithecus*, *Paranthropus* and *Homo*. Eventually a branch of this family tree developed into *Homo sapiens*, the only remaining hominin species.

The evolutionary story of our species remains incomplete, but some things are known. We know that our ancestors adopted bipedalism early, probably between 6mya (Richmond & Jungers, 2008) and 7mya (Daver *et al*., 2022). We know that early hominin brain size did not differ significantly from that of the *Pan* clade, remaining at about 450cc until about 2.5mya, when a threefold increase began in the genus *Homo* (Dunbar, 1993). We know that tool use is typical in our lineage, and several technological breakthroughs have been identified: the roughly-hewn stone tools of

Lomekwi, about 3.3mya (Harmand *et al.*, 2015); the Oldowan stone choppers, from about 2.6mya (Faisal *et al.*, 2010); the Acheulean teardrop axe-heads, from about 1.75mya (Shipton *et al.*, 2019); the Levallois prepared core tools, from about 400kya (Shipton *et al.*, 2013); the Mousterian flake technology, from about 200kya (Hovers & Belfer-Cohen, 2013); the Palaeolithic technological revolution, which was fully established by about 40kya (Conard, 2005); and the Neolithic domestication revolution of about 10kya (Johanson & Edgar, 1996, 250-261). In terms of art, there are indications of body adornment from about 80kya (Henshilwood *et al.*, 2004), sculpted objects from 32kya (Conard, 2003), and depictive art (cave paintings) from about 30kya (Balter, 2009).

Yet, despite the importance of these evolutionary capacities, they do not address our particularly human weirdness. Foremost, we are highly socialised, which affects what and how we share: we work together in joint ventures, cooperating in long-term tasks and sharing the long-term gains; we share our social models to build socially agreed reputations about others and ourselves; and we have language, making cooperation and gossip possible (Key & Aiello, 1999). Our society is contractual, based around reciprocal altruism – generosity received today creates an expectation and obligation for reciprocal generosity (Taylor & Day, 2004); and our social contract is supported by altruistic punishment – we have property-related concepts such as theft and cheating, and take revenge against freeriders (Chiappe *et al.*, 2004). Altruistic punishment is itself a social act implemented at the group level – we have rules and laws and we recognise ownership. Individuals do not punish other individuals, the collective abstraction of society punishes. Because individuals rely on the approval of others to prosper, we have developed a genetic predisposition for submission to the group: in human evolution, social group cooperators have largely outbred loners (Vugt & Schaller, 2008). This, in turn, brought us the advantages of specialisation: I do not have to make everything I need, I can rely on the social contract of reciprocal altruism and altruistic punishment to exchange my labour for yours. This exchange may have started as a continuous cycle of reciprocal giving to build personal reputations (Mauss, 1950); but, as the circle of giving widened, an individual’s commitment to the social contract became symbolically stored into agreed units of exchange, or money.

This leaves the human weirdness we explored in the last chapter: we make models of ourselves. This chapter explores how our objective self-imaging relates to our capacity for altruistic punishment, and to our reification of the group as an entity superior to ourselves. Both these traits are problematic in evolutionary terms, although for different reasons; but they both have important roles in the process of becoming human.

The Value of Altruistic Punishment

Altruistic punishment is uncommon in nature, but not unknown; and it is often a response to inaccurate or deceptive signalling (Tibbetts & Dale, 2004). In many cases, signal honesty is maintained by its cost to the sender: the threshold at which the receiver accepts the signal is so high that cheating by the sender is too expensive (Zahavi & Zahavi, 1997). Altruistic punishment is not needed in costly signalling.

Costly signalling is useful where the receiver must make judgements based solely on the signal value (such as mating signals), but it is less useful for kin-selected or reciprocally altruistic signals. In these cases, a "contract" outside of the signal means the signal can be reduced in cost without affecting value (Bergstrom & Lachmann, 1997); and this is what we see in the warning signals and exchange of social information between members of human groups, especially between relatives. These cheap signals are open to cheating, however, so the receiver must be able to evaluate the signal/signaller combination rather than just the signal itself. However, Számadó *et al.* (2022) have shown that cheating is not a simple matter of "if I can, I will": the cost of cheating is measured in humans not by immediate advantage but by marginal costs and trade-offs; and manipulating trade-offs had a greater influence on honesty, and therefore signal reliability, than manipulating signal costs.

Humans are exceptional in the honesty of their signals. Despite living in groups with limited relatedness, our signals are cheap to make; we rely, instead, on complex social systems to enforce signal honesty (Knight, 2008). Not all our signalling needs to be factually honest, we also value the shared fantasy of storytelling; but if the receiver expects signal/referent veracity and the sender does not deliver, there are group-level punishments available to use against the cheat (Gintis, 2008). These punishments are initiated by the group, so individuals must be willing to punish those who have offended against unrelated others as well as their own offenders. Group-driven altruistic punishment is difficult to explain in evolutionary terms: how does supporting unrelated others enhance my personal reproductive fitness? Yet it is what humans do. We extend our definition of kin to include unrelated members of our groups – or, at least, we treat them in the same way we treat our kin. The affinity of group membership seems to be as significant in determining and enforcing cooperation as the consanguinity of kin.

There is also some evidence that altruistic punishment need not be inevitable. When modelling natural groups, where mistakes are often forgiven rather than punished, a small number of altruistic individuals who are more forgiving than punishing can stabilise cooperation in the group. Most members of the group can be conditional cooperators, withholding cooperation equally from cheats and those who make mistakes; but the leavening of forgiving individuals makes the whole system more sociable and more trusting (Battu & Rahwan, 2023). Pure altruism can directly disadvantage the individual in isolated transactions; but in a socially transactional environment the same pure altruism can advantage the individual indirectly through their membership of the group. The advantage to the group becomes an advantage shared between the individual members of the group. *Pay it forward* may be more humanlike than the consistently rational and narrowly self-interested *Homo economicus* model allows.

The Value of Metaphor

When discussing metaphor, we must first make clear what we mean by the term. The traditional view is that metaphor is a tool of language, a trick by which we can describe something in terms of another. It is seen as a figure of speech, a sophisticated but superficial way of expanding the descriptive power of language by borrowing between “semantic realms”. However, Lakoff (1993 [2006]) looks at metaphor in a different way:

> As a cognitive scientist and a linguist, one asks: what are the generalizations governing the linguistic expressions referred to classically as 'poetic metaphors'? When this question is answered rigorously, the classical theory turns out to be false. The generalizations governing poetic metaphorical expressions are not in language, but in thought: they are general mappings across conceptual domains. Moreover, these general principles which take the form of conceptual mappings, apply not just to novel poetic expressions, but to much of ordinary everyday language.
>
> In short, the locus of metaphor is not in language at all, but in the way we conceptualise one mental domain in terms of another. (185)

Lakoff calls this new type of metaphor a conceptual metaphor, to distinguish it from the traditionally restrictive poetic metaphor.

Communicating using metaphors does seem to be distinctively human. Yet, when we look at metaphor in cognition, it is widespread throughout nature. For instance, intrinsically recognising colourful striping, particularly black and yellow striping, as an arbitrary sign of danger saves a lot of painful experimentation; and aversion to open stretches of water is a fit strategy if there are likely to be predators in the water. Cognitive metaphors of this type existed before humans, indicating that there is a fundamental conceptual metaphor of metaphor: Y, AS A PROXY FOR X, IS X[9].

Conceptual metaphors initially rely on personal experience: *once bitten, twice shy* is the rule. However, these experiential metaphors can easily become innate knowledge: they are simple and reliable enough to be encoded at the genetic level, because individuals for whom the metaphor is innate do better than those that need to personally experience it. If we instead imagine the metaphors being shared communicatively then we immediately encounter the two communication dilemmas: why should the sender give away valuable information for free, and why should the receiver believe free information? If I tell you that black-and-yellow stripes indicate danger then it has exactly the same "truthfulness" for you as if I tell you bananas are dangerous: you can only assess the truth of both statements by observation and experimentation. Yet the first is useful free information, while the second is a deliberate attempt to keep all the bananas for myself. The algebra

[9] Conceptual metaphors are traditionally written in uppercase.

of costly signalling tells us that the value of a signal to the receiver is correlated with its cost to the sender: the cheaper the signal is for the sender, the more often it is faked; so the less value it has for the receiver. The transmission of metaphor, as with any language utterance, is an almost cost-free exercise for the sender, therefore it is an almost value-free signal for the receiver.

While metaphor as a cognitive mechanism is commonplace, as a communicative mechanism it seems, at first view, to be exclusively human. Yet mimicry, where one species copies the defensive signals of another species, is common throughout Nature and can be viewed as a kind of metaphor. Wickler (1965) describes three main types of mimicry: Müllerian, where two defensively armed species adopt similar identification marks, such as wasps and bees both using black-and-yellow stripes; Batesian, where a defensively unarmed species adopts the identification marks of an armed species, such as hover flies looking like wasps; and mimesis, where a defensively unarmed species disguises itself as an inanimate object of no interest to the predator, such as a stick insect resembling a twig. A fourth type of mimicry is aggressive mimicry, which comes in two forms: a predator can resemble a species to prey on that species, as is the case for the caterpillar of the Large Blue butterfly (*Maculinea arion*), which mimics the nestmate recognition pheromones of *Myrmica sabuleti* ants to be taken into the nest where they can feast on the ant larvae (Hayes, 2015); or predators can resemble harmless objects, such as a praying mantis resembling a leaf to allow it to seize unsuspecting prey (Holen & Johnstone, 2018). Mimicry can also be used to subvert otherwise honest signals: drongos produce predator warning calls both as genuine warnings and to get other birds (and meerkats) to drop their food and run for cover (Flower, 2011). Mimicry does not need to be visual, it can subvert warning signals in any channel. However, while mimicry is a good example of the metaphor of metaphor at work throughout nature, it is also deeply deceptive, so not a good model for the negotiation toward meaning involved in human language.

As we saw in chapters 2 and 5, language is embodied; which simply means a body is needed to understand the metaphorical links made in everyday language. For instance, the utterance *I see what you mean* works because both speaker and listener understand the concepts of seeing and meaning as physical processes in their own bodies. Their two understandings need not be fully congruent – a congenitally blind person is unlikely to have the same experience of seeing as a sighted person, and a philosopher may have an understanding of meaning different from that of a statistician – but their experiences are close enough that, when a blind philosopher hears the utterance from a sighted statistician, they both recognise that the statistician is claiming an understanding of the intentions behind the philosopher's previous utterance. The statistician is also indicating a level of agreement with the philosopher's previous utterance and simultaneously moving the discussion on; and they are often performing other pragmatic functions enabling the discussion – all of which are only possible because the speaker and listener share a common socio-cultural approach to the universe, enabled by their shared embodied humanity (Lakoff, 2014).

Negotiation toward meaning and the embodiment of language are why humans need the metaphor of metaphor. Between minds, metaphor is a source of obfuscation and confusion (Searle, 1993); but inside a single mind it has practical superordinating, subordinating and coordinating functions. In its superordinating role it offers the cognitive shortcut that an instance, X, is part of a related group of instances, Y, so the features of X must predict the features of Y (this elephant has big ears, so all elephants have big ears). In its subordinating role it offers a different shortcut: the features of Y must predict the features of X (elephants can be dangerous, so this elephant could be dangerous); and it also offers the shortcut that a shared nature of X and Y predicts that X is Y (elephants have trunks, this has a trunk, so this is an elephant). In a coordinating role it offers the shortcut that X shares some features with an unrelated object Z, so it is likely to share other features, too (wasps are black-and-yellow and sting, this is black-and-yellow so it probably stings). All these metaphors do not identify what is logically or actually true; they provide shortcuts that work in enough cases to give a fitness advantage to the animal that uses them.

Cognitively, metaphor is essentially self-deceptive, allowing immediate action based on incomplete information. Sometimes, lying to ourselves is worthwhile, letting us react when our evidence is below the threshold for reaction. While over-reaction does incur costs, the cost of under-reaction may be much higher: better to spook at branches that look like snakes than ignore the snake that looks like a branch. Communicatively, metaphor is not a good vehicle for transferring true information between minds, so it had to wait for minds which could productively deal in deceptive information: minds that can make models, negotiate toward meaning, speculate and confabulate; minds that can disinterestedly treat self as other, willingly believing in six impossible things before breakfast; and minds that can treat modelled abstracts as actualities, treating Popper's World 3, Reality, as actual.

THE GROUP IS AN ENTITY: 1 + 1 = 1?

One of the conceptual metaphors that we humans use in our shared definition of Reality involves treating groups as if they are entities. using the Cognitive Linguistic notation of metaphor (Lakoff & Johnson, 1980), this can be expressed as THE GROUP IS AN ENTITY, a conceptual metaphor key to understanding human socialisation. THE GROUP IS AN ENTITY is also a coordinating metaphor: just like an ENTITY, the GROUP can be treated as a cohesive singularity because of the shared nature of its members. The metaphor also has superordinating features: the individuals in the group are each an entity themself; so the attributes of the individuals, particularly the shared attributes, can be treated as attributes of the group – they are not just shared, they define membership of the group, and establish a shared responsibility by the members for the activities of the group.

By itself, the ENTITY concept appears widespread throughout nature. Segmenting the universe is vital for many fitness-enhancing cognitive activities, such as identifying food, predators, and sexual partners, so it is a skill which enhances the evolutionary fitness of individuals who have it. A primitive version probably evolved early, possibly with multicellularity. Using the social modelling notation of chapter 7, the ENTITY concept can be equated with *sense of other*, the subconscious knowledge that there are other animate objects in the universe. This basic sense of other would have elaborated over evolutionary time, making it possible to classify entities in various ways on a range of interrelated scales; for instance, moving versus stationary, living versus non-living, friend versus foe, in-group versus out-group.

These four classifications have been deliberately selected to illustrate increasing sophistication in identifying entities. Moving versus stationary allows the differentiation of things of interest (moving) from less interesting aspects of the environment (stationary): interesting things (sexual partners, predators, and sometimes food) tend to be mobile. However, interesting things also tend to be living, and so move differently from less interesting things; and it becomes worthwhile to develop a way to distinguish between different ways of moving. Living versus non-living becomes a difference worth identifying (Corning, 2000).

The first two classifications are basic and most animals can make these distinctions. The next classification, friend versus foe, is more sophisticated: it relies on identifying objects as individuals as well as simply classifying them. A social animal must classify another individual according to their reactions, especially if they have extended contact with each other; but identifying that individual requires attention to the characteristics making them individual – the individual is recognised holistically through a range of distinctive traits (Tibbetts *et al.*, 2008).

Recognising friend from foe by individual identification requires a new level of cognitive complexity, corresponding to *awareness of other* in chapter 7 notation. Nonetheless, most warm-blooded animals (mammals and birds) seem capable of identifying individuals. This is particularly so for social mammals, where the imperatives of feeding and breeding require complex relationships with other individuals, as well as the cognitive overhead of individual identification. The Machiavellian intelligence of apes (Byrne, 2000) adds a new level to this *awareness of other*: the value of using my knowledge of others to manipulate them.

The last classification, in-group versus out-group or *Us Versus Them*, is the most problematic, not least because humans generate arbitrary groups incomprehensible to other social animals. Eusocial insects, for instance, have behaviours we describe as group recognition, consensual voting and even cultural conformity; but these behaviours emerge from convergent genetic imperatives, not conscious cognition. The only way for most individuals in a eusocial group to get their genes into the future is to support and defend the fertile few (Bourke & Franks, 1995, 56-66). At the individual level, *in-groupness* and *out-groupness* is recognised simply by type, usually by scent: the right scent indicates a friend while the wrong scent identifies a foe. In humans, classification of in-group versus out-group needs conscious recognition of the group itself (de Waal, 2006b, 52-58); and the evolution of this type of group recognition is not easily explained in terms of Darwinian fitness. In chapter 7 notation, it requires *awareness of self*, which in turn requires the sharing of social calculus, or A-Relationship-B modelling. Conscious group recognition needs a protolanguage with sufficient complexity to share A-Relationship-B models, which probably appeared only in more recent species of *Homo*.

Recognising the concept GROUP would seem uncomplicated: things occur in multiples which can be treated either as a single thing or a set of things – a bunch of grapes is both a single bunch and a group of individual grapes. This concept of GROUP is impersonal, however, with no emotional significance for the individual. The concept of GROUP in terms of in-group versus out-group is considerably more complex.

First, understanding the concept of in-group requires the self to know they are an entity: there is an object, *group*, and another object, *self*, and the *group* contains the *self*. However, this self-object is, in a vitally intimate way, me: both as a cognitive model made by me of myself, and as the person who is generating the cognitive model. This lets me see myself dispassionately – as if I were an external third person (Jordan, 2003) – which poses an evolutionary conundrum: how is being dispassionate about myself a fit strategy when everyone around me is passionately self-serving? Treating myself objectively would seem to put me at a genetic disadvantage; and yet, somehow, this capacity defines our species, making an inability to take a dispassionate personal stance problematic – with those on the autism spectrum (Jordan, 1998) and those diagnosed with sociopathy (Pitchford, 2001) providing quite different examples.

Self-objectivity does help to explain how we objectify the group: as a third party, the self can be treated in the same way as any other third party, as a component of a group. The objective grape is part of a bunch, the objective me is part of a tribe. There is no fitness implication here, unless the tribe is seen not only as a product of its component individuals but also as a superordinate of those individuals. If the tribe is reified and anthropomorphised as being itself an individual, then the imperatives of the individual, me, are subordinate to the imperatives of the entity, the group. Or, to put it another way, my actions become subject to sanctions, not just at the group level but in a self-censoring way, too. I begin to see myself as no more important than any other member of the tribe, and I must therefore be willing (and cognitively able) to subordinate my needs to those of others.

Subordinating self in the emergent metaphor, THE GROUP IS AN ENTITY, is startlingly un-Darwinian. So what is the countervailing fitness advantage that makes self-effacement (or morality) a good strategy for the individual? Living in a group gives many advantages and is widespread throughout nature; but the socialisation that makes self-sacrifice a viable strategy is rare. The eusociality of ants, bees, wasps and termites is one obvious example: with the reproductive capacity of individuals reduced, sometimes to zero, individuals rely on the breeding success of close relatives to get their genes into the next generation. This is not the case for humans: we individually retain our full reproductive capacities, which should impose a standard imperative of nature on us: everyone else is either a potential mate or a potential rival and should be treated as such. How did we evolve a pseudo-eusocial group structure which lets us work together on vastly complex projects?

Social cooperation seems to have evolved in humans before self-modelling: our offspring are feeble and require extensive cooperative nurturing, and our individual capacities are limited and specialised, requiring transactional negotiation with others to meet our needs. However, specialism and coordinated effort do have multiplier effects: we can share skills by teaching and learning, allowing those skills to be passed on to others in new generations; and we can remember and react to the reputations of others, allowing their present actions to be judged against their past reliability.

Together, these markers of socialisation create an environment where the self is better served by cooperating than by selfishly pursuing personal ends. For cooperation to work it need be only slightly better than individualism, because small individual differences in fitness can, over time, become powerful trends for a species. Cooperative capacity appears to have become genetically established early in the *Homo* clade: we are happier in groups than by ourselves, we are happier working with others than working alone, and we are happier conforming to group norms than rebelling; and happiness is nature's reward for not fighting the genetic programme (Pressman *et al.*, 2005). Yet our niceness is not the whole answer; our nastiness, our willingness to altruistically punish, must have also played a productive role.

Altruistic punishment poses another evolutionary conundrum: what fitness advantage do we get from punishing others? To punish others I must use my own time and resources, and accept any costs generated if the punished individual fights back (Fowler, 2005). Altruistic punishment is most effective when a team works together to punish the transgressor, but this raises the problem of the single mutant bottleneck: all evolutionary change must start with a single mutant; but a single mutant altruistic punisher pays all the punishment costs in a population when others pay nothing; so they are compromising their own fitness and advantaging the fitness of others. How can their genes become sufficiently dominant in a population to create reliable team punishment? (Fehr & Fischbacher (2005.)

Despite the single mutant bottleneck, altruistic punishment has been observed in several species, and the more socialised the species, the more likely team punishment becomes. It may be that altruistic punishment is not a mechanism for socialisation but a product of it. If an individual needs group membership to survive and thrive then the mere withdrawal of that membership may be sufficient to punish the individual. For instance, if grooming is necessary for good health, and individuals in a group groom only individuals they like, then altruistic punishment occurs if individuals view social transgressors unfavourably, giving them less grooming and thereby negatively affecting their genetic contribution to the next generation. A simple mechanism of Vigilant Sharing (Erdal & Whiten, 1994) – each individual guarding their share of the benefits from shared enterprises –

accompanied by low-cost social ostracism can be sufficient to create an environment where stronger altruistic punishment can develop.

Vigilant Sharing does not favour alpha individuals. Anyone who tries to dominate the group by seizing more than their share of resources is likely to suffer a reduction of fitness caused by their ostracism, reducing their capacity to act in an alpha role. This lets Vigilant Sharing develop into Boehm's (1999) Reverse Dominance, where alphas are suppressed by active group punishment, and being modest about altruism becomes a costly signal of fitness: I am so fit that I don't even need to advertise my generosity. Modesty is an unusual form of altruistic punishment: the individual is punishing (or censoring) themself to prove their value to the social group. They are making a costly signal by putting the needs of the group before their own, valuing the group's needs above their own – their own needs are a trivial draw on their resources, so they have spare capacity to give to the group (Barclay & Willer, 2007).

Otracism, inflicted on social cheats as an altruistic punishment, can also be used on signalling cheats, discouraging deceptive signals and non-signalling. The high level of socialisation allowed by Reverse Dominance means that signal honesty becomes particularly significant: deception is discouraged because it can create significant losses for the group. In eusocial insect species such as the *hymenoptera* and *isoptera* (Queller, 1994), and in the mole rat species *Heterocephalus glaber* and *Fukomys damarensis* (Burland *et al*., 2002), deceptive signalling is rare and usually punished. There is, however, usually only one dominant fertile female per group in these species, producing high relatedness between group members; and relatedness is a powerful Darwinian incentive to keep signals honest.

Altruistic punishment enhances socialisation only in limited circumstances, and only where considerable socialisation is already present. It cannot, by itself, generate socialisation in an unsocial species, and it is not a fit evolutionary strategy in those circumstances. Once harnessed, however, altruistic punishment can promote socialisation to levels of cooperation that create un-Darwinian effects in the individual, such as self-sacrifice.

With altruistic punishment, Vigilant Sharing, Reverse Dominance, and modesty, the group becomes a superordinate of the individual, allowing self-effacement and self-sacrifice to evolve as advantageous traits in the individual. Human children do seem to have an innate sense of fairness, preferring equal sharing over the mechanical demand model of modern Economics (McAuliffe *et al.*, 2017). THE GROUP IS AN ENTITY is more than a metaphor of superordination and subordination, more than an acquiescence in power relationships; it is an evolutionary mechanism that defines how humans cooperate and communicate together.

THE GROUP IS AN ENTITY: Building Social Structures

The metaphor, THE GROUP IS AN ENTITY, is both a product of social structure and a source of it; but the concept of GROUP is only relevant to the individual when they are consciously aware of it; and the individual need not consciously recognise the group for the group to exist. For instance, it is not necessary for a eusocial insect to have a conscious concept of nestmates or nest in order to work together with others in what appears to be a highly organised way (Hölldobler & Wilson, 2009, ch3); all they need is a sense of almost-self, a simple in-group chemical marker which can be used to recognise other members of their nest. Insect eusociality can even produce, from the viewpoint of humans, the illusion of hierarchy and central organisation. This is possible partly because of the range of morphs and roles that different individuals in the nest can have – for example, queens, drones, food collectors, scouts, brood nurses, soldiers, nest border control – and, in some species of honeypot ants such as *Camponotus inflatus*, living larders (Islam *et al.*, 2022); or, in phragmotic ants such as *Carebara phragmotica*, even doors (Fischer *et al.*, 2015). However, the hierarchy and central organisation we see in a fully eusocial species is mostly an illusory metaphor imposed on the species by our human pseudo-eusocial social models.

Eusocial insects show that group affiliation can emerge from a series of genetically driven small cooperations between individuals, it does not need to be cognitively recognised by its participants: the group can exist without being seen to exist by its members. So how do we bridge the gap between the genetically driven emergent groups of eusociality and the consciously recognised groups we humans have? The first step is likely to be consciously recognising others as intentional individuals (awareness of other), and maintaining individual accounts of cooperation to identify individuals who regularly cooperate with me and those who are not – necessary for any species with conscious control over their cooperation. Conscious control creates the possibility of freeriding, so recognition of others as individuals, and keeping accounts of their cooperative behaviours, lets me identify freeriders around me (Gardner & West, 2004). social accounting relies on Relationship-A cognition, establishing a hub-and-spoke model of my relationships with others (the spokes), around the fixed, subliminal (and therefore unmodelled) hub of me.

The next stage of social accounting identifies not just how individuals cooperate directly with me, but how they cooperate with each other, differentiating those who are generally freeriding from those who are not cooperating with me because it's me. This two-argument, A-Relationship-B cognition is essentially abstract and dispassionate – I am assessing the relationship between A and B separate from my personal relationships with A and B. This is not hub-and-spoke modelling, it is node-and-link modelling, with each node being an individual in my social orbit, and each link being a relationship between two of those individuals. Where I am the unmodelled gravitational centre of my Relationship-A cognition, I am not part of my A-Relationship-B cognition: I am neither an A nor a B in any of the relationships modelled.

I can, however, combine my A-Relationship-B modelling with my Relationship-A modelling to identify individuals with whom I can form alliances, based on our mutual interests and shared cooperators. These alliances are initially quite static, heavily influenced by genetic relatedness – like the alliances between chimpanzee males, who are often brothers or first cousins (Mitani, 2009). In human evolution, however, alliances became more *ad hoc*, changing as frequently as the node-and-link web of cooperation itself changed (Tomasello & Gonzalez-Cabrera, 2017). Individuals who quickly adjusted their alliances to match changes in their node-and-link social models were likely more successful at getting their genes into the future than less socially adept individuals, which in turn solidified the reality of the group; but initially only the boundaries of the group were needed, dividing the universe into individuals with a place inside the social modeller's model and others. Initially, the group would have been recognised as a set of individuals without the group being recognised as an entity itself.

Consciously modelling the group as an entity was a later development. Relationship-A and A-Relationship-B modelling can both be conscious processes, but A-Relationship-B modelling is also dispassionate: modelled others have their own intentions and agendas, and I need to model them as intentional beings. Each node of my modelled network has intentions about other nodes; they are all instances in a repeating pattern of *others*, each with their own selfness. Understanding that other individuals have intentions – evidenced by their relationships with others – creates a more nuanced and sophisticated cognitive model of the group around me. With A-Relationship-B social modelling I have knowledge of, and control over, my environment, making evolution toward more complex modelling possible: the costs of A-Relationship-B modelling are less than the value it brings, so they are worth paying.

New possibilities appear when the group is treated as an entity. The group is no longer just the physical actuality of its members, it becomes both an internalised Popperian World 2 virtual concept, and a shared World 3 entity given reality by consensus. How in-groups and out-groups are differentiated may appear reasonable in these World 3 cases, but it is often based upon arbitrary mutualities: why do individuals good at moving bits of wood across a tessellated board according to arbitrary conventions need to group together in chess clubs? What fitness advantages, other than the joy of chess, do they get?

Treating a group as an entity composed of individuals allows it to be subdivided into smaller groups, which are simultaneously aggregates of individuals and subgroup entities within the larger group. This introduces hierarchy into social modelling, allowing membership of a group to be situational: an individual can simultaneously be a member of several entities, and membership of one entity does not automatically preclude membership of another. Treating the group or subgroup as an entity not only makes membership of multiple entities a fit strategy, it can also make conformism to the arbitrary rules of the entity a fit strategy. The arbitrary meta-rules of culture determine the nature of the entities available for individuals to join, and the arbitrary rules within the group or subgroup determine which individuals join them.

We used to believe that only humans live in a social environment of hierarchical groups (groups within groups), but we now recognise that it is frequent enough to be unremarkable (Grueter *et al*., 2020); and there is now evidence that we may not be the only species to use third-order hierarchies (groups within groups within groups). Connor *et al*. (2022) have shown that dolphin males of the species *Tursiops aduncus*, who live in a single group of over 200 individuals, form social groups with at least three levels of strategic alliances. First-order alliances consist of two or three males who work together to corral and mate with receptive females. These alliances then form alliances with other first-order alliances to make second-order alliances of between four and fourteen males which, unlike the first-order alliances, are much more formalised. An individual seldom moves between second-order alliances, although he may be a member of several first-order alliances, moving between them to maximise his breeding opportunities. Within the second-order alliances females are often allowed to mate with other second-order allies, and males who are part of several first-order alliances therefore tend to get more genes into the future. The second-order alliances also form third-order alliances with other second-order alliances, calling on them in competitions with rival second-order alliances, but not sharing females with them. Third-order alliances can comprise 30 or more individuals; but the alliance is temporary, limited to the duration of the competition; and, while there is some reliability in third-order alliances, they tend not to be long-lasting. Nonetheless, this example from dolphins demonstrates how risky it is to draw a line and declare modern humans to be the only species on our side of the line.

THE GROUP IS AN ENTITY: Not Just for Humans

Using metaphors in general cognition is unlikely to be just a human trait; it offers cognitive short-cuts which provide tangible advantages even for simple life-forms. Some metaphors must, therefore, have their origins in pre-human cognition. For instance, the classic metaphor MORE IS UP is a simple product of gravity: as a pile grows it also rises. Identifying tall things as more and short things as less therefore becomes a cognitive short-cut which enhances fitness more often than it diminishes it.

Unlike MORE IS UP, however, the metaphor THE GROUP IS AN ENTITY does not have a natural relationship with the actual world. Indeed, it seems counter-intuitive in a selfish Darwinian universe. Subordination of the self to the aims of the group is explicable in eusocial animals, but not in a species with individual fertility, no matter how socialised they may be. There are indications that other primates treat their group as an entity; and their friend-versus-foe recognition shows that in-group versus out-group is an important dichotomy. Goodall (1990, ch10) has even reported that chimpanzees go to war – although it is a different type of war from the formal conflicts that even gatherer-hunter human groups can fight; and it is unlikely that any chimpanzee gives cognitive houseroom to the idea, *dulce et decorum est pro patria mori* (it is sweet and proper to die for your country). Yet *Homo sapiens* and *Tursiops aduncus* are not the only animals with a sophisticated conception of groupness: among others, elephants are aware of both their herd membership and their family group membership within the herd (Bradshaw, 2004).

The origins of the metaphor, THE GROUP IS AN ENTITY, must lie in the complexities of socialisation – it is an essentially social expression. To make it work, Machiavellian socialisation (I get more advantages in the group than out) must have been supplemented by cooperative altruistic punishment of unsocial activities, itself a product of Vigilant Sharing: every individual guarding their personal share and individually punishing the greedy – and, indeed, punishing those who do not themselves punish the greedy.

When greed is punished, whether simultaneously by the group or serially by individuals, the greedy suffer in terms of breeding success; which feeds back into Vigilant Sharing, enhancing its fitness as a strategy. Vigilant Sharing in turn leads to Reverse Dominance, where prospective alphas are suppressed by group action – the self-aggrandizing displays of alphas are not tolerated by the rest of the group, who work together to disfavour prospective alphas.

While the metaphor THE GROUP IS AN ENTITY is not a product of human evolution alone, it does have more significance for human groups than for others: it is significant for any species with high levels of socialisation and significant cognitive capacity, and humans have the greatest socialisation of any non-eusocial species and a substantial cognitive capacity. THE GROUP IS AN ENTITY has become a conscious moral imperative for *Homo sapiens*, because humans see their groups as both extensions to and containers for their individuality, and not just environmental events affecting the individual. The individual defines and chooses the group, and the group defines and chooses the individual; and without the conscious moral imperative that the group can be more important than the individual, the complexities of our modern social systems, including language grammar, would be impossible.

9 Language Grammar: From Sources to Complexity

Neither Aristotelian nor Russellian rules give the exact logic of any expression of ordinary language; for ordinary language has no exact logic.
(Peter F. Strawson)

Human society is built around cooperation. This is unusual in nature, because most species are deeply Darwinian: the primary directive of the organism is to survive; and the secondary directive is to get genetic copies of the individual into the future. The phrase “survival of the fittest” merges these two directives into one, because individual survival is sometimes secondary to reproduction; for instance, male spiders (Andrade, 2003) and praying mantises (Fisher *et al*., 2020) offer themselves to the female as a meal to ensure a successful mating. However, this final self-sacrifice would not be possible without all the self-serving predator avoidances which succeeded on the arthropod’s way to the prize.

While intraspecies close cooperation may be unusual nature, it is far from unknown. Eusocial insect societies are nests of closely cooperating individuals, where individual survival is often made moot by rendering most of the individuals sterile; the only way an individual can get their genes into the future is to ensure the survival of the few fertile individuals in the nest. This creates an environment where sterile individuals become expendable in defence of the fertile individuals, a situation known as extraordinary self-sacrifice. Krupp & Maciejewski (2022) show that this evolves where neighbouring individuals are closely related, so self-sacrifice for neighbouring others becomes a fit strategy: the primary cost of self-sacrifice becomes a secondary benefit to close relatives. It is easy to see how communal nesting of close relatives can develop toward selective sterility and eusociality. This is not, however, the case for primates, and it does not explain how humans evolved toward self-sacrifice for strangers, which is here called pseudo-eusociality.

To understand pseudo-eusociality, we need to look at the role of reputation in human societies. Reputation is a game of two sides: how do I determine your reputation, and how do I build my own? We have already seen examples of how we manage the reputations of others: Vigilant Sharing lets me compare your behaviour against my expectation of being treated fairly; and Reverse Dominance lets me compare your behaviour against my expectation of your modesty. However, a third mechanism is available to manage reputations, described by Dessalles (2014a) as the Political Singularity: if early humans had tools capable of injuring or killing prey animals, they could use those tools to injure or kill rivals attempting to assert dominance. The altruistic punishment of the Political Singularity allows me to take action against your unfairness or tyranny, either by myself or, with greater certainty of success, in cooperation with others. This has considerable influence on how I build my own reputation and the type of reputation I build. I need a reputation that is notable, otherwise I am just another unremarkable member of the group; but if my reputation is too remarkable then I am liable to go beyond notable into notorious, and thus become a candidate for the Political Singularity.

In a society without human language, building my reputation is governed by what I do. How I publicly interact with other individuals determines my reputation within my group: am I agreeable or irascible, conscientious or lackadaisical, approachable or unsociable, confident or neurotic, receptive or aloof? And, if I am a chimpanzee, am I unaggressive or dominant (Altschul *et al.*, 2018)? Observation of personality to determine reputation has been recorded in chimpanzees: a small group of chimpanzees saw one human regularly give food to others (humans and chimpanzees), and another human regularly refuse to share. When offered the opportunity, half of the group preferentially begged for food from the generous donor, while the rest solicited both donors equally (Subiaul *et al.*, 2008). In another experiment, chimpanzees also cooperated preferentially with individuals who cooperated with them. The experiment required two or more chimpanzees to cooperate in pulling a tray of food into grabbing range, at which point there were several opportunities for cheating: one of the pullers could take all the food, a dominant could forcibly displace one of the pullers to take a share of the food, a freeriding individual could steal the food, or a dominant could seize the food from a successful puller by force. The experimenters found that freeriders were soon identified, and pullers stopped pulling when they were close by; and dominants were discouraged by pullers seeking closely related pulling partners, or those of similar rank. In more than 80% of over 4,000 trials, pulling teams cooperated successfully in getting and sharing the food, while freeriders were often punished by the pullers or even by third parties (Suchak *et al.*, 2016).

Human language gives us another method for building and assessing reputations: we can share our social knowledge with others, and they can share with us – we have access to Shared Social Calculus. Building my reputation is still governed largely by what I do, but it is also influenced by what others share about me; and those others are willing to share information about me because this information-sharing is also a form of cooperation, so they can build their own reputations while building – or ruining – mine. The information shared is all about the relationships between members of the group, which means that it has a particular structure, or grammar: cognitive social modelling has a structure of A-Relationship-B, so the sharing of this modelling needs to have the same structure. The grammar used by language has a basic structure of subject-action-object, or *[things]-[do things to]-[things]* (Edwardes, 2018), which maps closely to the A-Relationship-B cognitive form.

Cognitively, however, sharing social models also relies on recognising both the differences and similarities between my viewpoint and those of others. This requires Theory of Mind – knowing that others have minds, that then do not necessarily know what I know, and vice versa (Fenici, 2012). The sharing of social models is not, therefore, about sharing truths: it is just as easy to share untruths and fantasies, and it is often in the sharer's interest to do so. As Dessalles (2000) shows, this means that the currency of exchange in this system is not accuracy but relevance. It is not important whether the information offered to me is accurate but whether it is useful to me; which, in turn, means that it is not necessarily the supply of inaccurate information that must be discouraged by punishment, it is the supply of irrelevant information. Inaccurate information may still be useful to me: it may not accurately tell me about the relationship between A and B, especially if I already have a model of that relationship; but it can be used to interpolate the speaker's relationships with both A and B. From the extended cognitive form, *[speaker said A-Relationship-B]*, the listener can build information about A-Relationship-B, speaker-Relationship-A and speaker-Relationship-B.

Mostly, the relevance of shared social information is not in the information itself, it is in the information-supplier's reputation, which the information-receiver assesses using their social calculus. The first language grammar mechanisms are therefore likely to have made the exchange of reputational information easier. These mechanisms include: object-action differentiation, so that the nodes (or individuals) can be differentiated from the links (or relationships); the one-, two- and three-argument forms to allow the node-and-link constructs to be presented propositionally; hierarchies of meaning within an utterance to allow differentiation between actor, action, patient, and information source; descriptors (adjectives and adverbs) to indicate the speaker's stances on the actor, action, patient, and information source; and the identification of the three persons involved in a signalling event (speaker, hearer and other; or first, second and third person).

There are two routes by which this language grammar could have developed: the individual mechanisms could each have had their own genesis, or the whole structure could have arisen as a single event. If we look for a single origin for the whole system then we have a problem with multiple dimensions of difficulty. The first question to answer is, what fitness pressure required the grammar system to spring fully-formed into the minds of humans? The second question is, what evolutionary event made the grammar system a necessity in communication? The third is, what underlying single principle produced all the complexities we see in language grammar? And the fourth is, how does a species-wide language grammar system produce so much variation between individual languages? Any theory identifying language grammar as the product of a single cognitive system should at least attempt to answer these questions.

Generativism, which espouses this approach, has had mixed success in providing answers to all the questions, and solutions offered for one question have often encountered difficulties with other questions. The current leading Generative solution for the sources of language grammar is the single cognitive mechanism of recursion or Merge, the capacity to merge base materials (nouns, verbs, *etc.*) together to make compound components with the same functions as the base components. For instance, the previous sentence starts with a noun phrase, *The current leading Generative solution for the sources of language grammar*, which could be replaced by *it* or *this* without affecting the grammatical integrity of the sentence (although the semantic integrity may be compromised); and it contains three other noun phrases (*The current leading Generative solution*, *the sources*, and *language grammar*). Recursion, or Merge, is without doubt an important mechanism of language grammar; but it is neither an early development, nor is it fundamental to the working of grammar. The answers given by Merge to the four questions are as follows: first, Merge is a random mutation not subject to evolution before its appearance, it fundamentally reorganised human cognition, and it was cognitively so advantageous that recursive humans quickly replaced non-recursive humans; second, because communication is sharing cognition, the adoption of Merge into communication needs no special explanation; third, language structure is only superficially complex, Merge makes it both universal and simple; and fourth, the grammatical variation between languages is illusory – to a Martian, they would all seem to be one language.

Rather than a single evolutionary macromutation, an incremental approach to the sources of language grammar implicates our social cognition as an indirect but consequential source of cognitive complexity (Tomasello, 2003a, 282-283). For non-Generativists, the interpersonal nature of linguistic communication seems to intrude at every level. A semiotic approach, studying the forms of language in terms of the tasks performed, must address the specifics of human socialisation (Hurford, 2007, ch7); an evidenced grammatical approach, looking at different languages to identify common forms, identifies many idiosyncratic forms (both between and within languages) which, nonetheless, are acceptable to sender and receiver (Sampson, 2005, ch5); and an anthropological approach, explaining the development of signalling complexity from cultural complexity, remains largely unexamined by Generativism (Steels, 2003). The phonological approach, examining the rise and stabilisation of speech complexity in human evolution, also poses problems for Generativist analysis (Port & Leary, 2005); but that is not covered here.

The Generativist grail-quest for a single, elegant solution to explain the whole of language remains a viable objective; but, in the face of the difficulties already encountered, this book abides by Einstein's dictum: "If you are out to describe the truth, leave elegance to the tailor."

Grammar is a Moving Target

Even with a single macromutation, early humans could not have suddenly started to share fully formed language grammar, there must have been a process that negotiated it into existence. This would have involved both the invention of grammatical forms to meet new communicative needs and the realignment of existing forms to accommodate increased complexity, a process of grammatical change known by linguists as grammaticalization. No language is fully fixed and stable, even dead languages are subject to change as new texts are discovered; and in live languages, changes often lead to new dialects or even new languages, a process we see at work today.

For instance, one small grammatical change in colloquial British English is the appearance of the term *innit*: originally a contraction of *isn't it*, itself a reanalysis of *is it not*, *innit* was originally used as a pragmatic marker or tag question seeking agreement for a preceding statement (*it's hot, innit?* or *this is the way*, *innit?*). It has now evolved into a general marker for sharing any information (*you're happy*, *innit* or *he's going home, innit*). The normal rule of agreement, that the subject of the main sentence dictates the form of the subordinate subject and verb (*you're happy, aren't you* or *he's going home, isn't he*) has disappeared, and *innit* has become an adverbial (Block, 2008, 194). This lexicalisation is becoming more formalised, leaving *innit* no longer analysable into *is it not* (Martínez, 2014). It is following a similar path to *isn't it*, where the uncontracted form, **is not it*, is currently seen as ungrammatical.

Grammaticalization changes seem to follow rules. These rules are not invariable, but they do have statistical significance. Hopper & Traugott (1993, ch3) show that language change happens pragmatically (in terms of social language use), semantically (in terms of meaning), morphologically (in terms of language structure) and phonologically (in terms of sound or gesture representations): language forms previously communally agreed as grammatical are reinterpreted by groups in the community and then shared with and adopted by other members of the community. Grammatical reinterpretations include phonetic erosion (such as the loss of a syllable between *isn't it* and *innit*), semantic bleaching (the loss of meaning, such as the loss of intention in the future auxiliary verb, *will*), obligatorification (making unnecessary forms obligatory, such as using the auxiliary verb *do* in questions), widening (gaining more usages, as in the expansion of *tree* from a biological object to any structure with multiple branching) and narrowing (losing usages, as in limiting the meaning of *meat* from any food to just animal flesh). The English noun *takeaway* shows how these reinterpretations can work together: originally a general-purpose phrasal verb (*take away*), it was adopted early in the 20th century as a single word (phonetic erosion) to refer to a restaurant where meals could be purchased for consumption elsewhere; but its usage soon came to refer to the food itself (narrowing). It now also refers to cooked food delivered to the home (widening), and which therefore involves no taking away (semantic bleaching). Another example is the pronunciation of the verb *police* and the noun *police*: there is no difference in Standard English. Many English two-syllable words which are both nouns and verbs emphasise the first syllable of the noun form (*e.g.*, *decrease*) and the second syllable of the verb form (*decrease*). This is not the case for *police*, which has led to reinterpretation of the noun sound in some dialects; for instance, in parts of Scotland the noun is pronounced *polis*, with a shortened /i/ (obligatorification).

One aspect of grammaticalization of interest here is directionality: do grammatical changes occur in one direction only, or can change be multidirectional? Hopper & Traugott (1993, ch5) are neutral on this, detecting both processes at work; but Dixon (1997, 41-43) sees change as cyclical. He describes grammaticalization in terms of typological change: analytic languages, in which each word is a single meaning-unit, tend to become agglutinating languages, in which words contain multiple but separable meanings; agglutinating languages tend to become fusional languages, in which single syllables can perform multiple meaning functions; and fusional languages tend to become analytic.

The differences between the typologies do not mean that there is a hierarchy of complexity between the three typologies. Indeed, if the typologies are not equally complex then grammaticalization has a “slope problem”: a need for increasing complexity would mean that all languages would tend toward the most complex typology, and this does not accord with the data. For this reason, typological change is not seen as a route to language grammar complexity. Dixon’s cyclical grammaticalization is a one-directional system, but it does not have a slope problem: all typologies must be equally effective if the cycle is to be continuous.

Typologies are useful for classifying languages, but it is really the functions within languages that are typological. Most languages contain examples of all three typologies, although one may be dominant. In the case of modern English, most linguists identify it as analytic – each word has a single semantic or grammatical role. However, it also has many fusional aspects: for instance, the suffix *-ed* indicates that the attached word is a verb, that the action of the verb is in the past, and that the action has been completed at the time of speaking. English also makes extensive use of agglutination: *down* and *cast* can be combined to make *downcast*; *broad* and *cast* can similarly be combined into *broadcast* (to sow seeds by throwing them across the ground), and has been widened to describe the distribution of radio and television programmes; while the words *never*, *the* and *less* can be combined to make *nevertheless*, although why it means “despite what has just been said” is opaque.

Grammaticalization can also be seen as culture imposing itself on language: to ensure that increasing cultural complexity can be shared, language needs to become as complex as the culture. Complexity is not just a function of language typology, it is generated by grammaticalization itself; which, in turn, is driven by the need to communicate cultural complexity. The typology of a language is not itself a marker of complexity, complexity can grow (or diminish) with each move to a new typology – multidirectional typological systems can be steplike, and cyclical systems can be helical.

Grammaticalization and the Sources of Language Grammar

Grammaticalization requires the pre-existence of language: language grammar cannot exist without language. Yet grammatical structures must have always been a part of language because language without grammar is not language, it is just communication. There seems to be a chicken-and-egg conundrum of which came first; and, like the chicken-and-egg conundrum, the solution is that a false dichotomy has been created. Just as eggs preceded chickens because chickens are not the only source of eggs, so grammar preceded language because language is not the only source of grammar. Eggs were being laid by the precursors of chickens, and grammar was being used in the private cognition of individual brains before it was exapted for public communication between brains.

So, what do proponents of grammaticalization say about the sources of language grammar? For Dixon (1997, 63-66), cyclical typological change provides little scope for languages to increase in complexity: language appears suddenly, swiftly developing from nothing to high complexity within a few generations as all the required cognitive grammar was negotiated into communicative use. Language complexity then became largely fixed, and all that was left for grammaticalization was a continual negotiation of usage between the three language typologies. Deutscher (2005, ch7) takes the view that we cannot know the sources of language grammar before the two-word stage. Symbolic usage began with two-word utterances, and we cannot usefully go back before this point. Pre-symbolic communication cannot be called language, and we cannot know the sources of language grammar except through language itself. Grammaticalization is an ongoing response to the need for an increasingly complex communicative structure, but how it came about remains a mystery.

In contrast, Heine & Kuteva (2007) see language complexity growing slowly and steadily through a series of grammatical layers: nouns; verbs; adjectives and adverbs; demonstratives, adpositions, aspects, and negation; extensions of form to produce complex constructs; reduction of meaning to convert meaning words into marker words; and the reduction of markers into morphemes. The earliest language grammar contained only nouns to label objects and people. This layer of nouns was then supplemented by labels for actions, or verbs, giving us the one-argument grammar form; then came qualifiers for nouns and verbs, followed by various types of marker words; and then, in a series of steps, the other grammatical resources of language developed. For Allen & Seidenberg (1999) words must have been present before grammar could begin. Grammaticality is a statistical relationship which emerges from the negotiation toward meaning between speakers and listeners – a negotiation toward grammar.

Hopper & Traugott (1993, 33-38) do not speculate on language origins, but they do make an important point about early languages: we should not assume that an early language contained something not evidenced in at least one currently known language; we cannot know about grammatical features that no longer exist. Nettle (1999, ch2), however, suggests that language change follows the same rules as genetic change: first, when the speakers of a language split into isolated groups, variation occurs; second, language contact can produce consolidation, where the two separate languages become more alike, or differentiation, where they move further apart. Language changes result from social contact, not from the languages themselves. Early languages would have had fewer contacts with other groups, and so would have undergone less change. We cannot know for certain that the grammar features described today are definitive; but we can expect modern language change to be faster, and therefore variation in modern languages to be greater than that of early languages. The chance that a grammatical feature was present in the narrow variation of early languages and not in today's wide variation is small.

Carstairs-McCarthy (1999, ch5) takes the two-argument form as the basic syntactic structure, and he compares the Generative analysis of this form with the traditional analysis of the syllable. A syllable consists of an onset, the sound that announces the syllable, and a rhyme which completes the syllable; so the syllable *dog* has an onset of *d* and a rhyme of *og*. The rhyme is also divisible, consisting of a nucleus (*o*) and a coda (*g*). A Generative analysis of a two-argument sentence such as *Alf likes Beth* produces a similar tree structure: a noun phrase or subject (*Alf*) links to a verb phrase or predicate (*likes Beth*); the verb phrase is then divisible into a verb (*likes*) and a noun phrase or object (*Beth*). However, while the syllable and the two-argument sentence share the same logical structure of [A[uB]], what does this signify? The collocation of the two forms indicates a human capacity to identify analogies, but it is not evidence that the two forms share a common mechanism: they have the same logical form without sharing a common source.

The Beginnings of Language Grammar

If grammaticalization is not a primary source for language grammar, what is? To answer this, we must differentiate cognitive grammar, the algorithms we use to make our mental models, from communicative grammar, the algorithms we use to share our mental models. Both grammars let us organise objects together by establishing the actions, or relations, which exist between them; but they are different in one important way. Cognitive grammar is not shared, so idiosyncrasies are not important; but communicative grammar must be interpreted by others, so idiosyncrasies impose limits on what can be expressed.

Just as eggs preceded chickens, so cognitive grammar must have preceded communicative grammar. The cognitive forms used by social modelling calculus prefigure and determine the forms used to exchange social models: for successful communication, communicative grammar must map to cognitive grammar. However, not all communicative grammar is prefigured in cognitive grammar: the act of communication requires forms and structures not just for communication but for negotiating toward meaning. So which forms would have been required for sharing social calculus, and which for negotiation toward meaning?

A consensus from grammaticalization indicates that nouns, verbs and the one-argument (action-object) form would have been early requirements for grammatical communication. These three functions enable a range of communicative actions: Naming individuals to get attention; Semanticity, or agreeing terms to represent concrete objects (*e.g., wolf, apple*); Agreement and Negation (*e.g., yes, no*); Manding, or using imperatives to indicate an action and a target object (*e.g., shake tree*); Stating, to indicate an existent state (*e.g., meat cooks*); and Coordinating, or checking agreement about meaning (*e.g., meat cooks* said with a questioning tone). However, while this level of grammaticalization is language-like, many linguists dismiss it as pre-language; I therefore call it first-stage protolanguage, or protolanguage 1.

If we add the exchange of social models to protolanguage 1 then we need at least the two-argument form (the subject-action-object form) to express the relationship between two individuals. This form is also the simplest basic structure requiring syntax: if the relationship between the two individuals is bi-directional then order is unimportant – *Alf likes Beth* is the same as *Beth likes Alf*; but if it is one-directional then the actor (subject) and the patient or receiver (object) must be differentiated – *Alf hit Beth* is quite different from *Beth hit Alf*. Once again, though, many linguists dismiss this as pre-language, so I call it second-stage protolanguage, or protolanguage 2.

While Protolanguage 2 may not be full language, it does need syntax, making it an important stepping-stone between Protolanguage and complex language. Cognitively modelling relationships in a social group requires the two-argument form; so, if language is about sharing cognitive models of social relationships, it must be able to express and comprehend two-argument forms; which means that the two-argument form, like the one-argument form, must have been prefigured in cognition before it was used communicatively. The important first steps of language grammar were not the creation of new cognitive systems, they were the reuse of pre-existing ones (Edwardes, 2014b).

Naming and Semanticity should not be underestimated, either. The capacity to label, or tag, objects must be present in cognitive social modelling before it is shared communicatively. Modelled objects must be tagged so the modelling mind can use the same object in different models and know it is the same thing. Inside the modelling mind, the tag represents a holistic object – an amalgam of physical attributes, emotional responses, associations with past events, and so on; and it is idiosyncratic – my tag for the object need not correspond to your tag for the same object. When social models are shared between minds, however, there must be negotiation toward common tagging, and tags need to become less holistic and more neutral. Simple nominal labels now become sufficient (and it can be argued, although it is outside the scope of this book, that the vocal channel is superior to the gestural channel for expressing simple nominal labels – *e.g.*, Napoli & Sutton-Spence, 2014 – and for communicating abstract meaning – *e.g.*, Zdrazilova *et al.*, 2018). The label, or noun, must be prefigured in cognition before it can be shared; but there must also be negotiation as part of the sharing to establish common labelling.

The relationship between objects must also be prefigured in cognition. if [A[uB]], the one-directional two-argument form, is the basic syntactic structure then, as well as nouns (A and B), there is a need for verbs (u) to establish the relationship between the nouns. Verbs represent an important cognitive trick that social cognition both requires and makes possible. If my model of Alf is accompanied by fear, and my model of Beth is accompanied by fear, how do I generate a model of Alf and Beth together that is not dominated by fear? To model an alliance between Alf and Beth I must be able to do this. I must see Beth through Alf's eyes and Alf through Beth's eyes to model the relationship between them; which means their relationship must be modelled in a different way from my own relationships with each of them. Adopting the viewpoint of another requires the relationships between others to be separate from my personal relationships with those others: the cognitive linkage of verbs must include emotional disinterest.

The negotiation toward common labelling, or tagging, required by shared Naming and Semanticity is also prefigured in cognition, but not by cognitive social modelling. Instead, it is an instantiation of the cognitive capacity for metaphor, particularly the conceptual metaphor of metaphor: Y, AS A PROXY FOR X, IS X. Tagging creates a shorthand label which cognitively stands for the object itself, such that any attribute of the object can be evoked by the tag. From a modern linguistic perspective, this process seems obvious: words represent objects in an arbitrary but agreed way, so the use of a word evokes a similar set of impressions in speakers sharing a language, even though their emotional reactions to those impressions may differ markedly. There is a *thingness* to words even when they are abstract and physically unbounded.

Cognitively, the dislocation needed to treat Alf both with subconscious fear and with conscious disinterest means my model of Alf is abstracted away from my personal reaction. All the representations I have of Alf are associated with a single “Alfness” which is completely arbitrary; and the arbitrary tag *Alf* stands for my whole knowledge of the Alf-object in ways that my personal Relationship-A viewpoint cannot. In the conceptual metaphor of metaphor, the tag Y is a proxy for the object X because it is dislocated from my personal Relationship-A viewpoint of Alf. It is easy to see how disinterested, dislocated tagging could have become useful in disinterested, dislocated negotiation toward meaning.

So, for communication grammar we need to segment utterances into separable meaning-units which, as tags, are interchangeable within the communicable A-Relationship-B form. There are also different meaning types in the structure – at minimum, objects or nouns and relationships or verbs. Finally, we need hierarchy, allowing different parts of the structure to govern other parts: in the one-directional two-argument form we need to describe one object acting upon another. What does not seem to be present at the origins of language grammar, contrary to the prediction of Hauser *et al*. (2002), is recursion.

Moving on from Beginnings

The requirements at the beginnings of grammar are: an understanding of nouns and verbs; a capacity to model disinterested, one-directional, two-argument forms; and a capacity to negotiate with others toward shared meaning. Together, these mechanisms let us share with others our models of the social relationships within our group. They existed as cognitive capacities before they were pressed into service in social model communication; they did not emerge just to enable language.

An aspect of modern language absent from this list is linguistic complexity. That, however, is unremarkable: adverbs, adjectives, multi-argument forms, subordination, iteration and recursion can all be explained as later increments in the development of modern language. There is, however, another aspect that we tend to see as basic to language, but which is not part of the origins picture painted so far.

Selfhood, naming, and the use of persons or pronouns (*me*, *you*, *they*) would seem to be a minimal requirement for linguistic communication: when exchanging social models, the ability to tag, or name, the speaker and listener and those who are neither seems basic. Yet there is no need in non-communicative cognition for recognition of the performers in a speech act: unshared social modelling consists of A-Relationship-B constructs with all the A and B objects as third persons identified by personal tags. The non-communicative mind does not need the special object *you* because there is no external listener, and it does not need negotiated tags, or names, to identify the objects. There is only one voice, so the special object *me* does not need to be identified, either: there is no point in viewing the only voice as *my* voice, or anything other than *the* voice. Cognitive social modelling requires models of others, so it must consciously recognise the border between self and other; but the self itself does not need to be consciously modelled.

The communicative mind, in contrast, needs to recognise that the *you* receiving my models and the *I* providing them are special, not least because *you* can become the speaker and *I* the listener. Awareness of *you* means awareness that the third persons in my offered models are sometimes second persons – I am offering the receiver information about themself; and it means being aware that the models being offered to me are being offered intentionally. The role of *you* becomes significant in both what I say and what I hear.

Being aware of *I* and *me* has far-reaching effects. First, my voice is no longer the only voice, it is one among many and, although usually the most important voice for me, there are times when it isn't. Second, others may offer me their models of me; so, to interpret those models, I must make my own model of myself, treating myself as a third party like all my other cognitive models. As we have seen, however, an important feature of social modelling is treating modelled entities dispassionately, when treating my self dispassionately seems contrary to good evolutionary sense. What is the advantage in treating myself objectively when everyone else is passionately self-interested?

Self-modelling seems an unwelcome side-effect of sharing cognitive models, yet it is also a significant marker of humanity; how could this have happened? Self-modelling only has value in an environment where social models are being shared, so awareness of self is ultimately an outcome, not a source, of sharing. However, the main reason for exchanging social models is to assess the reputations of other group members; and, in a social environment driven by reputation, sharing models is a way to demonstrate honesty, informativeness and cooperation, enhancing the sharer’s reputation and therefore their fitness. When associated with awareness of self, self-disinterest allows me to model the effect of my actions on my reputation, letting me manage my own reputation.

Developing sufficiently language-like complex communication involves several steps. Initially, vigilant sharing, altruistic punishment and reverse dominance generated a society in which sharing social models was a viable fitness strategy; and sharing social models led to awareness of self (the discovery that others are modelling me), and awareness of my own reputation (the discovery that others are judging me). This, in turn, led to disinterested egalitarian self-awareness (the view that I am worth no less, but no more, than any other group member). Finally, awareness of my own reputation and disinterested egalitarian self-awareness together allowed complex social modelling, and therefore complex language, to develop.

Reputation, Iterative Hierarchy and Complexity

When exchanging social models became a fit strategy, the sharing of one-dimensional two-argument forms created its own evolutionary pressures. Individuals proficient at sharing two-argument models became fitter than those less proficient, so the species evolved toward more effective signalling of two-argument models. Odling-Smee & Laland (2009) include this in niche construction: grammatical communication generates its own enhancements, driving the signalling environment toward further complexity by its very existence. For instance, individuals who can tag received A-Relationship-B models with the sender's identity do better than those who cannot. Knowing the source of a particular model allows the sender's reputation to be factored into the bare A-Relationship-B information, enriching the receiver's social calculus. The supplied information is no longer just about A and B and their relationship, it is now also about the sender, C, and C's relationships with A and B; and it is about the receiver themself, and how the supplied information fits with their current models of A, B and C. The offered model becomes valuable whether it is accurate or not. What makes gossip valuable is not just what is said but who is saying it, and why.

Attaching a sender-tag to a shared model gives the form [C[A[uB]]] (*Alf-Relationship-Beth-by-Gemma* – or, in English, *Gemma said that Alf likes Beth*); but understanding this construct means the receiver must be able to interpret a multidimensional social web involving several different relationship forms. First are the receiver's own one-argument Relationship-A models with Alf, Beth and Gemma; second is the receiver's own model of the A-Relationship-B between Alf and Beth; third is Gemma's shared model of the A-Relationship-B between Alf and Beth; and fourth is the receiver's triadic model of the [A-Relationship-B]-by-C between Alf, Beth and Gemma.

The triadic [A-Relationship-B]-by-C structure is hierarchical, so has the potential to become iterative (*e.g.*, *Del said that Gemma said that Alf likes Beth*), but the need to compute this arises only if the triadic cognitive modelling becomes expressible communicatively: Del must tell me that Gemma said that Alf likes Beth before I can know that Del said that Gemma said that Alf likes Beth; and I must be able to model the four-node construct before I can share it with others. The tetradic [[A-Relationship-B]-by-C]-by-

D construct must be modelled in my social calculus as [D[C[A[uB]]]] before I can share it.

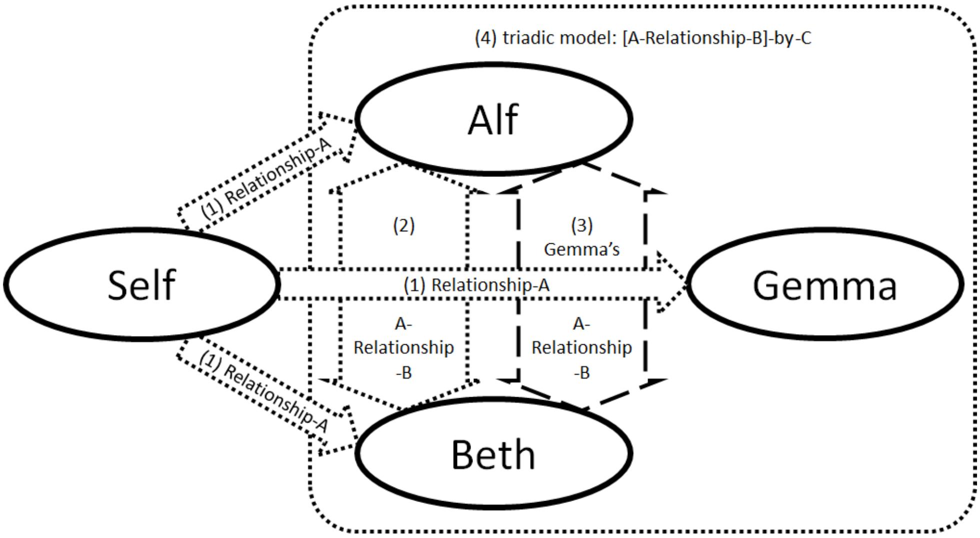

Figure 9.1: The receiver's (self's) multidimensional cognitive social modelling

This iterative hierarchy is theoretically infinitely extensible, but in practice it is quite limited. Each new level dilutes the signal's reliability because it introduces a new reputation to be assessed. If I trust everyone in my group 90% of the time then Gemma telling me that Alf likes Beth is 90% reliable. But Del telling me that Gemma says that Alf likes Beth is only 81% reliable, and Evan telling me that Del says that Gemma says that Alf likes Beth is 73% reliable. By the time we get to Iona, or Iona says that Theda says that Rita says that Zeta says that Evan says that Del says that Gemma says that Alf likes Beth, or [I[T[R[Z[E[D[C[A[uB]]]]]]]], the information is less than 50% reliable. This may be why, while some people seem able to map nine nodes in their social calculus (that is, handle up to seven levels of reputation), most are more limited, with seven nodes (five levels of reputation) being the average (Miller, 1956). However, all but a small minority of humans can keep track of five nodes (three levels of reputation): *Evan says that Del says that Gemma says that …*

How the receiver benefits from iterated labelling seem obvious: it provides information on the reliability of the message, making cognitive social modelling more subtle; and it makes the receiver aware of the sender's sources of knowledge. Why the sender should provide this information is, however, more difficult to explain. Reciprocity may be a factor, you tell me yours and I'll tell you mine; but the main value to the sender is that it safeguards their reputation for honesty. If social information is valuable, and I become known as an honest source of social information, my reputation improves: there is a fitness pressure for sharing reliable information. Alternatively, my reputation suffers if I offer inaccurate third-party models as my own; so a mechanism which allows me to offer the information, but distances me from its accuracy, also benefits me.

Other factors generate complexity in language. For instance, if complexity indicates fitness, it can itself become a costly signal. Using complex grammar can indicate proficiency in language; or it can indicate proficiency in complexity itself, especially when the complexity is in the form and not the meaning of the utterance. As a costly signal we would expect males to use complex utterances more often than females, to impress potential mates and overawe potential rivals; and, statistically, this does seem to be the case (Tannen, 1994; Cameron, 1998a). However, as Cameron (1998b) points out, while males seem to use more linguistic complexity, females can usually understand it; they just use it less. Using language complexity to establish dominance is also subject to reverse dominance: the US Navy KISS doctrine (Keep It Simple, Stupid) is a recent example of a continuing fight against language complexity. Cicero's statement, "Brevitas optima commendatio sermonis" (Brevity is the best recommendation of speech), shows this to be an ancient battle.

Complexity can also differentiate between in-group and out-group individuals (Nettle, 1999). If language is a measure of group membership then the capacity to handle local-language complexity is a badge of group membership. An example of this is the slaughter at the fords of the Jordan (The Bible, Book of Judges, 12:4-7). After losing a battle, the Ephraimites tried to cross the River Jordan back to their homeland; but the Gileadites seized the fords, demanding that everyone who wanted to cross say the word *shibboleth*. The Ephraimite language lacked a /sh/ sound, so the Gileadites slaughtered everyone who said *sibboleth*. There is some evidence that the word *Scheveningen* was used in a similar way by the Dutch in 1940 to identify German infiltrators (Goodman *et al.*, 2023).

However, Darwinian pressures toward language complexity rely on a pre-existing language system: they can exapt an existing language system for signal fitness, but they cannot generate a new language system. Language can be an effective Darwinian fitness mechanism, but its original evolutionary function is to enhance socialisation and cooperation, not to signal fitness; and this may be why the animals with communication systems closest to human language in terms of complexity are eusocial insects (*e.g.*, Gould & Gould, 1995; Sudd & Franks, 1987), and not songbirds (Searcy & Nowicki, 2008).

Utterance from Pre-grammar to Complexity

There seems to be a ratchet effect between cognitive social modelling and model sharing. At first, only one-argument Relationship-A grammar ([uA]) was needed, to model the conditionally cooperative social interactions involved in Joint Ventures. However, enhanced socialisation meant that knowing the relationships between other individuals became useful; and this could be determined simply by observing the reactions of individuals to each other. Modelling those relationships was even more useful, but it required the development of a two-argument calculus (A-Relationship-B or [A[uB]]). This calculus developed as humans became more socialised, moving from Machiavellian Intelligence through Vigilant Sharing, Political Singularity and Reverse Dominance. Close cooperation became a fit strategy, and sharing the two-argument calculus became a fitter strategy than not sharing, leading to a realisation of selfhood. Sharing meant that tagging two-argument forms with the sharer's identity became a fit strategy, making three-argument calculus ([A-Relationship-B]-by-C or [C[A[uB]]]) necessary. Tagging two-argument forms with their source also allowed the receiver to build reputations into their models of individuals. The cooperative human signalling environment then made it a fit strategy to share these three-argument forms; and this, in turn, introduced an iterative capacity into human social modelling.

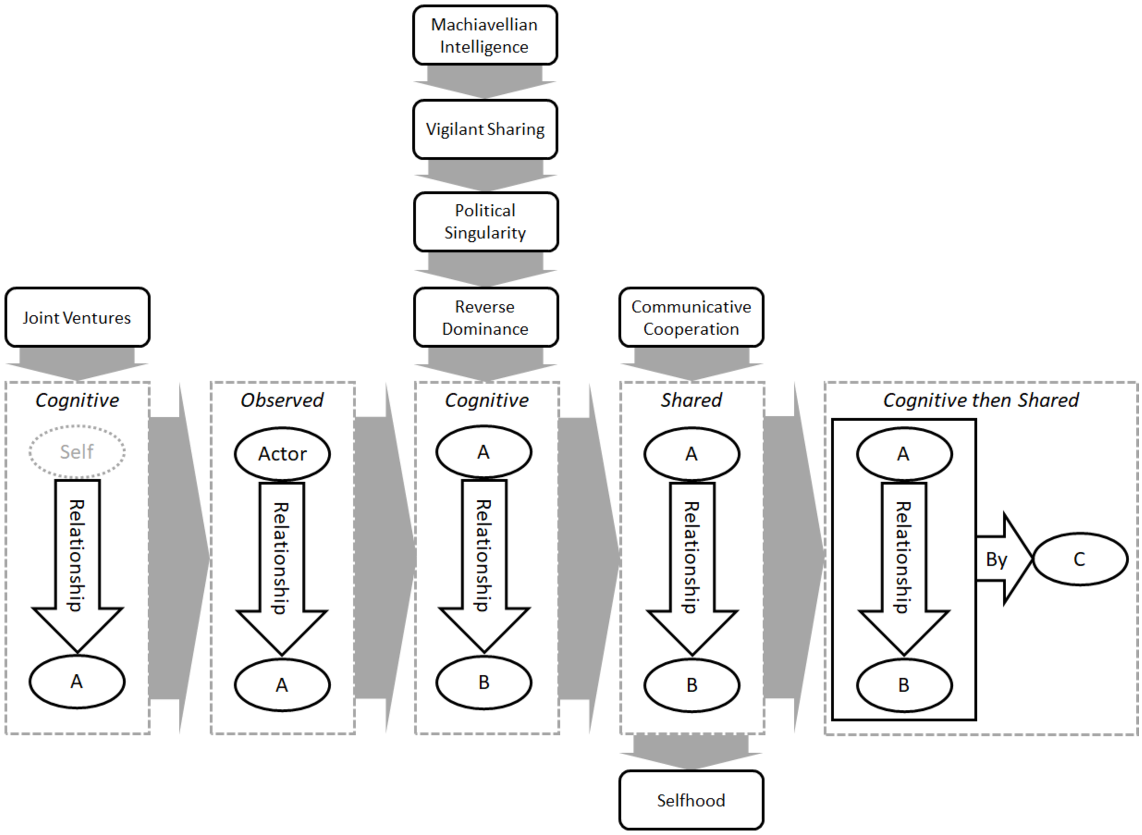

Figure 9.2: From Pre-Grammar to Complexity

This iterative capacity is not the infinite recursion of Hauser *et al.* (2002), it is a limited side-effect of sharing three-argument social calculus. The language structure used to share this calculus is indeed theoretically infinite, but it only needs to work with a limited number of iterations. Dunbar (2004, ch3) shows that five levels of iteration, a seven-argument form, is the comprehension limit for most people.

For Hauser *et al.*, recursion is the minimum requirement for a communication system to be called language; in which case, everything described here is pre-language or protolanguage. If, on the other hand, language is defined as the manipulation of symbols then two-argument communicative social modelling can be called it language. The point at which language begins depends on the definition of language used, and the definition used here (formal grammatical utterances) produces its own point of origin (the first formal grammatical utterance). This definition is no better, but no worse, than others, and it does provide a reasonable explanation for the sources of language grammar.

However, while this explanation of the developmental path from pre-grammar to complex grammar is reasonable, it is only one of several possible paths; and it leaves some questions only partially answered: for instance, what makes exchanging social models a fit strategy, how does a language propagate through a community, and how does negotiation toward meaning happen? Language complexity has been addressed somewhat abstractly and discussed only in terms of nouns and verbs. Multi-argument structures have been addressed, but only in linear form; and little has been said about the truly complex mechanisms used in modern language. This developmental path does, however, provide a plausible route from pre-grammatical communication to grammatical human language.

10 Language Grammar and Nonhumans

> ***During long or difficult sessions, Alex may request students to "go away"; or he occasionally states "I'm gonna go away," climbs off his training chair, and tries to leave. After such an interaction, attempts to continue training are usually fruitless.***
> *(Irene M. Pepperberg)*

A common question in language origins research is, can nonhumans use human language? This question, however, makes the assumption that language is a monolithic thing; and the assumption in this book is that it is not. So this chapter will look at a slightly different set of questions: which nonhumans can use aspects of human language, what are the aspects they can use, how do they use them, why can they use them, and what does this tell us about the sources of language grammar?

The first question gives us a wide range of species: among others, some chimpanzees and bonobos have been taught to use a sign language, both with humans and between themselves, and understand and use syntactic structure (Fouts with Mills, 1997; Savage-Rumbaugh & Lewin, 1994); Alex, a grey parrot, recognised number symbols, did simple arithmetic, and understood adjectival properties (Pepperberg, 1999); and dolphins parsed and responded to multi-argument sentences (Herman & Uyeyama, 1999). The purpose behind these experiments was to prove the ordinariness or uniqueness of human language as a communication system: can other animals learn human language, thus proving it is not solely human; or are animal signalling and human language irreconcilably different? While *Homo sapiens* is just another type of animal it has, like every animal, certain unique characteristics; and studying only those characteristics leads to the conclusion that language is “a complex of capacities that seem to have crystallised fairly recently … that sets humans apart rather sharply from other animals, including other hominids, judging by traces they have left in the archaeological record” (Chomsky, 2005, 3).

Chomsky seems to equate grammar with language, such that language is grammar is language: the essential *cognitiveness* of grammar cannot be differentiated from the essential *communicativeness* of language. Reboul takes a similar view on this collocation, treating language as the container within which, and for which, grammar appeared: “That language is eccentric among animal communication systems cannot be seriously disputed. It has a core combination of features – semanticity, discrete infinity, and decoupling – that is found nowhere else in nature to our present knowledge.” (2015, 2).

Reboul’s first sentence sums up the dilemma: language is a communication system; compared with other animal communication it is eccentric, but it is a communication system nonetheless. However, Reboul’s second sentence indicates a way out of the dilemma, although she does not explore it: semanticity, discrete infinity, and decoupling could all be separate from language, part of a purely cognitive grammar system. Here, we explore language as precisely what the Generativists say it isn’t: a communicative system which lets us share the cognitive structures used in social modelling. This approach insists on continuity with the rest of nature, not exclusivity: the cognitive grammar needed for social modelling is also needed to communicate those models. Language grammar is, however, not just social modelling; other mechanisms, like holistic, grammar-free utterances (*e.g.*, *yes*, *no*, *soon*) show that language grammar is not completely explained by formal grammar systems. Dingemanse *et al*. (2013) found that communicative repair initiation, an important language function, is represented in many languages by the questioning grunt, *huh?* To discuss the universality of the word *huh?* with linguists, Dingemanse *et al*. first had to argue that it really is a word inside language, and not a non-linguistic sound outside language. This is explored in more detail in chapter 13.

Evolution is the development of a species toward its environmental niche through the reproductive success of the individuals most fitted for the niche. Humans, however, are also good at engineering unpromising environmental niches to meet their needs. This adaption of a niche to the individual, or niche construction, is a more complex way of matching species to environmental niche: individuals able to adapt an environment do so, making more space for their offspring; and the species develops toward niche construction through the reproductive success of niche constructors. In humans, cooperation (Fuentes *et al*., 2010) and social intelligence (Sterelny, 2007) are viable niche construction tools as well as important components in socialisation and communication. So, the establishment of language among humans could be the result of the evolutionary fitness it offered – either by making individuals better able to handle their environment, or by making the environment more friendly. If language is an outcome of human socialisation – because what was being communicated were the mechanisms of human socialisation – then its appearance in humans and its non-appearance in other animals are both unremarkable. Human socialisation is, by definition, species-specific; so its products, such as language, are therefore also likely to be species-specific.

However, if language is species-specific then it meets the communicative needs of humans only; which raises the question, why should other animals want to learn human language? What is remarkable about nonhumans using human language, therefore, is that it happens at all. Somehow nonhumans use our signalling system, however incompletely, to understand our intentions and convey information to us. Some, but not many, humans can also do this, producing sounds which are recognisable by nonhumans as conventional signals; and it may be that, in our past, this capacity was a widespread and important hunting skill. However, the usual reason why human hunters use animal signals is to distract an animal or deceive them, not inform them (Lewis, 2009).

In our studies of nonhumans using human language, veracity is of most interest to us: if the animal responds to our utterance in a contextually predictable and coherent way we judge communication successful. What we expect from the animals in these experiments is therefore different from our expectations about humans using language. With humans, we treat the deceptions of metaphor, obliquity, hyperbole and fiction as acceptable signs of communicative success; and we also treat defiance, disengagement, or ignoring the utterance as valid responses, part of everyday discourse. In the animal experiments (although not in the social linguistic interactions with the same animals) linguistic deceptions by the animals were usually treated as communicative failures. Is this because the experiments were assessed scientifically, requiring provable coherence, rather than part of the informal negotiation toward meaning we use in our everyday communication; or does it represent an unspoken prejudice against animal minds being similar to ours? (Savage-Rumbaugh & Lewin, 1994, 256-258).

Can Animals Share Social Modelling?

The theme of this book is that humans need language grammar to exchange social models. The cognition behind this, treating social groups as a series of interlinked bipartisan relationships, may not be exclusively human, it may be a general tool of Machiavellian manipulation; but communicating social calculus is rare – so far, we have identified only one species doing it habitually (Donald, 2001). While communication is vital for social networking, there is only anecdotal evidence of nonhuman social model communication, either as a natural signal or by those trained in human language. However, we have not tested for it, simply because we have no way of asking the equivalent of “what do you think of Beth?”, or “what does Alf think of Beth?”.

While we do not yet know whether other animals are exchanging social models, the social systems of chimpanzees and bonobos support the idea that they are capable of Relationship-A cognition, modelling the relationship between their unmodelled self and another individual (Tomasello & Call, 1997, 338-341). Some who have been acculturated to human society also seem to cognitively model the relationships between individuals (Premack & Premack, 1983, ch3). So social modelling involving both Relationship-A and A-Relationship-B forms seems within the cognitive capacity of our closest primate relatives, even if they do not have communicative systems for exchanging models.

In a Machiavellian social environment, being able to manipulate others through their friendships and rivalries offers a fitness advantage. Modelling relationships between others lets me tailor my approach to them, helping me build coherent alliances and establish or enhance my position in the social hierarchy. Modelling social information lets me enhance my social knowledge and reduce costly confrontation; but giving it away reduces its exclusivity and therefore its fitness advantage for the giver. Giving away social information is not an evolutionarily fit strategy – and it seems that only humans actively do so: the capacity to model social relationships is a product of Machiavellianism, which precludes the sharing of those models.

If we look for sources of language grammar in nonhumans, therefore, we must look for social modelling in cognition rather than in communication. Fortunately, considerable work has been done in this area recently, and we are beginning to get a clear picture of social modelling in our closest primate relatives.

For instance, while symbolic communication is not obvious in the natural communicative repertoires of other Great Apes, they nonetheless show symbolic competences in laboratory studies; these competences must therefore be cognitively available even if they are not communicatively apparent. Cissewski & Luncz (2021) look at cooperative nest-building by chimpanzees of the Taï South group for indicators of arbitrariness (a gesture has no obvious connection with its intended meaning) and conventionalisation (a gesture is socially learned and not genetically instantiated), and they show that the two reasons for nest-building – playing or mating – mean that the invitation to play and the invitation to mate essentially have the same ambiguous meaning. From this, they show that gestural symbolic competence may have arisen in two ways: a gesture can change meaning, creating a semantic shift; or a meaning can be attributed to a new gesture, creating new semanticity. At least one of these is at work in the nest-building of the Taï South group.

In another study, Hobaiter *et al.* (2022) show that, like human language utterances, many gestures made by nonhuman Great Apes are intentional. However, while first-order intentionality (knowing that others have separate minds) has been identified, evidence for second-order intentionality (knowing those minds have different knowledge) remains uncertain: individual gestures do share some characteristics with human language words, but fully wordlike gestures have not been conclusively identified. Hobaiter *et al.* suggest we are approaching problem wrongly:

> It may seem like a theoretical stretch at times to ask – how is an ape gesture like a human word, not in its shape or structure, but in its use? But we suggest that this is the approach required to move forward in asking the questions that are key to understanding why human language emerged. Rather than, 'What were the first words like?', we suggest asking, 'What were they used for?' (*ibid.*, 6).

Vocalisation is not the only way to produce sound: many species use non-vocal acoustic methods for long-distance communication (Eleuteri *et al.*, 2022). Wild chimpanzees drum on the buttress roots of trees, generating low-frequency sounds that can travel over a kilometre. This drumming seems to encode the drummer's identity in signature beat patterns, so they can act as "I am here" signals, with selfness symbolically coded in. Drumming sessions were more frequent in small groups, with frequency dropping as group size increased, indicating some flexibility about when to drum. Drumming skill also seems to vary between chimpanzees (Dufour *et al.*, 2017): one male, Barney, part of a group at the Biomedical Primate Research Centre at Rijswijk, the Netherlands, gave a performance that was rhythmical, decontextualised and well-controlled, and not equalled by other chimpanzee drummers.

Finally, Roberts & Roberts (2018) show that chimpanzees negotiate their social networks with both gestural and vocal signalling, and larger social networks signalled more frequently than smaller networks. It seems that, throughout hominin evolution, larger group sizes are supported by increases in both vocalisation and gesture; although vocalisation, capable of addressing larger audiences, tends to dominate signalling.

Can Animals Know Themselves?

Above the entrance to the Temple of the Oracle at Delphi was carved *know thyself*. The Oracle's pronouncements, however, were noted for their treacherous semantics, so this has several meanings: know you have existence, know you are a person, know the person you are, know yourself as others see you, and know your social role, among others. The need for these different ways of being self-aware varies between species, affecting both awareness of self and awareness of other. As Budiansky (1998) says:

> There is a certain flavor of anthropocentric bias in the very hunt for self-awareness in other animals, a hint that conscious self-awareness is the best thing evolution has yet to produce – and we want to know how animals stack up against this standard of ultimate perfection. Yet it is no insult to animals that they might do what they do without self-awareness as we understand it; nor is it a particular compliment to animals to see how closely they share our peculiar cognitive abilities. (162)

So, when we look at how nonhumans show – or do not show – awareness of self and awareness of other, we should be clear about what we are testing, and why.

One of the earliest tests for self-recognition was Gallup's mirror test (1970). He found that some, but not all, chimpanzees recognise their image in a mirror as being themself, while monkeys cannot do this. The experiment involved first acclimatising a subject to a reflective surface; then anaesthetising the subject and placing a water-based mark on their face in a position they could not see directly. The test was whether, upon noticing the mark in the mirror, the subject touched the mirror or the mark on their face, using the mirror to find the mark. While several chimpanzees checked out the facial blemish by touching their face, monkeys either ignored the mirror image or treated it as another monkey. Recognising the image in the mirror as their physical self seems to be a capacity available to some chimpanzees and most humans, but not other animals.

Further experiments, however, have clouded the issue. Bottle-nosed dolphins (*Tursiops truncatus*) pass the mirror test (Reiss & Marino, 2001); one Asian elephant, Happy, has also passed (Plotnik *et al.*, 2006); and even some monkeys (*Cebus apella*), excluded from self-recognition by Gallup, have shown mirror-awareness, at least in terms of other objects in the room (de Waal *et al.*, 2005). Pigs have also shown they recognise mirrors as reflectors of the real world (Broom *et al.*, 2009); and now, even mice (Yokose *et al.*, 2024). Perhaps most intriguing is evidence for mirror recognition in magpies (*Pica pica*), which are as reliable as chimpanzees in passing the mark test (Prior *et al*., 2008). Recognising the image in a mirror as representing the physical self – what Povinelli (2000, 328-337) calls the kinaesthetic self – seems to be widespread in nature; but if physical self-recognition extends widely beyond the *Pan-Homo* family, it cannot be an indicator of human uniqueness.

Instead of differences, perhaps we should consider cognitive similarities between humans and other species. For instance, Locke (2021) finds that what and how we communicate are not that different from other primates. Like us, other primates use both direct signals (intentional information) and indirect signals, or cues (incidental, often subliminal, information). Direct signals display the intention of the signaller, while cues provide a communicative background supporting interpersonal relationship-building, needed by a species with complex social systems. Gestures and pragmatic signs like facial expression, touching, and loudness initiate both interpersonal activities like grooming and cooperative ventures like hunting, managing predators, and challenging rival groups. Primate communication cannot involve only informative signals, it must support social interactions and negotiations toward joint actions.

Primates share many complex interpersonal emotions with humans; but De Waal (2011a) shows that animal emotions are often sidelined in comparative biology, with interactions between individuals being described and analysed without reference to their emotions. He proposes that animal emotions should be studied as the causes and outcomes of social interactions: from unshared subliminal arousal comes either excitement, creating desires and hopes, or apprehension, creating fear and anxiety. The emotions that excitement creates are mostly internal rewards, and some also act as subliminal cues for others: for instance, self-pride and affiliative affection are both sources of personal physical pleasure and conspicuous confidence.

Apprehensive emotions can also be cues, but they often directly generate other effects: shame, anger, guilt, grief, and jealousy all generate sadness and pain. De Waal challenges animal behaviourists to "move from … 'ill-defined categories' to replicable, objective methods to document the emotional deep structure of behavior." (203).

Nieuwberg *et al.* (2021) looked at emotional contagion, where one individual adopts the same emotion as another to indicate solidarity, and cognitive empathy, where an individual recognises the emotional relationships between others as different from their own relationships with those others – labelled A-Relationship-B cognition here. They show that aspects of cognitive empathy are common in primates, indicating that emotional semantics probably developed early in the primate clade. In contrast, emotional contagion seems to occur unevenly: it is present in highly socialised species, but absent in less socialised species, indicating it is produced by interactive socialisation and not cognitive social modelling. Looking at communication of emotion instead of emotion recognition, Parr *et al.* (2005) implicate neurological features (spindle cells and mirror neurons) in the evolution of emotional awareness or empathy, suggesting it appeared late among primates – perhaps only with the *Homo* clade.

These two accounts are not contradictory: intuitive reaction does not need to be a conscious cognitive activity, it can occur subliminally; while emotional awareness, by definition, requires awareness or consciousness. However, the two accounts do indicate that empathy needs to be more closely defined: if it can be both subliminal and conscious, then maybe it is not one thing but two. Méndez *et al.* (2022) support this, showing that intuitive reaction is a subliminal process involving a deep neural network based on the superior colliculus in the midbrain, and not involving the cortex; and Kret *et al.* (2020) show that Great Apes do indeed seem to use both involuntary intuitive reaction and voluntary emotional expression, and may well use the same mechanisms in emotion recognition.

Dielenberg (2013), however, identifies one area of cognition where humans do seem to differ from other Great Apes. He sees other apes as having, like humans, a *first order awareness of the unseen world*: they can infer causation into environmental changes and seek out that causation. *Second order awareness of the unseen world*, inferring properties into things based on the effects they have on the environment, is limited to the *Homo* clade. For humans, environmental change is not just explicable after the event, it is predictable based on the properties of things in the environment. Dielenberg identifies this capacity to predict as part of the human capacity to visualise finished tools in raw rocks, potential relationships in current social interactions, future events caused by current events, and possibilities emerging from current actualities. He further argues that this second order awareness of the unseen (or conscious modelling) is a product of conscious self-recognition, making awareness of self a difference – perhaps the key difference – between humans and other Great Apes.

Empathy, Theory of Mind and False Beliefs

The relationship between empathy, subliminal and conscious, and Theory of Mind is explored by Seyfarth & Cheney (2013). They refer to the two types of empathy as reflexive (subliminal) and reflective (conscious). Before full Theory of Mind, empathy is reflexive: heavily reliant on instinct to predict the behaviours of others. However, understanding the minds of others enhances reflexive empathy, and understanding the emotions of others enhances Theory of Mind: there is a ratchet effect between reflexive empathy and Theory of Mind. Full Theory of Mind becomes possible when conscious, or reflective, empathy begins: reflective empathy involves consciously modelling the emotional states of others, and full Theory of Mind involves consciously modelling their cognitive states. In a study of over 3,000 conflict interactions in 44 chimpanzees, Webb *et al.* (2017) found evidence for a spectrum of reflexive empathy. They identified a range of empathic behaviours in the group, but they also found a positive correlation in individuals between the complexity of their consolation behaviour and the sophistication of their social integration.

In a review of the seminal paper on Theory of Mind (Premack & Woodruff, 1978), Call & Tomasello (2008) describe Theory of Mind, like empathy, as being of two types: understanding the perception, knowledge, goals and intentions of others, which chimpanzees seem able to do; and realising that the perception and knowledge of others may be wrong, that their goals and intentions can be based on false beliefs. There is no good evidence that chimpanzees identify false beliefs in others, so Call & Tomasello conclude that:

> In a broad construal of the phrase 'theory of mind', then, the answer to Premack and Woodruff's pregnant question of 30 years ago is a definite yes, chimpanzees do have a theory of mind. But chimpanzees probably do not understand others in terms of a fully human-like belief-desire psychology in which they appreciate that others have mental representations of the world that drive their actions even when those do not correspond to reality. And so in a more narrow definition of theory of mind as an understanding of false beliefs, the answer to Premack and Woodruff's question might be no, they do not. Why chimpanzees do not seem to understand false beliefs in particular – or if there might be some situations in which they do understand false beliefs – are topics of ongoing research. (*ibid.*, 191).

Recent research (Barone *et al.*, 2022) indicates that, in human children, there seem to be stages and nuances in the way they understand the false beliefs of others, and in the way they exploit those false beliefs. As they grow, human children develop a theory of how others think, interpreting and manipulating the beliefs of others with increasing sophistication; and this capacity seems to include understanding of both conscious and subliminal false belief systems. Human children develop a conscious understanding of the false beliefs of others and factor this into their planning and modelling, while chimpanzees do not have the belief-desire psychology to make the shift from subconscious knowledge to conscious understanding.

Empathy is sometimes treated as a subset of Theory of Mind, but they are two different ways of cognitively modelling others. Empathy is not just about understanding the needs of others, there must be an attempt to meet those needs; and this differentiates it from Theory of Mind, which can operate in a Machiavellian environment in which the needs of others are exploited, rather than accommodated. Premack & Premack (1983, ch3) showed that chimpanzees can model the probable future actions of others based on their previous actions. However, chimpanzees do not seem able to model another individual modelling the beliefs of a third individual –

something which humans do regularly, with sentences like "Alf thinks Beth is unhappy". Dunbar (2004, ch3) shows that, in terms of Theory of Mind, chimpanzees can impute motivation to others, but not to impute motivations about motivations. Most humans can work at about five levels of motivation (which Dunbar describes as "A believes that B thinks that C wants D to suppose that E imagines …", *ibid*., 48). This is a clear difference between the mental modelling of humans and other primates. However, Tomasello (2008, 342-345) shows that this advanced Theory of Mind is not sufficient to get us to language by itself; sharing our models of the intentions of others is vital. It is not our intelligence that demands language but our cooperative sophistication; and empathy is the system that powers our cooperation.

Read *et al*. (2022) identify another significant cognitive difference between chimpanzees and modern humans: while human working memory allows most individuals to simultaneously retain between five and nine items (Miller, 1956), the working memory of chimpanzees seems to be limited to between one and three items. The human working memory capacity maps to Dunbar's (2004, ch3) explanation for the human capacity for recursion, a key feature of human language. Where Hauser *et al*. (2002) discuss a theoretical capacity in human language for "infinite recursion" – thus ignoring the fact that the cognitive complexity cost of each level of recursion is not zero – Dunbar makes the more modest claim that humans can achieve between five and nine levels of recursive cognition. The fact that the chimpanzee working memory capacity (one to three items) does not even overlap with the human range (five to nine items) indicates that the quantitative difference is likely to reflect an important qualitative difference. To have the thought "Alf likes me" requires three levels of modelling: Alf's model of me within my model of Alf within my model of me. So it requires a capacity for three levels of recursion, holding three things in working memory simultaneously. If the average chimpanzee can hold only two things in memory then this A-Relationship-Me modelling is beyond the capability of most chimpanzees.

Different species in the mammalian clade use different empathic strategies, indicating that empathy is not a binary on-off trait (de Waal, 2011b); and it also seems that empathy in the human lineage has been subject to considerable evolutionary pressure. Yet there may be no all-or-nothing difference between humans and other primates, as many typically human empathic behaviours have also been recorded in nonhumans. This is especially true of animals who have learned to communicate with humans using versions of human language: it seems that exposure to human language may have highlighted or even enhanced their empathic skills. This may be because humanlike communicative behaviour allows us to see their other humanlike qualities; or it may be because their exposure to human culture has allowed them to express, rather than suppress, all the humanlike qualities they have. For instance, Kanzi, the bonobo taught to communicate with humans via a special keyboard, has established friendships with other primates in the Yerkes Primate Centre, and has been recorded requesting visits with his friends and taking gifts for them (Savage-Rumbaugh & Lewin, 1994, 155). Nonetheless, even with language-competent nonhumans like Kanzi, the distance between human empathy and that of other animals remains significant.

So how did humans move so far along the scale of empathy and cooperation, ending up in a quite different species-niche from that of our closest relatives, the chimpanzees and bonobos? This is an issue that any evolutionary theory of human origins must address: cognitive social modelling used empathically leads to greater cooperation, and greater cooperation leads to the advantages of enhanced communication, such as specialisation and social tolerance; but cooperation also leads to effective cheating by individuals who use cognitive social modelling for Machiavellian ends. How did we humans get past this and become the highly cooperative species we are?

Sober and Wilson (1999, ch4) offer one solution: a species does not evolve toward species-benefitting behaviours, it evolves toward behaviours that benefit individuals. A sharing behaviour can benefit an individual if it reduces the individual’s stress from living in a socially complex group. Maintaining a cognitive network of interpersonal relationships is costly; but sharing behaviours make encounters with others less confrontational, and so emotionally less costly. If enough individuals adopt altruistic behaviours then social reciprocity develops, letting individuals share personal surpluses today to encourage others to share tomorrow – a social contract described by

Trivers (1971) as *reciprocal altruism*. This ensures that troughs of foraging and health are mitigated by others' highs, reducing individual stress and stabilising the group's size and fitness. Individuals are disadvantaged by failure to establish an interpersonal support network, and cheating is counter-productive: cheating the contractual exchange mechanism works for only a short while – support for the cheat quickly dries up. In this environment the evolutionary pressure is toward cooperation, and toward enhanced empathy to anticipate others' needs.

Enhanced empathy generates two evolutionary pressures toward cooperation: the individual benefits from reduced stress, and the group benefits from consensus, which also benefits the individual. Empathy reduces in-group confrontations, allowing larger, more concentrated, and more cooperative groups with less stressful interrelationships. Conforming to the *group morality* (essentially, to obey the golden rule, "do as you would be done by") becomes subject to evolutionary pressure; and group-driven altruistic punishment of uncooperative individuals becomes possible (Singer *et al.*, 2006). Empathy creates a tyranny of the cooperative as well as individual cooperation.

Can Animals Show Empathy?

An important difference between humans and other apes is the tolerance shown to conspecifics: humans have greater interest in, capacity for and skill at exploring the emotions and intentions of others. Hrdy (2009) describes how chimpanzee and macaque infants, raised by human carers after being rejected by their birth mothers, initially imitated their carers just like human infants. However, where humans continue to imitate others throughout their childhood, and possibly their life, the chimpanzees ceased to do so after about eleven weeks, and macaques after a mere seven days. The willingness to copy the activities of others shows an interest in both the physical acts they are performing and the reasons why they are performing them; it give the infant access to both the physical and the cognitive lives of those others. The early suppression of this curiosity among chimpanzees and macaques indicates an important difference in socialisation. Tomasello *et al.* (1993) list three types of learning: first is imitative learning, which provides the base for the other two, instructed learning and collaborative learning. Together they make up cultural learning, which is unique to humans – other primates do not retain imitative learning, so have no base on which to build other types of learning.

Fossey (1983, 70-71, 218-219) describes how some male gorillas use infanticide to remove a rival's offspring and bring females back into oestrus; and this infanticide is not accompanied by any apparent mourning by the mothers. It makes sense in a male-dominated gorilla group for a mother to discontinue investment in offspring that cannot earn her the protection of the dominant silverback, and for a silverback to encourage this transfer of a female's investment to ensure that, during his reign, his reproduction is maximal. Empathy in this circumstance is not evolutionarily fit. However, while infanticide by incoming males is not rare, it is also far from universal; it appears to be used only against unweaned infants who are interfering with the fertility of their mothers (Packer & Pusey, 1983).

Cheney & Seyfarth (1990, 235-236) state that the few descriptions of empathy in the actions of monkeys are likely to have other explanations. Monkeys just do not seem to exhibit compassionate traits such as care for the elderly, the sick, the bereaved or the defeated. While they do care for infants, it seems to be largely innate caring, and they do not adjust their caring regimes if their offspring becomes sick or distressed. When grooming others, monkeys tend to treat wounds as points of interest and probe them without consideration for the wounded animal. There seems to be no compassion in the actions of monkeys, and any cooperative behaviours seem to be products of innate mechanisms rather than conscious cognition.

This does not mean that empathy is missing from other primates, just that it has a different quality. Warneken & Tomasello (2006) conducted tests on young chimpanzees and human infants to assess their willingness to help others. The tests were on an increasing range of difficulty to assess the capacities of the subjects to mentally model the needs of the person needing help. Both children and young chimpanzees proved willing to help, but their actual support, and their understanding of the need for assistance, differed markedly. The human infants helped in more circumstances, provided help that better modelled the needs of the person needing assistance, and helped regardless of whether they knew the person needing assistance. Chimpanzees were less likely to help, less effective at helping, and tended to help only individuals they knew.

Yet there are some examples of nonhuman empathy which are eerily humanlike. De Waal & Lanting (1997, 154-160) discuss a strange proto-symbolic activity of bonobos which indicates a respect for the emotional states of others. Bonobos build night nests for sleeping; but they sometimes also build day nests, and the purpose of these nests seems to be mainly to establish personal space – a purpose which is respected by other bonobos. The nests are built to provide private feeding spaces or just to deter others from approaching; even close allies do not invade the sanctity of the day nest, and offspring beg at the edge of the nest for their mother's permission to enter. It has even been recorded that one male successfully used nest-building to deter an aggressive opponent (Fruth & Hohmann, 1993). It seems that there must be, on some cognitive level, recognition of, and respect for, the Garboesque message of day nest-building: "I want to be alone".

De Waal (1996, 148-150) also describes an experiment involving capuchin monkeys, in which monkeys in twin cages were selectively fed and allowed to share the food with another known capuchin. The unfed monkey had no direct access to the food, so relied on the fed capuchin to give them a share. The experimenters found that the capuchins shared readily if they had a pre-existing good relationship, but refused to share with their enemies or unknown monkeys. In the wild this would translate to assisting kin: capuchins mostly stay in small, closely related kin groups, so familiar monkeys are also kin. There does seem to be recognition here of the needs of others, and a willingness, albeit limited, to provide for those needs; but it is a partisan recognition which can be explained in terms of Neo-Darwinian kin selection, it does not require conscious choice.

From the examples of cooperation given above, it seems that altruism among nonhuman primates is considerably more constrained than among humans. Does this constrained altruism amount to empathy? It is hard to say for certain, but it is clear from the evidence that there is a significant difference between human altruism and the constrained cooperation of other primates.

How Intelligent Can Animals Be?

When making comparisons between humans and nonhumans there is a danger that a spurious difference is selected as a yardstick. For instance, early comparisons between humans and other Great Apes relied heavily on measuring intelligence, without attempting to define what intelligence is. The first formal measure of intelligence (Binet, 1903) treated it as a nebulous “cleverness”, often associated with a capacity for learning or reasoning or understanding or knowledge or high mental capacity, all of which are themselves underdefined. The idea of measuring intelligence, despite the lack of definition, became somewhat of a preoccupation in psychology, and Stern (1912) introduced the *Intelligence Quotient* as a comparative measure. We had no idea what we were measuring, but we weren’t going to let that stop us.

As a counter to the growing tyranny of IQ as a single measure of intelligence, Gardner (1983) proposed the theory of multiple intelligences. It is not possible to judge a human as evolutionarily more or less fit based on the standard IQ test because it tests only one fitness-relevant capacity, what Gardner called logical-mathematical intelligence. To this he added five other intelligences, fitness-relevant capacities which can be described in the wider sense of *things known*. These are linguistic, musical, spatial, bodily-kinaesthetic and personal (or self-) knowledge. Each intelligence in this suite marks a measurable potential difference between humans and other animals; and they could all, therefore, be involved in defining our species genetically.

Yet even with all these intelligences, the most significant difference between us and other animals, our level of socialisation, still seems to be missing. As early as 1927, Thorndike had used the term *social intelligence* to describe the fact that some academically gifted (officially high-IQ) students were nonetheless failing the social side of university life, because the Universities were wrongly equating logical-mathematical intelligence with life success. For this reason, Goleman (1995) added a seventh item, emotional intelligence, to Gardner's list.

At base, every intelligence relies on curiosity. Intelligences need the capacities to discriminate and retain knowledge, but there must also be a will and a capacity to acquire new knowledge; curiosity is the mechanism powering that acquisition (Kashdan & Silvia, 2014). In terms of social intelligence, curiosity about the cognition of others is represented by an interest in their states and relationships, and it works on three levels: curiosity about the physical states of others, their strengths, weaknesses and habits (kinaesthetic knowledge); curiosity about the mental states of others, what they are thinking and how this can be used to advantage (Machiavellian knowledge); and curiosity about the emotional states of others, what they are feeling (empathic knowledge).

Kinaesthetic and Machiavellian knowledge are capacities well-documented in other primates (*e.g.*, Arbib, 2005; Whiten & Byrne, 1988); but only humans seem able to use empathic knowledge unselfishly in their relationships. This empathy, however, is difficult to explain in evolutionary terms: to empathise we must "feel the pain" of others; and how does it advantage an individual to take on others' problems? It would seem more effective to concentrate on solving your own problems. In social mammals we often see consolation being offered to the loser after a confrontation (*e.g.*, Baan *et al.*, 2014), but this is more emotional calming than sharing the pain of the defeated individual. |In contrast, humans often model the emotional states of others and attempt to empathise. The change from passive sympathy to compassionate empathy means human social interactions are more complex than those of other social mammals, and our societies are correspondingly quite different. Could our unusual level of empathy, therefore, be an indicator of what makes us human? Jaeggi *et al.* (2010) think so, identifying the difference between humans and chimpanzees as the human capacity for greater social awareness and prosociality:

> While it is parsimonious to assume that human ancestors shared […] aspects of cooperation based on direct reciprocity with chimpanzees, humans seem to have acquired several derived features, some of which may have

> evolved convergently with other taxa. These derived features combined to make human cooperation more stable relative to chimpanzees. (2730)

Animals and Human Language

A key question in language origins studies is, can nonhumans use human language? However, while this is not a trivial question, it is not necessarily helpful when looking for the sources of language grammar. To use language, nonhumans must first understand that a sign is arbitrary, having meaning only through negotiation with others; and there is evidence that signs as arbitrary symbols are not limited to humans (Addessi *et al.*, 2007). The nonhumans must also share a signalling system complex enough to allow segmented signs to be coherently exchanged; and there is some evidence of nonhumans doing this, too (Ackerman *et al.*, 2014, 543). Third, the nonhumans must share a hierarchy of meaning, allowing separate signs with individual meanings to be combined into new signs with new meanings; and, once again, there is evidence of some nonhumans doing this (Pepperberg, 2005, 141). Finally, the nonhumans must understand that the combinatorial rules of shared communication are themselves arbitrary; and this has also been evidenced in some nonhumans, particularly Kanzi the bonobo (Schoenemann, 2022). Handling arbitrary combinatorial rules is of particular relevance to the sources of language grammar; but, as a mathematical-logical skill, its role in solitary puzzle-solving most likely preceded its role in shared communication. Nonetheless, the fact that any nonhuman uses combinatorial rule systems in their communication with humans is nonetheless sufficiently remarkable to merit further investigation.

Humans have long histories of commensal or cooperative relations with other species, and there is considerable anecdotal information about the complexity of our interspecies communication. Yet the early scientific attempts to introduce human language to nonhumans were unsuccessful. Aitchison (1998, 35) described two early failures to teach chimpanzees to speak English. Gua, trained by Kellogg & Kellogg (1933) never uttered a single English word, although she appeared to understand about 70 words; and while Viki, trained by Hayes & Hayes in the late 1940s, managed to produce vocal approximations of the words *papa, mama, cup* and *up*, that was the sum total of her spoken vocabulary. Again, though, she seemed to understand many other words (Hayes & Nissen, 1971). The problems with vocalisation are simply mechanical: chimpanzees do not have the laryngeal and orofacial control mechanisms to handle human speech, so their inability to produce it says nothing about their capacity to understand it.

Fouts (Fouts with Mills, 1997, ch2) describes a different and more successful form of language communication between humans and chimpanzees: during the late 1960s, Gardner & Gardner (1969) taught American Sign Language (Ameslan) to Washoe, a female chimpanzee. After one year she had a vocabulary of about 25 signed words, according to stringent rules determining when she had learned a new word. Washoe also began combining her words in ways similar to human children at the two-word stage of language development, although many linguists have refused to accept Washoe's signed communication as language-like (Fouts with Mills, 1997, ch5).

Fouts himself worked with another chimpanzee, Lucy. She gained a wide and productive vocabulary, combining signs to describe new items, especially food. For instance, watermelon became CANDY-DRINK, while radish became CRY-HURT-FOOD. Fouts was also instrumental in the adoption by Washoe of an infant chimpanzee, Loulis, to discover whether Washoe would teach Loulis to sign. So that Loulis could not learn sign from the human helpers, all but seven signs were replaced by spoken English – Washoe had no difficulty understanding that a change of signalling channel did not affect meaning. The experiment proved successful, and Loulis learned his first 55 signs from Washoe (Fouts with Mills, 1997, ch10).

Another Ameslan experiment, conducted by Patterson & Cohn (1990), involved gorillas rather than chimpanzees. Because gorillas have limited motor control over their hands, Patterson referred to the gorillas' signing as GSL (Gorilla Sign Language) rather than Ameslan. Patterson's claims for Koko, her star pupil, are impressive (Tanner *et al.*, 2006), but so different from those for Washoe that they represent either a major difference between chimpanzees and gorillas, or a major difference in what is defined as a sign.

For this reason, Patterson's approach has been somewhat discounted by other scientists as lacking scientific rigour (*e.g.*, Savage-Rumbaugh & Lewin, 1994, 148).

These positive studies contrast with Terrace's work with the chimpanzee, Nim Chimpsky (Terrace *et al.*, 1979). Terrace concluded that Nim was only copying the signs of his tutors; it was a simple copy-for-reward behavioural response, there was no evidence of original signing. Terrace took this to mean that not only could Nim not sign linguistically, neither could any other ape. However, Terrace's position is criticised by Fouts, who placed Nim in a group of Ameslan-signing chimpanzees in Oklahoma after Terrace had completed his experiment. At Oklahoma, Nim showed a dramatic increase in spontaneous signing as he integrated with the existing community of signing chimpanzees (Hess, 2008).

Savage-Rumbaugh (1999) has also taken issue with Terrace's conclusions. Her team at the Yerkes Primate Centre in Georgia (USA) used a keyboard of arbitrary symbols called lexigrams to study language-like communication with two chimpanzees, Sherman and Austin. This was moderately successful, and the experiment was later expanded to include a group of bonobos (Savage-Rumbaugh *et al.*, 1993). However, the first bonobo subject, Matata, proved disappointing: her understanding of the symbolic nature of the task remained suspect, and inferior to that of Sherman and Austin.

While Matata was training she was also mothering Kanzi; and, unlike his parent, he proved an excellent subject: he picked up the meaning of several lexigrams without direct tutoring, merely by observing the training given to Matata. By 17 months, Kanzi was producing novel combinatorial signs that he had not seen his tutors produce. The number of lexigrams available to Kanzi has dramatically increased over time, and his keyboard now contains of over 300, compared to the under 100 of Sherman and Austin. Kanzi has also supplemented the lexigrams on his keyboard with vocalised and gestural signs (Segerdahl *et al.*, 2005).

Figure 10.1: A section of Kanzi's Yerkes keyboard.
Note the *Keyboard* lexigram (bottom left) which uses a part of the keyboard to represent the whole keyboard. This is both recursive and a form of metaphor known as synecdoche.

In other human-chimpanzee communication research, Premack & Premack (1983) involved five chimpanzees, Sarah, Gussie, Elizabeth, Peony and Walnut, in a study in which the subjects attached *cards*, metal objects of different colours and shapes, to a magnetic board. Each of the cards had an arbitrary word-meaning, and they were either arranged into sentential structures for the subjects to interpret or were provided to the subjects so they could arrange them into sentential order themselves. The chimpanzees, with varying success, learned concepts such as *on*, arranging cards accurately to differentiate between GREEN ON RED and RED ON GREEN. They also showed understanding of the concepts *same* and *different*.

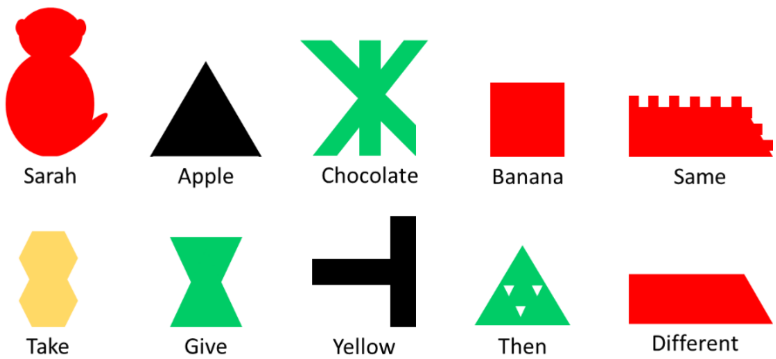

Figure 10.2: Some of the Premacks' word-card shapes with allocated English meanings

However, these ape experiments all rely heavily on the empathy of the human trainers. All the animals were taught language within a human cooperative environment, where playing the language game is a fitness-enhancing strategy for the individual. Many animals who cooperate with humans seem able to acquire language comprehension skills, which is why there have been successful language-related experiments involving non-primates: Alex the grey parrot (*Psittacus erithacus*), who learned a wide range of descriptors and numbers (Pepperberg, 1999); dolphins such as Akeakamai and Phoenix (*Tursiops truncatus*), who learned a communication system similar to that created for chimpanzees by the Premacks (Herman & Uyeyama, 1999); and Rico (Kaminski *et al*., 2004) and Chaser (Pilley & Reid, 2011), collie dogs who both learned hundreds of names for individual toys in large collections; they all show that many species have a capacity for cooperative communication – as long as a cooperative human is at the other end of the communicative act. However, the main question raised by these studies is not *can nonhumans learn human language?*, but *can nonhumans develop the same level of cooperation as humans without human intervention?* – and the evidence for that is still unclear.

The studies of nonhumans using human language do tell us some important things about the sources of language grammar, though: first, that several aspects of language grammar are prefigured in nonhuman cognition, and available for use in communication; second, that they are only used for communication after teaching within a human environment; and third, that language-trained nonhumans still cannot use human language at the same level as a human five-year-old. Evolutionarily, there are both old and new aspects to human language and grammar, and both nonhumans and young humans must be studied to fully understand the sources of language grammar.

11 Language Grammar and Young Humans

Children's minds need not innately embody language structures, if languages embody the predispositions of children's minds!
(Terrence W. Deacon)

Over time, much anecdotal evidence about human child development has become available: every parent has their stories. There has also been considerable scientific research, generating a pragmatic and evidence-driven discussion about the subject. This, however, does not mean that accepted ideas about human child development are wholly the product of practical experience. For instance, Chomsky's Poverty of Stimulus argument (1980) proposes that children are not exposed to sufficient correct language input to learn language grammar, so the fundamental forms of grammar must be innate. However, this idea remains a hypothesis, and has been challenged in principle (Sampson, 1997, 38-45; MacNeilage, 1998) and in detail (Yang, 2002, 101-124; Atran, 2005, 55-60). Consequently, poverty of stimulus argument, while a key concept of Generative Grammar, remains non-canonical in linguistics.

Mostly, though, research into child development has taken a pragmatic and evidence-based approach. There are aspects of language acquisition which are driven by innate capacities (phonology being a notable example), and these aspects therefore require a genetic, rather than a socially interactive, explanation (*e.g.*, Seidl & Cristia, 2012). Acquisition of other features, lexis and semantics, is easier to explain in terms of socialisation and culture (Ellis *et al*., 2021), while acquiring grammar requires both genetic and acculturation explanations. The syntax of the two-argument form seems intrinsic, as does three-argument syntax; whereas the use of determiners, absent from many other languages but a key feature of English, is clearly learned (although whether it is learned subliminally or learned explicitly and then sublimated remains unknown). This mixed model of language acquisition, incorporating both evolutionary and socio-cultural explanations, is often referred to as the Evo-Devo approach (Benítez-Burraco & Longa, 2010).

Cultural development in children is commonly measured against a list of achievements, some physical (*e.g.*, first step, holding and using objects, successful potty training) and some cognitive (*e.g.*, first word, understanding simple instructions, attempting dialogue). In more developed countries, these lists have often been formalised into prescriptive schedules, with children who do not meet the schedule being diagnosed with Child Developmental Delay. This has become a source of anxiety for parents, often unnecessarily so: only a minority of delays have any effect in adulthood. The high frequency of Child Development Delay diagnoses may indicate an overreaction to testing: the US CDC believes that "In the United States, about 1 in 6 children aged 3-17 years have one or more developmental or behavioral disabilities"[10], which makes it either a pandemic or within the limits of normalcy.

The prescriptive schedules mean that childhood is often treated as a steplike process. Piaget (1923 [1959]), one of the first to do this, identified four stages: the sensorimotor stage, from birth to about 24 months, in which children experience the world through sensation and movement; the preoperational stage, from two to seven, in which motor skills and most language is acquired; the concrete operational stage, from seven to eleven, in which children begin to think logically about concrete events; and the formal operational stage, after age eleven, when children become adolescents and develop abstract reasoning (Smith *et al.*, 2003).

Piaget's model proposes a relatively fixed schedule for language acquisition, but he interpolates many substages and autonomous events into the four main stages. For instance, he places the emergence of consciousness, a key event for cognition and communication, in the middle of the

[10] https://www.cdc.gov/ncbddd/childdevelopment/screening.html.

sensorimotor stage at about 9 months (Piaget, 1947 [1950], 113-114). At the other end, at least one stage is not included in Piaget's model: adolescence is unlikely to be the final stage of development because adolescent brains are still maturing and cannot model all the subtleties of social interaction (Sabbagh, 2006). While Piaget's four stages are still widely accepted, the details of his model are less used – particularly by Generativists, who view human language development as innate rather than learned (Parisi & Schlesinger, 2002). However, the view that children acquire language in a series of stages has become canonical.

Vygotsky (1934 [1986], 80-83) offered an alternative staged model in which child development was driven by learning. Most of this learning is not produced by intentional teaching, it is part of the natural human socialisation process: we are genetically ready to learn from birth, although it is how we learn that is innate, not what we learn; and we are genetically inclined toward teaching by example, although it is, again, how we teach and not what we teach that is innate. Where Piaget saw cognition as preceding and producing language, Vygotsky envisaged separate forms of pre-thought and pre-language merging together at about 24 months to produce verbal thought, a key feature of being human.

For Vygotsky, child development is continuous: independent threads of cognition develop both concurrently and serially through pre-adolescence, driven both by physical maturation and previous learning. New learning relies on pre-existing knowledge; and new knowledge, whether gained through physical maturation or previous learning, prepares the child for yet more knowledge. Vygotsky describes this cycle, learning generating new knowledge generating new opportunities for learning, as the Zone of Proximal Development (Vygotsky, 1935 [1978]).

Piaget's model emphasises developmental stages and provides a schedule for developmental events to happen; but it does not explain the transitions between the stages, relying instead on maturation as a hidden cause. Vygotsky's model emphasises continuous development throughout childhood, but it treats childhood development as an idiosyncratic response to specific learning circumstances; it does not effectively explain why human children tend to follow the same developmental path quite so predictably. There is more to the story of human maturation than is told in either of the two models – or even in a synthesis of the two.

Children and the Sources of Language Grammar

If language requires both a programmed acquisition schedule and an open-ended capacity for *ad hoc* learning then, evolutionarily, they must both have given humans a fitness advantage; so we need to know what made them fit strategies. Yet, despite the attention paid to modern human childhoods, we are only just beginning to realise how important the childhoods of early humans are in our story of becoming human. Shea (2016) suggests that looking at the debitage of flint-knapping will help us identify smaller items of debris which could have been produced by children playing at adult toolmaking. Modern children are happy playing at being adults in several ways, most of which involve imitation; and it is likely this was also true for early humans. One way to play at adulthood in the Palaeolithic would have been copying the flint-knapping activities of the adults – and, in the process, beginning to learn the skill. This learning-through-play is also common in modern skill acquisition, although here it is more about acquiring abstract knowledge than practical life-skills.

For Deacon (1997), the way modern children acquire language follows the same pattern as species language development. He sees both types of language attainment as learning processes driven by a need to communicate. While modern language seems elegant and integrated, early language was disjointed and complex, a set of independent responses to different signalling needs. The first languages probably had no structured grammar, only context-specific rules; but the rules were enhanced, honed and integrated over many generations until we reached our current, open-ended languages. Deacon sees this process as a product of accumulating errors made by children over generations, not as a collectively conscious attempt to make better language. The errors simplified and generalised the rules by making language more *child-friendly*.

This theory reverses the normally accepted train of events: language does not start simple and get complex, it starts complex and gets simple. Language is not unitary, and it did not suddenly appear as a complete system (Chomsky, 1988, 183-184); nor did it emerge as a single entity which then developed incrementally (Pinker, 1994, 366), or progressed in a short series of steps (Bickerton, 1990, 128); nor did it emerge to satisfy a single signalling need before expanding to address other signalling needs (Mithen, 1996, 213); instead, language merged several separate, pre-existing signalling systems, each with their own complexities. Language has no primogenitor, it is a by-product of socialisation. Deacon does not, however, address how the merging of simple pre-language signalling systems produced the complexities of modern language.

Deacon's theory matches evidence. For instance, children learning English often over-generalise standard case endings, such as the plural -*s* and past perfect -*ed*; between 12-36 months they produce constructs like *the childrens not speaked good*. Some children even go through a stage where already-learned irregular past tenses are reinterpreted after they discover the general rule: the child changes from saying *ate* to *eated* and then later back to *ate* (Plunkett, 1995). Deacon's theory suggests that some childhood forms do not always revert back to irregularity, and a new form following the general rule eventually replaces the irregular. Thus the past perfect form of *burn* used to be *burnt*, an irregular form tracing back to the Germanic roots of English; but *burned* has recently become increasingly acceptable. Another example would be the plural of *formula*: the anglicised *formulas* is steadily replacing the Latin *formulae*. With both regularisations, English is moving from complexity toward simplicity.

Deacon does not dismiss universality, but he does not see Universal Grammar as necessary. Instead, universality comes out of the interaction between linguistic possibility and phenotypic limitations (Deacon, 1997, 116). Universals are not a side effect of cognitive evolution, as Gould and Lewontin's spandrel theory (1979 [1997]) proposes, they are the physical barriers imposed on humans and therefore on human language by evolution, as Nettle (1999, 5-11) proposes; and the major difference between our minds and those of other apes is the ease with which we think symbolically. We do not need to think in a different way to enter a universe of symbolic thought, we are born into it; and can leave it only with difficulty. Deacon attributes our capacity for symbolic thought to our prefrontal cortex, which is larger than other primates (Schoenemann *et al.*, 2005).

For Deacon, what generated language was the appearance of symbolic thought; which excludes symbolic thought from other animals. Yet, while no other animal in its natural environment has been shown to use symbols in its signalling, it cannot be excluded as a cognitive capacity; and when we look at animals trained in human language we cannot explain their behaviour satisfactorily if we exclude symbolic representation. Symbolic representation may be more common in nonhuman cognition than we currently acknowledge, and it is the absence of symbols from their signalling that needs an explanation. What differentiates us from other animals may not be symbolic capacity but whatever required that symbolic capacity to become shareable.

Deacon proposes long-term pair-bonding, which he calls marriage, as the social structure which generated first language. He is careful to state that the marriage he is referring to is not the Western monogamous model, it can include multiple serial partnerships; but he takes the view that these partnerships must be monogamous and stable over long periods, and that “two males almost never have simultaneous sexual access to the same reproducing female” (Deacon, 1997, 385). This does not match the data available from pre-urban cultures, (*e.g.*, Bergstrom, 1994; Beckerman & Valentine, 2002; Goldstein *et al.*, 2002) where stable pair-bonding is neither the economic nor the moral basis of societies.

Deacon recognises the importance of meat to hominin groups, and that cooperative meat-sharing between males and females is advantageous; but the model he proposes to explain cooperation (males give meat in return for sexual fidelity from the females) is both unenforceable at the individual level, and isolating at the gender level: if males and females are part of a social group then males out hunting for meat to bribe their mates to fidelity have no way of enforcing fidelity on those mates during their absence. To enforce fidelity they would have to sequester their mates away from the group – which means that their children, necessarily sequestered with the mother, would have no social group in which to produce consensuses of new language forms.

However, problems in the details of Deacon's theory should not be overstated. The basic principle, that infidelities in transgenerational transmission mean that language is in a constant state of change, is clearly true. How human children acquire language not only reflects the way humans got to language, it *is* the way humans got to language. Grammaticalization began when the first segmented two-argument [A[uB]] utterances were made, and it is ongoing today (Hopper & Traugott, 1993).

How Children Cooperate

Humans cooperate, and cooperation is at the heart of language (Grice, 1989, 26-31). But does this mean we are naturally cooperative? Looking at the adult response to human infants it would seem to be so: infants activate an innate "cuteness mechanism" in adults, engaging all the senses into treating the infant as delightful (Kringelbach *et al.*, 2016). Cooperation is a major component in the evolutionary fitness of individual humans (Taylor & Day, 2004), mainly because cooperative adults raise more children to maturity than non-cooperators (Traulsen & Nowak, 2006). Cooperation, however, is an unexpected fitness trait: it is open to exploitation by non-cooperators. Yet, as we have seen, if altruistic punishments like social exclusion are not too costly for the punishers and sufficiently debilitating for the punished then cooperation can be a fitter strategy than aloofness or cheating.

We are not, however, born fully cooperative. We have some genetic features present at birth promoting our cooperation, such as an awareness of and attraction to other humans, especially our primary carer (Healy, 1994, 15-17); and we tend to treat most of our encounters with others, up to 24 months at least, as benign. We have an innate default assumption that people will cooperate with us even if we cannot yet understand how to cooperate with them (Gopnik *et al.*, 1999, 25-31).

Most of our cooperative behaviours, though, are acquired during childhood. Children do not play cooperatively until about 36 months: before then they use parallel play, playing together in the same place, sometimes with the same objects, but not at the same game (Brewer, 2001, 30-31). Three-year-old children, even though they are willing to share an imagined universe for play purposes, are cooperating because their desire to play their game must mean that everyone wants to play it: there is only one intention in the universe. By age four a child is usually aware that others may or may not want to play – they have their own intentions. This ability to view others as intentional is often regarded as the beginning of Theory of Mind, and therefore the beginning of language as true dialogue rather than vocalised thought (Foley & Thompson, 2003, 25-29). Before this, the child does not have an internal dialogue of “inner speech”, only external “social speech” (Meares & Sullivan, 2004).

Cooperation involves both nature and nurture, it is a product of both innate features and socialisation; and one way to understand how these two factors work together in the acquisition of language is by considering how sociological and pathological deprivation in childhood affect language. Sociological deprivation is a product of grossly abnormal childhoods, such as those experienced by feral children; and pathological deprivation involves non-standard genomes producing non-standard brains – for instance, Williams Syndrome or Turner Syndrome or Autism Spectrum Condition.

While feral children give an insight on sociological deprivation, recorded cases of such extreme child neglect are rare. The textbook case is that of Genie: born in 1957, she was kept isolated by her father in an environment without access to language or other humans (except her father, mother and brother) until age 13. Her subsequent treatment by well-meaning scientists, Social Services and her mother, meant that she never experienced anything like a normal life; and, as a test of nature and nurture in language acquisition, Genie’s case is hopelessly compromised. She eventually disappeared into obscurity in 1978 after access to her as a data subject was terminated by loss of funding and by her mother forbidding continuation of the research (Newton, 2002, ch7).

Scientific study of other feral children is similarly problematic: the extreme circumstances of their lives make it hard to identify specific causes for their language abnormalities. However, some general conclusions can be drawn. First, both socialisation and language were abnormal in all cases, and neither developed to acceptable normality. Second, the levels of socialisation and language finally achieved varied considerably, although in all cases high language achievement mapped to high socialisation, and *vice versa*. Third, although the willingness to cooperate with others was also highly variable, where language and socialisation were high then so was cooperation; but it is impossible to judge whether cooperation is a source or outcome of language ability.

When looking at autistic children, the effects of nature and nurture are easier to identify. Most autistic children have care regimes similar to those of other children, but their response to the care and support offered is unusual. The socialisation of autistic children is predictably different: they seem to have an unusual Theory of Mind and their linguistic comprehension is literal (Greenspan, 2001); their linguistic production also differs (Meir & Novogrodsky, 2019). They have difficulty cooperating with others in both shared ventures and play, preferring stereotyped behaviour to experimentation (Smith *et al*., 2003, 477-481). Where a neurotypical child finds routine boring, an autistic child finds it comforting. There is clearly something different about the neurodivergent autistic brain.

Frith (1993 [1999]) believes a single cognitive capacity is impaired in the autistic brain, identifying this component as “the ability to think about thoughts or to imagine another individual’s state of mind”. Baron-Cohen (1995, chs4-5) sees the problem as more complex: Theory of Mind involves two stimulus detectors and one modelling module. First is the Intentionality Detector, interpreting the desires of others from their expressions; for instance, a grimace indicates dislike. Second is the Eye Direction Detector, mapping the gaze of others to discover their goals; for instance, what the observed person is looking at is why they are grimacing. These detectors both give dyadic representations of the observed agent and their goal, but they are then combined in the Shared Attention Module to give a triadic model of the agent, self and goal. The Shared Attention Module sets the stage for neurotypical Theory of Mind.

For Baron-Cohen, autism is an impairment of the Shared Attention Module. Autistic children interpret the actions of others in a mechanistic way, being able to identify the focus of their subject's attention and how they feel about it; but they cannot interpolate a reason for their subject's attention, which makes triadic relationships between self, other and goal difficult to model. Theory of Mind is compromised before it begins.

Increasingly, it is possible to diagnose autism by examining brain function (*e.g.*, Morrel *et al.*, 2023). When performing tasks that involve modelling the minds of others, non-autistic and autistic brains have very different patterns of activation: the prefrontal cortex is heavily used by non-autistics but remains less active in autistic brains (Carter, 1998, 141-143). The prefrontal cortex is larger in humans than other animals, and it is associated with distinctly human cognitive faculties: planning, imagination, selfhood, awareness of other, working memory, and space-time cognition (Greenfield, 2000, 144-153). More recently, Villar-Rodríguez *et al.* (2023) have shown that individuals with atypically lateralised brains (where language is not centred on the left hemisphere) often have a mirrored inhibitory control network in their right hemisphere and stronger interhemispheric connectivity; and there seems to be a link between atypical lateralisation and autism. While Xin et al. (2023) found that, in non-autistic infants, language facility is significantly correlated with the grey matter volume in the bilateral prefrontal cortex and cerebellum; no correlation was found in autistic infants.

Studying autism lets us see both the importance of cooperation for humans and how the brain is involved. The prefrontal cortex in archaic *Homo sapiens* was almost as well-developed as in modern humans (Aiello & Dean, 1990, ch10), so we can say with some confidence that the appearance of modern *Homo sapiens* did not involve dramatic changes to the brain; and that the modern human prefrontal cortex generates pre-emptive cooperation through self-modelling and modelling others. Do we cooperate because of our modelling, or do we model to enhance our cooperation? The answer probably involves a feedback loop: modelling allowed our ancestors to anticipate and accommodate the intentions of others, which then enhanced our cooperation; and cooperation gave us better understanding of the intentions of others, enhancing our modelling.

How Children Acquire Selfhood

The role of selfhood in language acquisition has a long history. Wilson (1937, 143-146) discussed how awareness of self uniquely defined the communication of our species, while Piaget (1923 [1959], 39-43) described young children as linguistically egocentric: their early utterances mostly express an internal monologue instead of a social dialogue, because their unrefined model of selfhood makes their language use suboptimal. Vygotsky (1934 [1986], 217-235), viewed the emergence of consciousness as gradual, a series of emergences of “consciousness of”. Language begins as communicative, not cognitive: egocentric speech is the first step in language internalisation, developing into inner speech by about age five. Only when speech has internalised can narrative self-consciousness begin.

Awareness of self is not present at birth: few, if any, humans can recollect their early years, with any memories of events before age four being isolated from our self-defining narrative memories. Infants are unaware of their individuality during the first 6 months, treating other people and objects as physical extensions of themselves (Brewer, 2001, 17-23). Between 6-24 months, infants are usually being acculturated by their care-givers “from an organism to a person” (Kaye, 1982, 205). After 24 months there is an identifiable self being asserted, and this often creates a period of carer-child conflict referred to as “the terrible twos” (Brewer, 2001, 218-222).

Gopnik *et al*. (1999, ch2) describe the child’s awareness of self and others as developing in stages. In the first 6 months the infant is building on innate knowledge. They know the significance of human faces and voices, and they are learning to recognise their caregivers and how to identify others. They are also learning about facial expressions, and how they relate to caregiver’s behaviour. By 12 months, infants are beginning to see others as agents: other people attend to, and affect, the world. The child learns about pointing and eye direction, and they look at objects being attended to by others. Additionally, they are learning how to affect the world remotely by recruiting the muscle-power of others.

By 18 months the child is usually aware that the usefulness of others can vary – they sometimes help, sometimes not. The child is also beginning to realise that knowledge is personal, others may not know everything the child knows. To communicate their needs better and get what they want, the child begins using indexical word labelling.

At about 24 months, children begin to show empathy toward others. Initially, the competition between their own desires and those of others creates the frustrations and tantrums which have become known as "the terrible twos". However, the child quickly learns negotiating techniques which allow them to integrate their agendas with those of the people around them, and by 36 months the child seems to understand the important role negotiation plays in their social interactions. This capacity to negotiate is accompanied by greater emotional control and social competence, with the child fending for themselves in important ways, such as feeding, washing and dressing: given the raw materials, they will attempt to finish the job. At 36 months children are also becoming deceptive: others may not know what they know, and this lack of knowledge means those others can be manipulated. However, the child is usually hopelessly inept at taking advantage of others' ignorance: they know what the other person does not know, but they do not have the mechanisms to turn that ignorance to their own advantage. By age four, most children have an effective Theory of Mind: they make effective guesses about what others are thinking. This also teaches them two important lessons: first, others are aware they can be deceived, making successful deception a complex and costly cognitive process requiring deception in multiple modalities (Wray, 2002a, 128-129); and second, humans are primed to altruistically punish those they discover deceiving them (Fowler, 2005).

The Gopnik *et al*. (1999) model observes a steadily increasing awareness in the growing child, generated by increments in socialisation and cooperation. The young infant needs no concept of self or other; but, to co-opt the muscle-power of others – which parents and carers are usually willing to provide –modelling of others becomes useful. As the child develops it learns that others are more than objects to satisfy demands, they are agents who have their own reasons to help – or not. The Machiavellian intelligence of apes is sufficient for coopting others, but cooperating with them requires negotiation to understand and accommodate their intentions.

Humans, however, are raised in a highly cooperative linguistic culture; which means they are exposed to the knowledge that others are modelling them just as they are modelling others. For instance, if an adult says to a five-year-old child, "let's go to the park", the child is likely to understand that the adult has a model of the adult and child together in a different place and time than now. Comprehending that others are making models of both you and themself gives you all you need to make models of yourself. This self-model does not directly represent the self, though: your model of you is actually your model of their model of you (Gopnik *et al.*, 1999, 47).

The modelling by others of me and themselves makes my first-person model of myself possible; and, by treating this first-person model as a third person, I can examine both the attitudes of others toward me and my own attitudes toward myself, making me aware of my individuality. Cooperation, socialisation and culture drive us toward language, language drives us toward awareness of self, and awareness of self drives us toward further cooperation. Our innate need to cooperate is the mechanism that drives language acquisition and Theory of Mind; but it is the cooperative culture of negotiation and altruistic punishment that creates the environment in which these systems flourish.

How Children Acquire Language Grammar

Does the ontogenic grammar acquisition of modern children follow a path similar to the phylogenic development of grammar in the hominin clade, as Haeckel (1874 [1912]) suggests? The cognitive structures used in modern children's grammar acquisition are likely similar to the cognitive structures used in ur-languages; so, as childhood language acquisition builds up incrementally in stages, we can assume the same stages in the development of ur-languages. The cognitive capacity for social calculus, which develops throughout childhood, should also be mappable to the evolution of social calculus in the human clade because, as has already been shown, social calculus and language grammar are intimately linked. And, taking Piaget's approach, we can also say that the stages in phylogenic development of grammar, as in ontogenic grammar acquisition, are discrete – for instance, we do not encounter two-and-a-half word utterances.

As well as Piaget and Vygotsky, others have produced staged models of language acquisition. Brown (1970) proposed four stages: pre-language up to about 6 months; one-word utterances up to about 18 months; two-word utterances to about 36 months; and full language starting at about 36 months. Hirsh-Pasek & Golinkoff (1996) give a different schedule, based on language-as-communication rather than Brown's language-as-composition: from birth to 9 months the child associates sounds made by carers with objects in the world; from 9-24 months they understand the concept *word* and learn about attention; from 24-36 months they learn about combining words into two-argument forms; and after 36 months they begin to expand their language model into full language. Halliday (2004, ch16), using a Vygotskyan approach, divided language acquisition into three phases: language learning, dominant up to about 24 months; learning through language, from 24-48 months; and learning about language, from age four onward. For Halliday, language acquisition continues throughout life, it is not a skill perfected in the early years.

The models described here all agree that language is acquired in stages with significant cognitive changes in each new stage; and they agree that progression through the stages follows a permissive schedule – there is no strict calendar of acquisition to measure individual children against (Bates *et al.*, 1995). The models also view the stages as discrete: each stage adds capacities that were unavailable previously, and introduces new socio-cultural maxims to be learned and applied. Sometimes these new maxims replace older ones, and sometimes they supplement them.

However, while they support a wide range of transition ages for individuals, the models also recognise generally applicable schedules for childhood language acquisition. Each model has its own schedule; but by merging them a composite schedule can be generated. Based on the models discussed, a six-stage synthesis is proposed below.

- **The pre-language stage**: a sound is way of attracting attention. There is little differentiation between sounds in terms of meaning.
- **The phonetic stage**: certain sounds appear to elicit more or better attention from others. These are the first language-like sounds that the carers associate with “being human”, although the child is not aware of this (Berg, 1972, 7).
- **The word stage**: sounds have meanings. Objects can be requested or named with sounds, and personal wants can be better met by making the sound associated with the effect desired. At this stage, however, the sounds are indexical, in that they have pre-established meanings; they is no negotiation between sender and receiver over intended meaning, so they are not symbolic.
- **The one-argument stage**: words can be combined to produce more accurate requests and enhanced outcomes. Some of the sound combinations are recognised by the child as having a wordlike quality, which allows them to be used in a range of circumstances and constructions (Tomasello, 2003a, 139-140). The child is beginning to acquire a meaningful grammar, recognising that different word combinations have different effects: *kiss teddy* means that mummy is to kiss teddy; *teddy kiss* means that teddy is to kiss mummy (Gopnik *et al.*, 1999, 117). This seems to be the highest stage reached by animals taught humanlike symbolic communication in a human cultural environment (Tomasello, 2003b). It is also the stage at which the child begins to understand the need to negotiate toward meaning: the sounds are no longer indexical, they are symbolic.
- **The two-argument stage**: for every action there is something active (the actor) and something being acted on (the patient). *Mummy kiss teddy* is different from *teddy kiss mummy*; and while *juice in cup* is one outcome, *juice in bottle* is another, although they both involve getting juice.

- **The full language stage**: the child's grammatical knowledge is largely complete. While there are complex language forms and individual rules still to be learned, no further foundational grammar is needed. Also, while this stage largely completes the grammatical set required for full language, other intelligences, such as emotional intelligence, have not yet matured.

As well as the acquisition of language in children, this model maps a way to develop language as a species. The first two stages are pre-human, and the word stage may to be available to all apes, or even all primates. The one-argument form seems to be cognitively available to apes, but it is only communicated when they use humanlike symbolic communication. Apes may also have two-argument forms available in cognition to support their social intelligence; but communicative examples among apes using humanlike symbolic communication are rare and disputed. The final stage, full language, seems limited to modern humans, representing both a cognitive and communicative difference between humans and other animals.

However, there remains an important unresolved issue in mapping the ontogeny of grammar acquisition onto the phylogeny of grammar development: ontogeny emphasises intrapersonal psychology, or individual nature; phylogeny emphasises inherited interpersonal socialisation, or genetic nature. Despite Haeckel's dictum that "ontogeny is a brief and condensed recapitulation of phylogeny" (1874 [1912], 2:16, 179), they remain quite different in their trajectories. Some more recent approaches to ontogenetic language acquisition attempt to redress this imbalance.

For instance, Behrens (2021), taking a constructivist approach, suggests that children discover the rules of their first language by constant analysis of input. Morphology is discovered by testing new data against already-established regularities to identify and incorporate any new regularities found. Affixes and roots are treated in the same way, although affixes are also subject to ordering and functionality rules. Syntax, or word order, relies on schemata determining which words can – or cannot – be placed together. Language grammar is, in Behrens' model, an emergent social phenomenon, a convention rather than a genetically encoded rule system. While Behrens' approach does not fully explain the speed with which a child's language converges on the group language, it does provide a simple, cultural mechanism for generating new forms: they arise simply as fashionable memes which spread the new forms swiftly through a community. We should not be surprised by the constant changes that occur in a language, they are inevitable. However, as Behrens' model sees language as a social phenomenon, it is largely silent on the phylogenic sources of language grammar.

In contrast to Behrens' approach, Qi *et al.* (2021) see the developing human brain converging toward a fairly fixed functional architecture during pre-adolescence. Looking at brain development in 175 children aged 4-9, they found continuous elaboration in connectivity between the pars triangularis in the left hemisphere inferior frontal gyrus (also known as Broca's area) and the regions around the left hemisphere temporoparietal junction (also known as Wernicke's area). These two surface cortex areas are linked below the surface by the arcuate fasciculus which, compared with other Great Apes, is considerably enhanced in humans (Rilling *et al.*, 2008). The two cortex areas are also often larger and thinner in humans, with size and reduced thickness being associated with greater individual language ability. These features indicate physical and functional differences between the traditional cortical language areas in humans and other Great Apes, and show how the pre-human language-unready brain became the modern human language-compatible brain.

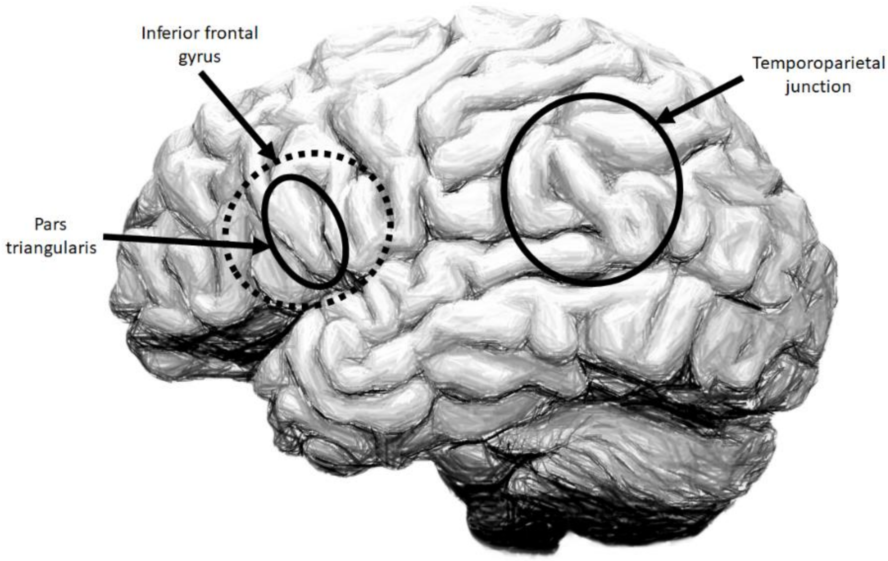

Figure 11.1: Traditional map of language areas in the human brain

Another recent approach reviewed gesture as both a support for spoken language and a parallel mode of communication (Goldin-Meadow & Alibali, 2013). Gesture elaborates and interprets the spoken word, allowing the speaker to engage in visuomotor thinking, whether directly or obliquely related to the verbal content. Children initially point to support their non-linguistic sound-making, but soon use pointing to indicate one argument in a more complex construct: pointing at a cupboard and saying “bikkit” indicates the carer should get the named thing from the indicated cupboard because the child wants it. Not only does gestural grammatical complexity keep pace with spoken grammatical complexity, they merge into a single multimodal communication system. Far from just adding emphasis, gesture is a key feature of the multimodality of human language, allowing the speaker as well as others to literally see the thought processes generating their speech, and even letting them change their mind on the fly. By observing gestures accompanying speech we get a better understanding of the evolution of cognitive grammar than is provided by just phonological, semantic and syntactic evidence.

Evolving over generations of human selection within a cooperative culture, humans are predisposed to acquiring language. This predisposition is at least partially encoded genetically: humans born predisposed to language probably had greater reproductive success than those born with less capacity for it. Language, and the forms of cooperation enabling it and enabled by it, have been fit strategies for generations of our species: we have adapted simultaneously to language as cognition and language as communication. Language, however, remains largely learned, incrementally building personal knowledge of how and what to communicate moderated by a constant negotiation toward meaning. The ontogenic stages of children learning “how to language” do not fully map to the phylogenic development of language as a communication system, but they do help us understand how humans evolved toward language.

12 How Language Grammar Manages Time

Time is like a river of passing events, and strong is its current; no sooner is a thing brought to sight than it is swept by and another takes its place, and this too will be swept away.
(Marcus Aurelius)

Time: a key component in the human interpretation of the Universe, and yet we still have only an indirect and metaphorical understanding of it. Human temporal awareness probably provides an incomplete impression of time, just as our visual awareness accesses only a small part of the electromagnetic radiation spectrum; but with vision we are now aware of what we don't see, while we are still unaware of what we don't know about temporality. Nonetheless, we can communicate about our shared understanding of time, regardless of its accuracy. Like gravity and light, time seems to be a key component of the Universe, defining everything in relation to itself. In cosmological terms, time is just another dimension, similar to the three dimensions that define space; but the nature of time means that humans perceive and model it quite differently from those other dimensions (Reichenbach, 1927 [1956], 109-113).

Primarily, we perceive time as a story: things happen in a set order, moving relentlessly from beginnings to conclusions. Every event occupies a delimited period, dependant on what happened before and determining what happens after; and movement through space also involves movement through time, because spatial movement relies on duration. There is some theoretical evidence that movement is possible from present to past, or to the future at an accelerated rate (Hawking, 1996, ch10), but this time-travel is not part of our everyday human experience. Unlike the spatial dimensions, we cannot control our position in time, which is always *now*; and we cannot control our direction or speed through time. Instead, we all seem to be heading inexorably into the future at the same rate: the story of time is both communal and inescapable.

We cannot see, hear, or feel time but we are nonetheless aware it is passing. Unlike the other dimensions, we can know time only in terms of metaphors, usually of the other three dimensions. We can model time from the inside as a line ahead and behind us, with either an immobile self in a moving stream of time, or an immobile path and the self moving along it. We can also model it from a position outside time, as a horizontal line from left to right in front of us, with either us or the line moving; or we can view the line as vertical instead. We even view time as cyclical, moving regularly through a repeating cycle – a representation used by rotary clocks and watches, creating a key Western metaphor of time, the cyclical day; and we can see time as helical, going through cycles that nonetheless move forward and do not repeat exactly.

Another time metaphor common in the West is transactional time: we can save, spare, spend, waste, give, take, run out of, or even ask for more time, all of which see time as a commodity with economic value. We can even invest time, doing something now to anticipate a future need (*e.g.*, making a packed lunch this evening to save time tomorrow). The transactional view of time is largely illusory, however: time cannot be saved today to give you more tomorrow, both days contain the same number of hours no matter what you do. Nonetheless, we rely on the metaphor *time is money* despite knowing the two are not equivalent: money is, by its nature, fungible, it is widely exchangeable; time is non-fungible, it disappears whether it is profitably used or not. No metaphor of time is truly commensurate, but some are better than others.

Time, for humans, has direction: we move through time (or it moves through us) from a remembered or recorded past to an unknown future. Past and future are quite different from the directions of 3-D space: we usually know where we are in space because we can see around us. In contrast, time is not bidirectional: we can know our past, but we cannot know the future. This is not trivial, it can generate different cultural views of time, as we saw in chapter 5: for the Aymara, the past is what they see, so it must be in front of them; the future cannot be seen, so must be behind them (Núñez & Sweetser, 2006). The Aymara embodied directionality reverses our Western experience: different cultural perspectives produce different analyses of time.

Metaphors of time let us create clocks, calendars, schedules and timelines as physical measuring-sticks of time, and our metaphorical relationship with time lets us model and forecast the future and remember and model the past. This modelling relies on the metaphor TIME IS SPACE (Evans, 2007), plotting the present as *here*, and the past and future as two different *not-here* places; which, in turn, this lets us model our self into the past, present and future, giving it continuity through time. Self-modelling must be part of cognition to make self-continuity through time possible.

Modelling the future is always imprecise, so we usually place more trust in memories of the past than models of the future. Yet the remembered past is only as accurate as our memory, which can become falsified as distance from the present grows (Johnson *et al.*, 2012). Recalling a memory involves more than just recalling facts; context and beliefs adjust our view of the memory to suit current cognitive needs (Heald *et al.*, 2023), and often these adjustments are consolidated back, replacing the original memory (Gilboa & Moscovitch, 2021). It is even possible to generate completely false memories, where imagined events are presented as memories (Miller & Gazzaniga, 1998). It turns out that both future and past can become falsified, because neither are present experiences, only models and memories.

Language makes the memory problem even more complex, by allowing others to present their memories to us as reported memories. Reported memory can be treated as factual if the source is treated uncritically as reliable. This gives us four types of memory: Scientific Facts, which include innate knowledge about and conscious experience of the physical world; Personal Memories and Forecasts experienced or generated by the self; Hearsay, reported memories treated as factual; and the shared fiction of Storytelling, where objective truth is not an issue. We treat Scientific Facts as factual and seldom challenge them, while Personal Memories and

Forecasts are treated as factual within the accuracy limits of memory and modelling, and hearsay is treated as only provisionally factual, or even as fictional. Storytelling is a peculiar recognition that there is more to human existence than being factually correct, and being human is, in large part, about socialisation. This gives us a relationship with truth and fact summed up in figure 12.1.

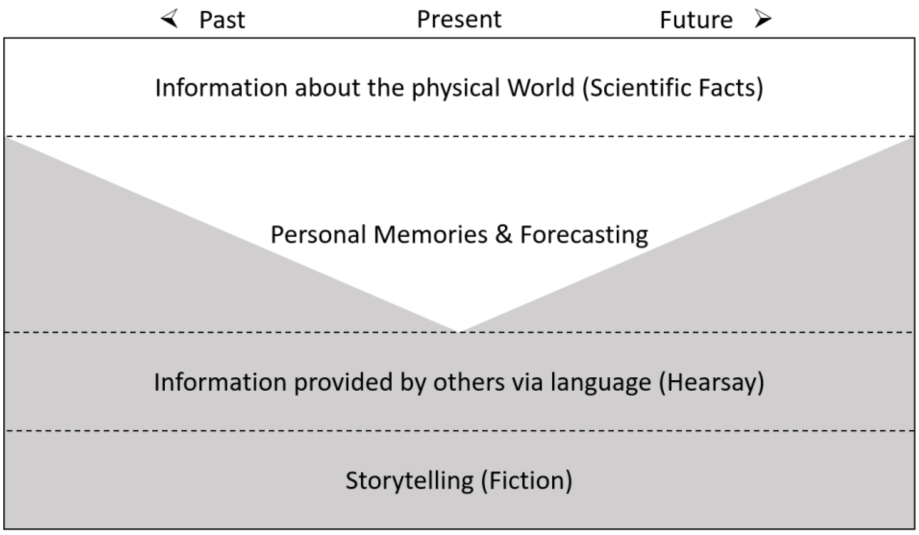

Figure 12.1: Information reliability varies, depending on source and distance from Now (White areas represent reliable information, grey areas represent less reliable information)

The Importance of Time in Human Cognition

How we perceive time determines the range of selves we can model: our models of our self in the past, present and future are reliant on our current views about the past, present and future. It is, however, possible to have a sense of time without self-modelling: the unmodelled actual self is fully capable of generating a subliminal predictive sense about future events based on prior experience. A simple punishment and reward system can generate a subliminal sense of time using memory, forecast and the instinct that past events predict the future. For instance, wasps will sting in the future because they have stung in the past, so avoid wasps. This needs no model of a self being stung to be effective, and this subliminal predictive mechanism may be how most life on this planet experiences time.

What happens, though, when a capacity to consciously model the self into the future or past appears? The two components of the simple temporal model, the unmodelled actual self tied to the present and the modelled past memory or future forecast, become extensible. A modelled self, once created, can be projected into the past or future where it can, like the unmodelled actual self in the present, have memories and make forecasts. For instance, the self modelled into the future can have memories of events that are in its past but still in the future for the unmodelled actual self in the present – and, therefore, unknown to the unmodelled actual self.

This ability to refer to events in the future as if they are in the past or future of a future time, and the ability to refer to events in the past as if they are in the future or past of a past time, are described by Reichenbach (Areces & Blackburn, 2005) as an interaction between Point of Speech (S, commonly the present), Point of Event (E, when the action of the verb actually takes place), and Point of Reference (R, the temporal location of the modelled self). This mechanism encapsulates the vast majority of what languages do with tense.

Reichenbach's three points give us seven natural tenses: the present tense, in which the unmodelled actual self, the event, and the modelled self are collocated; the simple past tense, a memory of an event that has happened; the simple future tense, requiring forecasting of an event going to happen; the past of the past, requiring a memory as a feature of a self modelled into the past; the future of the past, requiring a forecast as a feature of a self modelled into the past; the past of the future, requiring a memory as a feature of a self modelled into the future; and the future of the future, requiring a forecast as a feature of a self modelled into the future.

These seven tenses are represented in different languages in different ways. For instance, in English they are formed from a mixture of inflections and auxiliaries: *I do, I did, I will do, I had done, I was going to do, I will have done*, and *I will be going to do*. Other languages use different methods, but expression of these seven tenses is possible in most human languages – even languages described as tenseless can call on paralinguistic resources to replace direct tense marking (Bohnemeyer, 2009). Accurately expressing temporality is so important that it is often treated as a marker of fluency in that language (Paradis & Crago, 2000). For instance, some bonobos can generate simple past and future tenses when using human language, but none of the language-trained apes use more complex tenses (Savage-Rumbaugh *et al*., 2005). Using a grammatical feature, however, does not necessarily indicate competence in its use: a study of four children using the Japanese past tense marker at around age 2 (first language Japanese) showed that usage did not always match intended temporality (Shirai & Miyata, 2006).

Further tenses can be generated by setting the Point of Event in the present even though the Point of Reference is in the past or future – a process often called the historic present tense. This is not a mere rhetorical trope, it creates immediacy in the utterance or text, inviting the receiver to model themself as if their modelled self at Point of Reference were their actual unmodelled self. The historic present tense can create immediacy about past events (for instance, *It's the eve of Waterloo; Napoleon is in his tent; tomorrow he faces his final defeat*); while in Tom Paxton's *Daddy's takin' us to the zoo tomorrow*, the historic present tense is used for future events, with only *tomorrow* indicating that the events will actually happen in the future.

Tense	Point of Speech (S) Unmodelled self	Point of Reference (R) Modelled self	Point of Event (E)
Simple Present	Present		
Simple Past	Present		Past
Simple Future	Present		Future
Past of Past	Present	Past	Past of (R)
Future of Past	Present	Past	Future of (R)
Past of Future	Present	Future	Past of (R)
Future of Future	Present	Future	Future of (R)
Historic Present (Past)	Present	Past	Present
Historic Present (Future)	Present	Future	Present

Figure 12.2: The Seven Reichenbach Tenses plus Two Historic Present Tenses

All these tenses use Reichenbach's three points. Present tense merges all three points in the present, while simple past and future merge Point of

Reference with Point of Speech in the present. More complex tenses place the Point of Event in the future or past of the Point of Reference, which is in the future or past of the Point of Speech (Hornstein, 1990, ch1). With historic tenses, we know that the Point of Event is in the past or future, but we treat it as current.

Reichenbach's three-point analysis gives an effective way to describe the key tenses used in languages. The system can also be extended by adding another Point of Reference to the Point of Reference, producing four-term constructs such as *I will have been going to do*; but these constructs are not easy to interpret, and tend to introduce effects that are not strictly tense-related. For instance, the difference between *I have been going to do* and *I had been going to do* is that the first indicates an intention continuing into the present, while the second indicates a former intention now abandoned. There is no systematic, four-point tense analysis in *I had been going to do*, other temporal effects are at work.

Temporal Complexity in Language Grammar

Tense markers are one of several ways we express time in English. Further temporal marking can be added using adverbials to link events together (*e.g., before, after, since, etc.*), allowing temporal relationships of even greater complexity (Michaelis, 2006). This is particularly noticeable where a second event is entailed on the first event, as in *Alf is leaving when Beth arrives*. Both tenses, *is leaving* and *arrives*, must be historic present (future) because *when* indicates that Alf's leaving is contingent on Beth's arrival, and that Beth has not arrived yet – if she were already here then *when* would be replaced by something like *now* or *because*, and *arrives* would be replaced by *has arrived*.

Reichenbach's system shows how time is represented by modelling selfness into past and future; but it also illustrates several other linguistic temporal effects, often described as aspectual features (Comrie, 1976). The first of these is continuity through time: the perfective aspect represents a single event complete at the Point of Event (*I wrote a letter*), while an imperfective aspect represents a single ongoing event (*I am writing a letter*), one of a series of events (*I am writing some letters*), or even a series of events of which none are happening at the Point of Event (*I write a letter every week*). Continuity adds duration to the Point of Event, allowing expression of ongoing events as well as abrupt events.

A second linguistic temporal effect is imminence, which is about the distances between Reichenbach's points. A Point of Event can be close in time to a Point of Reference or more distant; and the distance between Point of Reference and Point of Speech can also vary. The nine tenses give the temporal order of the three points, but imminence determines their proximities.

However, while proximity can vary, many languages recognise only closeness and remoteness. For instance, some East African languages have two past tenses with different imminence (Lee, 1992, 9), as does English. The Point of Reference in the sentences *I wrote a letter* and *I have written a letter* is the same (the present) and the Point of Event is also the same (the past); but the Point of Event is closer to the present in the second sentence than in the first.

Imminence can occur in the future, too. In the sentences *I will write a letter* and *I am going to write a letter*, the Points of Reference and Points of Event are the same (present and future respectively); but, once again, the Point of Event of the second sentence is more imminent. In this case we can create even greater imminence with *I am about to write a letter*, indicating that, in English, imminence is not a strict binary dichotomy of near and far.

In English, imminence is often expressed with relative adverbials like *soon* and *just*, absolute adverbials like *tomorrow* and *last week*, or adpositional phrases like *by tomorrow* or *before next week*. Usually, relative adverbials affect the distance between Point of Reference and Point of Event, while absolute adverbials work between Point of Reference and the present. Thus, in *tomorrow I will have almost finished it*, we use *almost* to indicate that the Point of Reference of *tomorrow* is close to the Point of Event of *finishing*, while *tomorrow* fixes the distance between the Point of Reference and the present as one day. *Almost* also has a second role in this utterance, converting the event from completed (or perfective) at the Point of Event to incomplete (or imperfective), showing that temporal effects cannot always be isolated linguistically.

Building Temporal Complexity into Discourse

Temporality is not limited to individual events, it also defines order between events. Language lets us signal connections between events, thus sharing ordered timelines linking past, present and future events. This process, here called connectivity, is not an expression of temporality within a single utterance, it establishes temporal connections between utterances; and it therefore strongly identifies with the Logical metafunction of Systemic Functionalism. Connectivity facilitates the never-ending discourse of language by making the content of previous conversations accessible in the current conversation.

With temporal connectivity all shared events can be identified as simultaneous or sequential. For instance, in *he looked and listened*, *he looked before listening* and *he looked after listening*, the connective determines the order of the events. Connectives also place identities into a time series: in *he ate the plum, then the peach and finally the banana*, the event, eating, is applied to a series of objects in turn. While some connectors, like *after* and *before,* explicitly describe temporal relationships between events, not all do this. For instance, in *he jumped on his horse and rode into the sunset* we see *and* as linking two events serially: both events are in the past, but the first must happen before the second can occur. In comparison, in *he sat on his horse and stared at the sunset* the two actions are probably simultaneous; to serialise the events we would use *and then* or just *then*. Our shared implicit contextual knowledge is, therefore, at least as important as our shared explicit words.

Connectivity is no trivial side-effect of language; it is central to storytelling (Dautenhahn, 2003), and storytelling is central to being human. When we cognitively model past and future events into a timeline we are telling ourselves a story, extrapolating causes into effects to determine outcomes, or manipulating memories to recreate or reinvent history; and if our first attempt does not tell us the story we want, we can always model other stories.

Continuity works at the Point of Event, determining the duration of the event; and imminence works between Point of Event, Point of Reference and Point of Speech, determining the distances between the points. This gives a rich single dimension for linguistic expression of time. Connectivity adds another dimension to temporal space, linking individual events together into networks of events. It gives us narrative, and without narrative dialogue becomes an exchange of unlinked facts (Benson, 1993). Connectivity generates the continuous interpersonal narrative that language has become, making it a crucial differentiator between human language and other signalling.

Time, Conditionality and Imagination

Alongside the one-dimensional timeline of Past and Future and the narrative dimension of connectivity, there is another temporal dimension concerned with how language deals with uncertainty or modality – or, in linguistics, conditionality. Conditionality lets us place events onto a vector of probability, which works with the other two vectors of connectivity and serial time. In English, conditionality is mainly expressed through adverbials, with limited support from auxiliary verbs. For instance, *I may have done* and *I may do* are permissible English forms, but **I may had done* and **I may will do* are not. With adverbials the range of temporal expression is wider: *perhaps I had done*, *I will possibly do*, *I have likely done*, *I am probably going to do*, *maybe I will have done*, *hopefully I will be going to do* … These all add uncertainty into pre-existing verb constructs.

Because we experience past and future time differently, conditionality for past events differs from future conditionality. Events in the future already have uncertainty because the future, by its nature, is unknown, and adding conditionality only increases uncertainty. Past events, in contrast, have greater certainty, and adding conditionality converts certainty into uncertainty. For this reason, future conditionality tends to involve volition, establishing personal control over an undetermined future, while past conditionality is about review – and often regret.

Auxiliary conditional verbs (*may*, *could*, *should*, *etc*.) also show asymmetry between past and future, as the replacement of *will* with *may* illustrates. *I may have done* does not express the same temporality as *I will have done*: while *will have* is a past of future tense, *may have* is a simple past conditional tense. Conditionality seems to move the Point of Reference through the probability vector instead of the time vector, indicating that treating conditionality as separate from the temporality and narrative vectors provides a productive metaphor.

In terms of future temporality, the conditional vector probably existed before *Homo sapiens*. It allows us to plan, choosing between alternatives; and, as toolmaking seems to require a capacity to plan, cognitive conditionality must indeed be ancient. Experiments with New Caledonian crows (*Corvus moneduloides*) have shown them capable of planning the retrieval of difficult-to-access food by making the tool necessary for access, and then using it appropriately. In one experiment the crows worked with unfamiliar materials and an unnatural environment, but they were still able to bend a metal strip into a hook and use the hook to lift a pot of food out of an otherwise-inaccessible hole (Weir *et al*., 2002); attributing a planning capacity to the crow gives the simplest explanation for this behaviour. Experiments with chimpanzees have shown them working together in tasks requiring cooperative planning (Melis *et al*., 2006), although they do seem better at planning for competition than for cooperation (Hare & Tomasello, 2004).

Full conditionality, however, is only available when the self can be modelled into past and future. If the unmodelled actual self in the present can model itself into the past or future then the modelled self can also model a self in the past or future of the modelled self; and either of the modelled selves can be conditional – models of what the self could be or could have been. From this modelling into the conditional vector comes much of our fiction – and, indeed, much of our history. We can model from known facts about the past to unknown possibilities and, if enough people adopt a particular conditional model, a consensus view can be established about what probably happened. The old Soviet epigram that “the future is certain, it is the past we cannot predict” is, for historians, all too real.

Nonetheless, narrative fiction, generated by conditionality merged with narrativity, has defined what it means to be human. We are a storytelling species (Niles, 1999), probably the only one; and, if our storytelling is indeed unique, it is another important difference between us and other animals.

How Sharing Time Became Important

To properly understand how we relate to time we need to situate temporality into the sources of language grammar story. As we have seen, this is a contentious field: theories describing the developmental speed of language grammar range from slow and gradual (*e.g.*, Tomasello, 2019) to almost instantaneous (*e.g.*, Berwick & Chomsky, 2016). Another debate is whether grammar is a product of genetically determined cognitive changes (*e.g.*, Pinker, 1994), or a behavioural response to the communicative requirements of increasingly complex social systems (*e.g.*, Knight, 2010). Together, these alternatives produce four possible scenarios: slow and genetic, slow and sociocultural, fast and genetic, or fast and sociocultural. (Whether the primary source of language grammar is cognitive or communicative is a third controversy; but this is reviewed separately in chapter 13.)

This book takes the view that language grammar developed gradually and incrementally to facilitate communication about an increasingly complex social system. It identifies four stages of grammatical development, here called Protolanguage 1 and 2 and Complex Language 1 and 2.

The term *protolanguage* is used in linguistics for two quite different types of language development. Its first usage identifies unknown early languages which have diverged into a range of known languages – for instance, Proto-Indo-European, or PIE, is thought to be the common root language from which most of the languages of Europe and South Asia developed (Anthony, 2007). In contrast, Bickerton (2014) uses the term *protolanguage* to refer to an early state of human communication, more complex than nonhuman communication but lacking an effective syntax; and this corresponds quite closely to the definition of Protolanguage 1 used here.

Protolanguage 1 reflects the one-argument grammar stage of communication – the ability to indicate both an activity and a perpetrator or target or argument for that activity. It requires two shareable lexicons, of activities and of arguments, and it permits a range of simple utterances. These include manding, or using imperatives (*Do this*); stating, or using declaratives (*Is this*); coordinating, or using interrogatives (*Is this?*); and simple descriptions or attributions (*Has this*). At a minimum, agreement and negation (*Yes* and *No*) are also needed to allow the one-argument utterances to become dialogues; and, because attracting the attention of the intended receiver of the utterance is useful, personal labels, or names, can also be seen as part of Protolanguage 1. However, it is unlikely to have been used to share temporality; even today, one-argument grammar constructs mostly signal activities happening in the present.

Protolanguage 1 cannot be directly transmuted into full language, two intermediate stages are needed. To share our cognitive social modelling we need, at minimum, to communicate two-argument forms, as discussed in chapters 7, 8 and 9; and this is what Protolanguage 2 does. A-Relationship-B grammar is propositional, it represents relationships between individuals in the group; it drives social calculus; it requires abstract cognition (models of A and B); and it is dispassionate, because the relationship between A and B is independent of my own relationships with A and B. However, the most important aspect in terms of temporality is that Protolanguage 2 allows events to be placed in relation to other events: *before* and *after* become important concepts in the individual's social calculus, and they lead on to *because-of* and *thus*, and then onto the conditionalities of *therefore*, *despite*, *yet*, and so on. The increasingly complex modelling required in the individual's social calculus starts with Protolanguage 2.

Between Protolanguage 2 and Complex Language 1 are three paths, developing self-modelling, modelling others, and abstract reference, each with three steps. This process is set out in figure 12.3 below, and as part of the more complete EAORC Routes to Language map, available at http://www.martinedwardes.me.uk/eaorc/eaorc_languageroute.html. While the three paths are described serially here, the first steps in the three paths should be treated as simultaneous, as should the second steps, and the third steps.

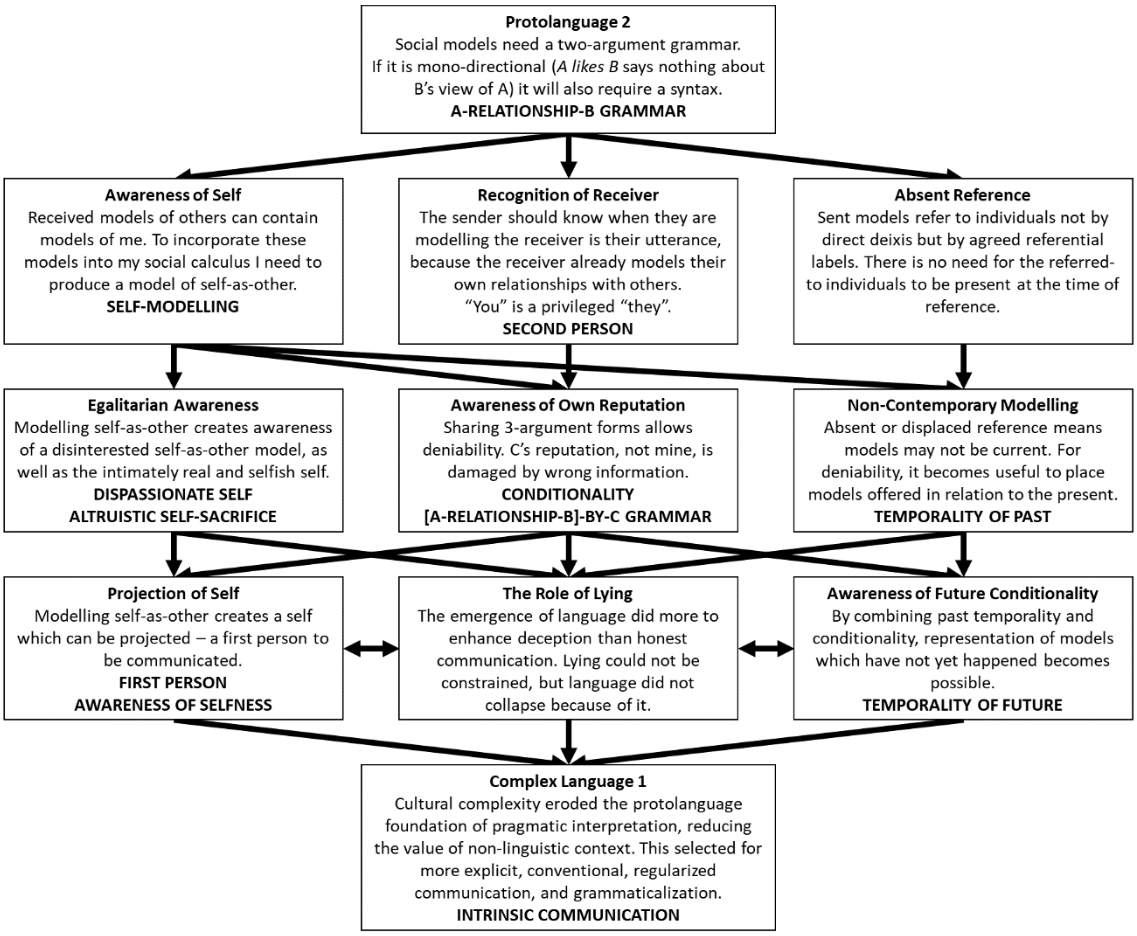

Figure 12.3: The three Routes from Protolanguage 2 to Complex Language 1

The first step on the self-modelling path is Awareness of Self (Edwardes, 2019). As described in chapter 7, the exchange of A-Relationship-B social models inevitably leads to circumstances where A or B represents the receiver of the utterance, and this generates a dilemma for the receiver: either they can ignore the utterance or they must create a model of themself as a placeholder in their social calculus. This is the self modelling the self; and, like all other models in social calculus, it is an externalised third-person model, a *me-as-them* representation which creates an awareness of self-as-other. My selfness is not just a subconscious expression of personal survival; *I* am *me*, consciously aware of myself as an objective thing. This capacity to make and project conscious models of myself seems to be impaired for people on the autistic spectrum (Frith, 2003, ch5).

Awareness of Self leads to the next step in self-modelling: Egalitarian Awareness (Engelmann & Tomasello, 2019). The modelling of self-as-other creates a dispassionate view of the self: it generates an awareness that, while I am special in my own social modelling, I am nothing special in shared social modelling. I am just another node in the social modelling of others; and, in my own social modelling, I can model myself dispassionately as a node shorn of self-interest. This makes altruistic self-sacrifice easier to contemplate: the subjective *I* making the sacrifice does not have to cognitively collocate with the objective *me* being sacrificed.

The final step in self-modelling is Projection of Self, a product of Egalitarian Awareness and Awareness of Own Reputation: when I share my model of me as A or B in an A-Relationship-B construct, I am telling the receiver how I represent myself to myself. This gives them a model of me they can reflect back to me; and, by reinforcing my own model of me, they can establish good relations between us. Projection of Self establishes a shared model of me that I endorse, and which can be used by my friends to bolster, and by my opponents to denigrate, my self-image (Henrich & Gil-White, 2001). In terms of communication, Projection of Self requires the first-person role to become a grammatical feature; and it also requires an awareness of selfness – a conscious knowledge that I am projecting a model of me which is not necessarily what I believe myself to be, it is what I want others to believe I am.

The second developmental path, modelling others, begins with Recognition of Receiver: the speaker recognises that the receiver of an A-Relationship-B utterance is a significant component of the communication process. Just as the receiver, when represented in another person's utterance, becomes both a *me* and a privileged *they* in the receiver's social calculus, so the representation of the receiver in the speaker's utterance creates *you* as a privileged *they* for the speaker. Recognition of Receiver requires the second-person role to become a grammatical feature, so the speaker can acknowledge the privileged role of the receiver in utterances (Heine & Song, 2010).

The second step in modelling others, Awareness of Own Reputation, is a product of Recognition of Receiver and Awareness of Self. The speaker becomes aware that the receiver is modelling the speaker to manage the relationship between them, and the actions and utterances of the speaker can positively or negatively affect this model. The speaker needs to cultivate this model by ensuring they are a reliable source of information (Edwardes, 2014a). There are two ways to do this, both of which mark some utterances as less reliable than others. The first method is deniability, marking some utterances as indirect, or second-hand, information; and, as discussed in chapter 9, the [A-Relationship-B]-by-C construct does this. By marking the A-Relationship-B information I am giving as indirect knowledge, I can transfer responsibility for inaccuracies to a third party, C. This, in turn, generates an iterative process where the receiver must keep track of who gave them received [A-Relationship-B]-by-C constructs by generating [[A-Relationship-B]-by-C]-by-D cognitive models.

The other method for preserving reputation is to mark utterances as speculative or hypothetical or conditional. We do this in English with a range of constructs, such as auxiliary verb forms (*can*, *may*, *might*, *etc.*), adverbials (*perhaps*, *probably*, *likely*, *etc.*), adjectivals (*uncertain*, *possible*, *occasional*, *etc.*), noun forms (*whoever*, *anyone*, *thing*, *etc.*) and grammatical constructs (the passive voice, intransitives, *etc.*). Together, these conditional indicators create a nuanced communication system where the provenance of utterances can be a more significant marker of veracity than the core statement in the utterance.

The final step in modelling others, emerging from Egalitarian Awareness, Awareness of Own Reputation, and Non-Contemporary Modelling, is the Role of Lying. This is not a default strategy, informing honestly remains the intention behind most human communication; but a capacity for judicious deployment of disinformation is an important feature of modern language use. Lying allows the speaker to decide between default honesty to preserve longer-term reputation, or dissembling to enhance short-term reputation, despite the damage to longer-term reputation. However, while language enhanced our capacity for deception at least as much as it enhanced our capacity for honest communication, the capacity to assess a speaker's reputation acts as a brake on lying – as the fable of *The Boy Who Cried Wolf* illustrates. Reputation ensures that, while deliberate deception cannot be constrained, language does not collapse because of it (Dor, 2017b).

The final developmental path is abstract reference, and the first step, Absent Reference, requires a capacity to remain aware of things which cannot be currently experienced. Our internal social modelling treats individuals who are absent the same as individuals who are currently present: with Absent Reference, out of sight no longer means out of mind. What is absent – whether real or speculative – can still be modelled, as the following poem illustrates:

> Yesterday, upon the stair,
> I met a man who wasn't there!
> He wasn't there again today,
> Oh how I wish he'd go away! (Mearns, 1899 [1922])

Because referred-to individuals no longer need to be present, Absent Reference becomes important in the sharing of A-Relationship-B models in Protolanguage 2 (Ganea, 2005). Shared A-Relationship-B models cannot usually rely on direct deixis to refer to individuals, instead, they use agreed labels, or names.

The ability to communicate about Absent Reference, when combined with Awareness of Self, generates a new class of shareable models: A-Relationship-B models in which the relationship is no longer current. Instead of up-to-date information, the speaker is sharing a memory of what used to be, a Non-Contemporary Model. This is useful for the speaker in terms of deniability: while what I am sharing is my own first-hand knowledge, it is possible that my knowledge has been superseded by events of which I have no knowledge and over which I have no control. The knowledge is being offered honestly, but it is from the past; and to understand it, both speaker and receiver must be aware of, and able to communicate about, the past. Subliminal understanding of the past is not unusual in nature; but awareness of the past as a *place* which can be modelled seems to be limited to the *Homo* clade (Dielenberg, 2013).

The final step in abstract reference is Awareness of Future Conditionality, a product of Awareness of Own Reputation and Non-Contemporary Modelling. The mechanisms for modelling the Temporality of Future require the same scepticism and deniability as the mechanisms for modelling the Temporality of Past, so the step from modelling past temporality to future temporality involves a change of perspective rather than a new cognitive mechanism (Schacter *et al.*, 2007). However, our unidirectional experience of time creates an important difference in conditionality between the two perspectives. The conditionality of past modelling comes from reinterpretation of remembered past events – our memory may be faulty but it still has a reality for us. In contrast, the conditionality of future modelling is simply speculative: we have no memory of the future, we have only the models themselves.

The steps in the three pathways all affect the form and function of language grammar. The self-modelling path gives us an objective awareness of our self, which leads to dispassionate self-modelling, altruistic self-sacrifice, and awareness of selfness; this, in turn, enables us to share our selfness with others as a self-within-an-utterance, a first-person voice. The modelling others path gives us a shareable second-person voice, conditionality, three-argument forms, and ways to dissemble without risking reputation. The abstract reference path gives us the mechanisms to share memories and plans which, because of their inherent conditionality, allow us to negotiate toward agreed meaning in the *what-if* universe as well as in the real universe; my and your memories and plans can become our memories and plans.

In their final steps, the three pathways merge into a coherent communication structure which is here called Complex Language 1; this is not a new thing emerging from the combination of the three pathways, it is merely a formalised descriptor for the three pathways working together in a systematic way. In the protolanguages there was a simple, pragmatic sharing of knowledge – a sharing continually enhanced by cumulative cultural evolution toward cooperation. However, this accumulation of cooperation was also responsible for an accumulation of cultural complexity, which eroded the pragmatic sharing behind the protolanguages, replacing them with an increasingly conventional communication system. The pragmatics of the protolanguages – negotiating toward actual facts by any means available – were replaced by an increasingly traditional and ideological interpretation of communication, which emphasised vocalisation over non-vocal content, reality over actuality, and negotiation toward meaning over negotiation toward factuality. The complexity of complex language was in part a response to the need to share complex cognition, but it was also a response to a need for more explicit, conventional and standardised communication (Sterelny, 2016). However, the standardisation conformed to local cultural norms rather than genetically-moderated species norms, quickly creating a range of local complex languages rather than a panhuman ur-language. Just as meanings in complex languages are usually intrinsic (they are contained in the definition of the language, not defined by the universe outside the language), so the grammar systems in complex languages are also usually intrinsic.

From Complex Language 1 to Complex Language 2

Complex Language 1 is, in many ways, identical to modern human language, and it can be used for almost all the social conventions to which modern human language is applied. One of these is Shared Storytelling and Fiction, the process of turning factual lies into non-factual truths. Among the Ju/'hoan (!Kung) Bushman tribe of Southern Africa this process has been further conventionalised, with daytime discourse being devoted to practical issues, such as economic and political discussions and the sharing of gossip, and night-time discourse being devoted to singing, dancing, religious ceremonies, and stories (Wiessner, 2014). Listening to stories allows the listener to adopt the perspectives of the protagonists in the stories, viewpoints that experience alone cannot provide (Krieken & Sanders, 2021). Having a good storyteller in your group improves socialisation by enhancing empathy and cooperation; and, if being a good storyteller has a genetic component, then a preference for storytellers as sexual partners would ensure that storyteller genes spread through a population (Smith *et al.*, 2017).

Combined with Egalitarian Awareness, Shared Storytelling and Fiction allow us to describe the actual world in virtual terms, building a shared and negotiated reality which defines *our group* and differentiates it from *their group*. Some of the stories on which we base our lives become more than just agreed truths, they become moral systems that demand the adherence of the individual, and which sanction punishments for non-adhering individuals. These stories make human communication unlike other communication systems: we exchange deceptive information to construct our sociocultural environment, not to subvert it.

An example of sociocultural construction based on story-telling is the Female Kin Coalition (FKC) proposed by Knight (1991) and set out in a succinct form by Knight *et al.* (1995). The FKC describes a female-led culture which goes through a monthly cycle, tied to the lunar cycle, of gender sharing and gender distancing: females live with their spouses for the two weeks of waning moon, and for the two weeks of waxing Moon the females form a solidarity with their brothers and fathers to exclude the spouses. The period of Dark Moon marks the separation from spouses; and the period of Full Moon marks a (hopefully) successful hunt by the spouses, followed by a feast and two weeks of conjugal relations. The monthly cycle is intended to line up with the female menstrual cycle, so that menstruation occurs at Dark Moon, followed by the follicular phase and ovulation, followed by the fertile or luteal phase starting at Full Moon. The FKC culture is symbolic, embodying a shared fiction about the way the universe works. The fiction is based upon an actual fact (fertility), but it is interpreted as a transactional relationship which cyclically transforms the symbolic nature of the people in the culture – from wives to monsters and from husbands to strangers, and back again. The FKC cultural model has been identified at work in many traditional gatherer-hunter societies, and it has left traces in the cultures of many less-traditional societies.[11]

It would seem that Complex Language 1 and Shared Storytelling and Fiction give us all that we need to define human language; but there remains one thing unaccounted for: recursion. As we saw in chapter 3, the generative linguistics Minimalist Program identifies recursion as the only significant difference between nonhuman communication and human language: all the features of modern human language can be traced back to proto-forms in the communication systems of other species – except the cognitive capacity for recursion (Hauser *et al.*, 2002). In the hypothesis presented here, Recursive Social Cognition is needed if [A-Relationship-B]-by-C constructs are being shared, because the receiver of the model must tag the offered construct in their social calculus with the label or name of the speaker offering the model, giving a recursive cognitive construct of [[A-Relationship-B]-by-C]-by-D.

Corballis (2007) identifies recursion at work in several cognitive systems as well as language: theory of mind, mental time-travel, selfhood, counting, and using tools to make tools. These are all examples of cognitive recursion limited to a single brain and requiring no conscious negotiation toward meaning with another brain. It may seem trivial to insist that recursion was initially a cognitive capacity, but it must have a pre-linguistic cognitive origin and be common to both brains if it is to be communicated and shared between those brains. Corballis' choice of cognitive origin, Theory of Mind, is closely aligned to the social modelling hypothesis proposed here.

Shipton (2019) sees recursion as part of a three-pronged development process, involving normativity (the group culture that dictates human relationships beyond one-on-one interactions), recursion, and abstraction. Recursion is placed in the middle of this process because it is viewed as a precursor for abstraction: Shipton says, "Middle Palaeolithic hominins seem to have sporadically expressed the ability to think abstractly; however, this

[11] See https://www.facebook.com/RadicalAnththropologyGroup/ for more on FKC theory, or Chris Knight's website, http://www.chrisknight.co.uk/.

ability is not consistently manifested until the late Palaeolithic after 70,000 years ago" (164-165). Here, we are concerned with the development of the cognitive capacity for abstraction rather than its application, which places it before recursion.

With the sharing of recursive information, Complex Language 1 develops all the attributes of modern language and becomes Complex Language 2. Chomsky describes the capacity for sharing recursion in language as Merge (Chomsky, 2007), and argues that this capacity must have emerged relatively recently in *Homo sapiens*. He sees it as the product of a single mutation, often called a macromutation because its effect upon the phenotype is extensive; and that raises problems, because beneficial macromutations are rare, although not impossible (Futuyma, 2015; Wood & Grabowski, 2015).

In a recent discussion, Martins & Boeckx (2019) have argued that treating Merge as a macromutation is unnecessary, because the idea that there can be no intermediate steps between no-Merge and full-Merge is itself fallacious. Berwick & Chomsky (2019) have doubled down on their claims that there is no half-Merge, but the two positions may be arguing from different premises. For Martins & Boeckx, the infinite applicability of Merge could have been preceded by a delimited form of Merge – or even several, becoming less delimited with each iteration. For Berwick and Chomsky, Merge is, by its nature, an instantiation of Humboldt's (1836 [1999], 91) view of language: "from limited resources it must therefore generate unlimited usage"[12]. Merge cannot be delimited because it is unlimited. This position, though, seems to use the definition of language to define Merge, while assuming that Merge can then be used to define language.

Merge allows structures to be reused within structures, and turns base grammatical forms (nouns, verbs, *etc.*) into special cases of phrasal forms (noun phrases, verb phrases, *etc.*). Outside of theory, however, grammatical hierarchy has no need to be unlimited. Unlimited usage remains a principle

[12] My translation of "Sie muss daher von endlichen Mitteln einen unendlichen Gebrauch machen".

and not a practice because, in the actual world, no unlimited language construct could ever be produced; and in actual language, grammatical limitations to limitlessness are constantly being identified. For instance, *[the boy [the girl hugged] cried]* is grammatically acceptable, as is *[the girl [the man saw] hugged the boy]*; but, while **[the boy [the girl [the man saw] hugged] cried]* is grammatically correct according to the rules used to generate the first two sentences, it is not really grammatically acceptable.

Another problem is that, while we can construct complexity, we cannot always understand the complexity we construct. For instance, one task set on the British TV Channel 4 *Taskmaster* programme[13] was: "Do the opposite of the following: you must, under no circumstances, not avoid not making the bell not ring." So should you ring the bell or not? Of the five contestants, three guessed and got it right, one guessed and got it wrong, and one worked through the nest of negatives – and got it wrong.

One of my favourite examples of complexity is: "Is it immediately obvious to you whether it is false or true to deceitfully say that it would be an erroneous indicator of fabrication to imply the fiction that George Washington could not tell a lie is untrue?" The simple answer is no, it is not immediately obvious to me; but whether the sentence is false or true has so far not become even eventually obvious to me. This is the key problem with complexity, and therefore with Merge: yes, it is probably present in no other species' communication; but is what it adds to Complex Language 1 significant enough to make Complex Language 2 decisively different? Recursion is a useful addition to language grammar, but it is not often needed, so not often encountered. Satik (2022) indicates that a strong Minimalist theory, in which all language grammar can be traced back to Merge, does not accord with the evidence of syntactic variation from case and agreement in some languages. Instead, he proposes a weak Minimalist theory where some syntactic forms are non-universal and do not trace back to Merge; they are proprietary to individual languages.

[13] Taskmaster series 12, episode 8, first task. Can be viewed at https://www.comedy.co.uk/tv/taskmaster/episodes/12/8/ at time of publication.

According to Hurford (2008), modern languages have a range of ten attributes, one of which can be found in nonhuman communication, some of which have precursors in nonhuman communication, and some of which occur only in human languages. The single attribute shared with nonhuman communication is the **Interpersonal Function**, which shows that all signalling carries social force: it is done to inform or deceive others.

There are three modern language attributes which are present as precursors in nonhuman communication. The first is **Stimulus-freedom**: while some nonhumans can have a simple memory of the past or even expectations about the future, only humans indulge in conditional thinking about *what-if*. The second attribute is **Mindreading, Manipulation and Cooperation**: while some nonhumans can anticipate the physical responses of others, only humans seem to see others as intentional beings. The third precursor attribute is **Reference**: while many nonhumans use alarm calls, some of which are specific to the type of threat, only humans seem to indulge in dialogues in which the speaker, the listeners and third persons can all be represented.

This leaves six attributes which are unique to human languages. The first is **Diversity**: humans have culturally defined languages which are different enough from each other to be mutually cryptic. The second is **Learning**: human languages must be learned, although they are constrained by biological factors like memory. Next are **Complexity**: human languages are complex, rule-driven and formal; and **Compositionality**: human languages can convey meaning on several levels (words, phrases, sentences, paragraphs, *etc.*). The compositional meaning is in the construction of the utterance, which must meet the intention of the speaker and attract the attention of the listener. With compositionality, the signal is a vehicle for intrinsic meaning rather extrinsically meaningful itself. The fifth attribute is **Double Articulation**: languages rely on essentially meaningless sounds which are combined into meaningful constructs; and the final attribute is **Self-organisation**: meaning structures and grammar constructs are organised by the interactions between language users as they negotiate toward agreed meaning during the discourse.

In regard to recursion, Hurford says:

> Hauser *et al.* leave it open whether such a capacity for recursion can be found in any non-human animals. If it can be, then the human faculty for language in the narrow sense is, in their view, actually empty, leaving us with a picture of FLB as a mosaic of factors all of which can be found in some form or other outside of the domain of human Language … The technical definition of recursion, and how to recognize whether it is in play in a specific animal activity, is not, however, satisfactorily pinned down, and there is room for argument about the use of recursion in animal activities (254).

The failure of Merge to provide a complete explanation for Complex Language 2 is why a social role for communicative complexity (the Female Kin Coalition theory used here) has been vectored into the sources of Complex Language 2. As we saw in chapter 2, humans have a capacity to treat their social interactions as games, particularly so when language is involved; and this, by itself, may explain a large part of the complexity in Complex Language 2. If, instead of Reichenbach's model of tenses, we look at Halliday & Matthiessen's model (2004, 340-342), we see a list of no less than 36 English tenses of increasing complexity. A study of the actual usage of the different verb forms on the Internet (Edwardes, 2011) showed that the last four tenses listed, *will have been going to be [verbing], had been going to have been [verbing], has/have been going to have been [verbing]* and *will have been going to have been [verbing]*, with 18, 16, 11 and 20 occurrences respectively, only barely registered in a dataset conservatively estimated to be about 4.8 trillion words and containing about 10 billion verb forms. Where they were used, 40 out of the 65 occurrences were on sites teaching English grammar. As the Taskmaster and George Washington stories above show, we play with language complexity; we also play with phonology (*e.g.*, Dorothy Parker's "I'd rather have a bottle in front o' me than a frontal lobotomy"); we play with meanings (*e.g.*, metaphors); we play with intentions (*e.g.*, sarcasm and irony); we play with names (my current favourite is Bogus Piffle Johnson) … basically, if we can say it, we can play with it. Perhaps what is missing in our understanding of language grammar sources is not a careful analysis of structure, but a more open approach to the crumbly edges of information exchange.

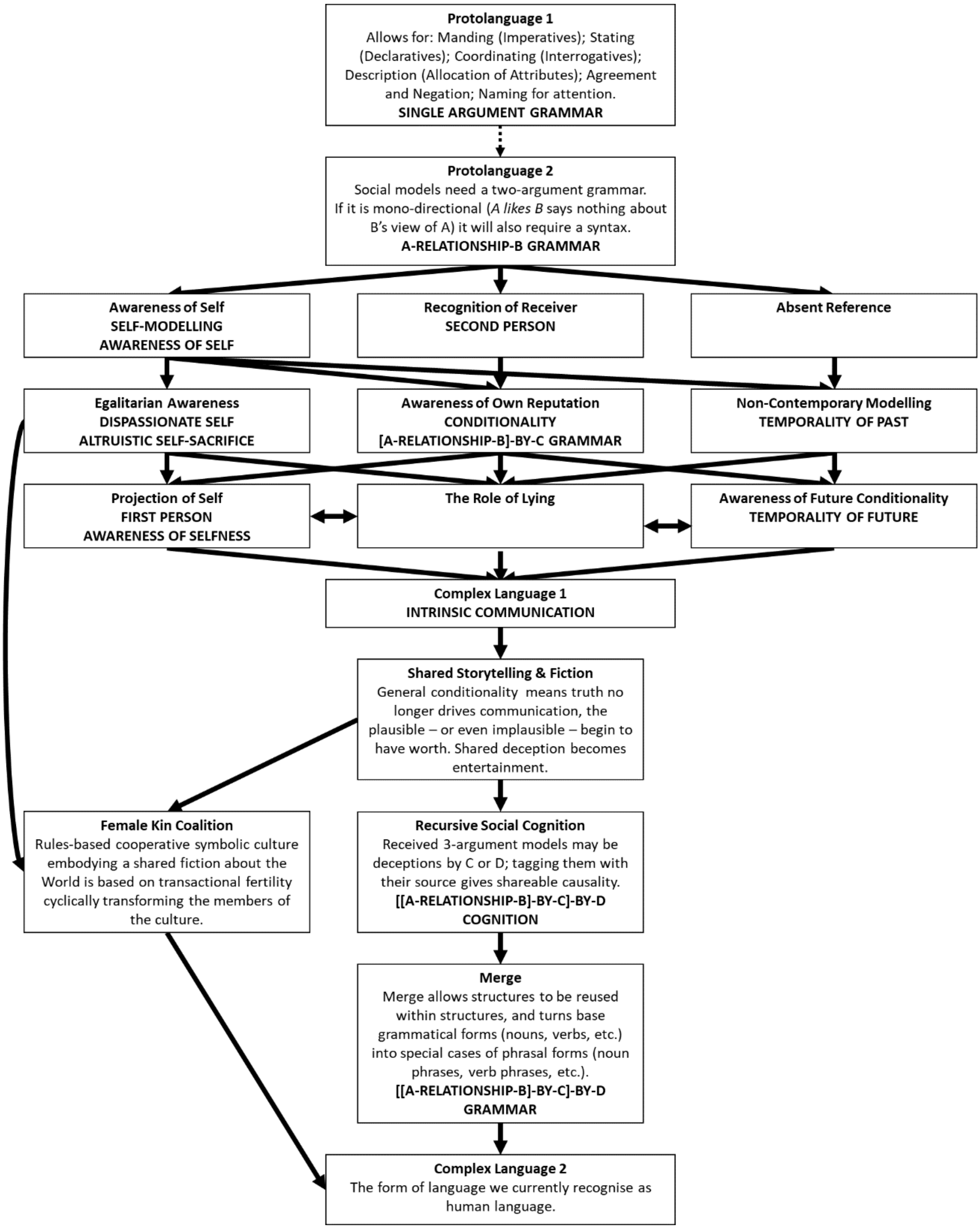

Figure 12.4: A possible route from Protolanguage to Complex Language

How Children Become Time-aware

Human children are not born with a fully competent memory system: before age four they lay down very few, if any, long-term memories, and do not have an autobiographic narrative of their life (Hayne *et al.*, 2015).

Instead, they seem to divide things into *now* and *not-now*, and events in not-now are considerably less valuable than now events. In a televised experiment based on the famous marshmallow experiment (Mischel & Ebbesen, 1970), children of various ages were asked to choose between having a single piece of chocolate now or a bar of chocolate in 10 minutes: the under-fours universally opted for the small piece now, while the over-fours universally opted to wait (Brewer, 2001, 179-183). The under-fours seemed unable to project a modelled self into the future to accept the whole bar, all they had was the unmodelled actual self in the present to accept the single piece.

Children do not achieve an instantaneous comprehension of temporality, their understanding grows incrementally, and their use of it in language is progressive. For first-language English learners, Weist (1986) describes four stages: for the first 18 months the child references current events; from 18-36 months they refer to past and future events using simple tenses (*-ed* morpheme and auxiliary *will*); from 36-48 months they add adverbials into their temporality, but only where they agree with the simple tense (*e.g.*, *We went to the park yesterday*); and from 48 months they begin to use complex tenses involving a separate Point of Reference. However, nobody expects children to meet this timetable exactly; and Shirai & Miyata (2006) show that, as well as individual differences, the language being learned affects the timetable. This schedule does indicate, however, that self-modelling, needed to use Reichenbach's Point of Reference effectively, is not available at birth.

Acquisition of temporality is not entirely systematic: tense seems to be acquired to one timetable, aspect (continuity and imminence) to another, and the order of acquisition can be idiosyncratic (Valian, 2006). Aspect and the nature of the action are also linked, with the perfective (non-continuous) past tense being used by young children for state-changing verbs, such as *stop*, *hit*, *give*; and the imperfective being used for activity verbs, such as *playing*, *doing* and *making*. Under age 4, children seem to find future tenses harder to understand than past tenses, and they also seem to process near-future events better than far-future events (Friedman, 2000). Additionally, many young children tend to use the *gonna* future tense only for their own plans, indicating other future events with *will* (Tomasello, 2003a, 217-224). Clark (2003, 258-261) shows that young children tend to treat events mentioned first as occurring first; so sentences like *he leapt after looking* and *before he leapt he looked* tend to be misordered at 36 months, although by 54 months this has become rare.

A study of four Japanese children found that, where the action of the verb had changed the described state so that the description was no longer true, the children initially used the past tense suffix *-ta* (*e.g., Alf dropped [the] book: Arufu wa hon o otoshi**ta***); where no change of state occurred, the present tense was used (*e.g., Alf holds [the] book: Arufu wa hon o motte imasu*) (Shirai & Miyata, 2006). The children's intended meaning was not about when the action took place (deictic use), it was about the current effect of the action (contrastive use). While contrastive use was first identified between 15-22 months, deictic use was not identified until 19-27 months, a gap in all cases of between 2-5 months. Although the study involved only one language, Shirai & Miyata suggest this effect is likely to be present in other languages, although deictic tense is probably easier to acquire in German and English than in French and Italian because the past participle is used as both a contrastive and a deictic marker in the Germanic languages.

How children acquire tense and aspect seems, therefore, more complex than just learning the rules, it relies on several things: how the action of the verb is viewed, the perceived difference between past and future, and reasonable but sometimes faulty syntactic hypothesising. Learning temporality in language relies on understanding the speaker's meaning and intention, the capacity to model events into the past and the future, and the flexibility of self-modelling in terms of recursive modelling.

Three Time Points, Three Persons?

Do Reichenbach's three points (Point of Speech, Point of Event, Point of Reference) correspond to the three persons in language? These persons are the self (me, the sender), the directly addressable non-self (you, the receiver), and the non-self that is not directly addressable (them, the referenced objects); Bloomfield (1933, 224) refers to these as “speaker, hearer and third person”, but they are more usually referred to as first, second and third persons. The three persons meet the need to explicitly reference the sender and receiver of a signal, and to separately identify the subject matter of the signal itself (Benveniste, 1970, 217-221). Every human language so far recorded uses a special class of words to indicate these roles, usually described as pronouns (or, more correctly, pronominals). A language often has its own characteristic pronouns, such as personal and impersonal first person (Modern English *I* and *one*), formal and informal second person (Middle English *you* and *thou*) and differentiation between human and nonhuman in the third person (contemporary English singular forms *they* versus *it*); but all pronominals indicate one of the three persons, sender, receiver or referent. Pronouns also commonly have singular and plural forms, although the range of plurals in each person varies between languages.

Language utterances are usually tailored to meet what the sender believes the receiver wants to hear: we address people using registers appropriate for the listener and the circumstances, something not identified in other animal signalling. In fact, most human communication is mediated by the sender’s modelling of the intended receivers of the signal: the capacity to model self and others is not just another way of being human, it is closely implicated in how we use language.

If we compare the three persons and Reichenbach’s points there does seem to be some correspondence. The sender is always at the Point of Speech, in the present; and the referent is always at the Point of Event. The receiver, however, is not so easily placed: they are in the present with the sender for speech (although in the future for writing), and the sender invites them to model the points of reference and event in the utterance. The receiver and the Point of Reference do share one feature in the mind of the sender: they are both models able to contain other models. The sender negotiates a shared model of the self at the Point of Reference, so that the modelled self can model the Point of Event. This negotiation toward modelling at the Point of Reference associates the receiver with the Point of Reference. The correspondence between Reichenbach’s points and the three persons is therefore not exact, but it does demonstrate that self-modelling could be behind both processes.

The three persons (I, you and they), three objects of a signal (sender, receiver and referent), and three points in Reichenbach's tense model (points of speech, reference and event) seem to be related, but there is no good evidence that they actually are. Correlation should not imply causation, and similarity is not identity. The correspondences identified here should be considered, for now, as metaphorically, but not actually, linked.

Time, Planning, and Being Human

Linguistic temporality is a complex set of functions for describing Point of Reference, Point of Event, continuity, imminence, connectivity and conditionality. English expresses temporality with inflection (*-ed*, *-ing*), auxiliary verbs (*was*, *have*, *will*, *going to*), and time-related words (*after*, *yesterday*, *tonight*, *eventually*, *ago*, and so on). Together, these functions express a wide range of simple and complex temporal relationships. Different languages express temporality in different ways, but behind the variations is the simple constraint of what we need to communicate – the noun-verb distinction, the three-argument form (*subject/verb/object/indirect object*), and the modelling of self and other.

Temporality is the expression in language of our experience of time. We do not have to view everything from the present because we can model ourselves; and by modelling ourselves as future or past entities we make planning and reviewing our actions possible. We can then predict the likely outcomes of our plans, or understand the actual outcomes, by modelling ourself as modelling itself. Unlike our unmodelled actual selves, our self-modelling is not limited to the present.

The self-awareness developmental path gave us nested iteration, the capacity to model a self modelling a self modelling a self. Sharing models of self and others relies on, and enhances, our already-exceptional level of cooperation, giving us access to the Plan-Execute-Review mimetic cycle

(Donald, 2001, 269-271). This, in turn, lets us plan cooperatively; and the enhanced cooperation let us use the Plan-Execute-Review cycle as a shareable experience as well as an internal, personal one. With Plan-Execute-Review we are able to share our plans and work together in new ways, developing more complex socialisation and cooperation. Simultaneously, the abstract reference developmental path let us project our first-level point-of-speech models into our past and future to create second-level point-of-reference models, and then project the second-level models into their own pasts or futures to create third-level point-of-event models. Language is enabled by, and enables, exceptional cooperation; and it is this extreme cooperation which creates the willingness to give truthful messages and receive less-than-truthful messages. Robert Burns was almost right when he said in his poem, *To a Louse*: "O wad some Pow'r the giftie gie us, to see oursels as others see us!" But it is actually seeing ourselves as we see others that gives us fully complex language.

Chomsky sees recursion as the key difference between human and other minds (Hauser *et al*., 2002). However, as Dickins (2003) points out, recursion cannot stand alone, it must operate in a system that tolerates recursion. It requires exchangeable tokens or symbols, a process of exchange which transmits meaning as well as symbols, and a structure where form can recur without meaning recurring. If the hypothesis presented here is correct, recursion emerges from the iteration of modelled selves, a property exemplified in grammatical temporality: the Russian dolls of self-modelled-within-self are recursive and potentially infinite – although actually profoundly constrained; and the intentionality-within-intentionality they create forms the basic structure for recursion throughout language (Dunbar, 2004, 47-69). Recursion, therefore, cannot be the sole source of language; instead, social modelling, a product of sharing multi-argument models through language, is the source of recursion.

13 The Sources of Language Grammar

There is no better tool for writing than experience. It has very little to do with grammar and everything to do with knowing.
(A. A. Gill)

This book argues that cognitively modelling self and others in a social setting provided a template for language grammar. However, it unusually suggests that modelling others is not an outcome of self-modelling; instead, self-modelling is an outcome of modelling others. although unusual, this approach is not without precedent: Nichols & Stich (2003) show that Theory of Mind does not require awareness of self, and that awareness of other and awareness of self may be the product of two distinct processes; and Carruthers (2009) shows that, in evolutionary terms, it is likely that mindreading (required for modelling others) preceded metacognition (the capacity to think about thinking, required for self-modelling).

Self-modelling and awareness of self both require metacognition; but sense of self does not. Subconsciously dividing the universe into self and not-self has its own fitness advantages, so would have been selected for by evolution. Yet even here, it is better to be aware of the not-self rather than the self: autonomic control reliably ensures the self operates optimally without conscious attention; it is the uncontrolled rest of the universe that needs active, and sometimes conscious, intervention. The ability to consciously model others, despite having a large cognitive cost, is fundamentally useful: it provides effective objective information which lets us anticipate the actions of others. Self-modelling provides subjective information which does not need the same level of accuracy: it is often evolutionarily fitter to have an inaccurate positive self-image than an accurate negative one (Epley & Whitchurch, 2008).

In the EAORC Routes to Language map, a series of developmental events have been plotted into a time series. We are, however, still in the early stages of establishing a comprehensive timeline for the development of all the necessary events, effects, links, and the intervals occurring between them (see http://www.martinedwardes.me.uk/eaorc/eaorc_languageroute.html). Nonetheless, the map does indicate that both self-modelling and modelling others are heavily implicated in sharing social calculus, and in several grammar features. The EAORC Routes to Language map is an ongoing living project, but it already tells a coherent tale of how language grammar came about.

Signalling before Language

It is reasonable to presume that signalling systems currently used by other primates were within the capacity of Australopithecus, 4mya; and that offers quite a wide range of signals, including simple manding, requesting conspecifics to change or prolong their behaviour. This type of primate signal is common; for instance, the chimpanzee lip-smack, used to encourage or prolong mutual grooming sessions, has also been implicated in the origins of human language (Fedurek *et al.*, 2015). Primate manding does not provide a generalised system for request signalling, it is a closed set of calls for specific actions which varies between species. Also, while dominance mands are often seen as exemplars of manding, submission mands are more common. Alarm calls, more complex mands used by many primate species, instruct receivers to carry out a particular course of action as well as referring to the cause of the warning (Cheney & Seyfarth, 1990, ch5). While these mands are not semantic in the strict linguistic sense (they are not symbolic and represent events categorically rather than conventionally – Rendall, 2021), they do have meaning to both the sender and receiver, and often cause the receiver to begin or end an activity. The prevalence of dominance and submission signals, courtship signals, territorial calls, threat calls and alarm calls means manding signals are fundamental to many nonhuman signalling systems (Bradbury & Vehrencamp, 1998, ch18).

Australopithecines probably also had a limited capacity to produce and understand segmented signals – signals containing more than one meaning-unit. Male diana monkeys (*Cercopithecus diana*) use different warning calls to indicate threats from leopards and eagles; but they also modulate the start of the call to indicate the direction and distance of the predator – and therefore, the level of threat (Zuberbühler, 2000). This modulation segments the call, separately identifying the predator and the threat level. Similarly, apes taught ways to communicate with humans have no problem segmenting their constructs into actions and objects; and Kanzi the bonobo seems able to apply syntax, making the order of segments significant (Savage-Rumbaugh & Lewin, 1994). Signal segmentation must be already available to wild chimpanzees and bonobos if human acculturation can make it consciously accessible for cognitive manipulation and signalling. It likely has a pre-existing cognitive role, probably for managing simple social calculus.

The segmentation into actions and objects indicates the *Pan* species must understand the basic typological distinction that some segments represent things and others represent actions done to things. Language experiments also indicate they understand comparators like *same* and *different*, colours, and descriptors like *sliced*; they correctly interpret adpositions like *on*, *under*, *above*, *below*; and they display some comprehension of conditionals, such as "*if* this happens *then* that happens" (Premack & Premack, 1983, ch8). In language terms we would say they apprehend the grammatical concept of word classes; but they are not necessarily using conventionalised rule-driven knowledge to do this. The word classes may be used appropriately because the words around them promote their correct use and suppress their incorrect use; word class is not driving usage, it is the word-forms themselves.

Grooming is another signalling system available to *Pan* species and therefore likely to Australopithecines (Dunbar, 1996). It is vital to the longer-term process of forming and maintaining alliances, and it is clearly pleasurable for both groomed and groomer. Mutual grooming is both interpersonal and interactive, and common in primates. It creates extended dialogues which are in many ways language-like, so an important precursor for language – and language, like grooming, is closely tied to the dopaminergic reward system which causes individuals to seek out opportunities to generate personal pleasure. The dopaminergic reward system is linked to many prosocial behaviours in modern humans (Arias-Carrión, *et al.*, 2010).

Manding (known in language as the imperative), segmentation, differentiation and grooming provide a rich environment for the development of hominin signalling; and, even if australopithecines had only a limited range of vocalisations, segmentation and differentiation supported by gesture can still generate a quite sophisticated communication system. We cannot know whether these capacities became actual communicative realities; but we can say that relatively complex meaning-making, even if non-symbolic and categorical rather than conventional, was within the competence of early *Homo*.

Social Modelling before Sharing

The appearance of A-Relationship-B modelling in social cognition let humans keep track of the interactions in their medium-sized groups. Representational tokens – labels or names – must have been available, as well as the capacity to use segmentation and differentiation in their cognitive constructs. A-Relationship-B modelling is more complex than the Relationship-A modelling used in Machiavellian social environments, allowing individuals to enhance their personal fitness with interpersonal alliances. Each primary relationship with another individual becomes a secondary relationship with their allies, making alliances vital for survival and successful reproduction. A-Relationship-B social modelling is important for any species composed of individuals living in moderately sized groups with high levels of social interaction.

There are two ways of cognitively encoding A-Relationship-B model sets: each set can be holistically encoded, in which case each set must be unambiguously labelled; or a label can be allocated to each individual and to each relationship between individuals. While the labels for the individuals must be unambiguous, the relationship labels can reflect both similarity and difference – they do not need to be identities, they can be metaphors. Each model set is a segmented meta-token with the same information content as its holistic equivalent, but the use of the same individual label in several sets creates a network of interpolated information: A's relationship with B and A's relationship with C tells me something about B's relationship with C. However, it does need new cognitive mechanisms to handle the network.

Relationship-A modelling identifies relations between my unmodelled actual self and others; and, as my relationship with each other individual is intrinsic to my model of them, it does not need segmented tokens. Simple reactions triggered by the presence of the individual greatly reduce cognitive load, which makes Relationship-A modelling very different from A-Relationship-B modelling. In Relationship-A modelling, the relationship is my intuitive reaction to A; in A-Relationship-B modelling, the relationship represents an abstract emotional state unrelated to my own reactions toward A and B. The emotion in A-Relationship-B forms should represent, not elicit, the emotion.

In contrast, I can have an intuitive reaction to the A-Relationship-B meta-token similar to my separate reactions to A and B. For instance, I may like A and dislike B, so I want to disrupt the strong relationship between A and B to establish a stronger relationship with A. This multi-level modelling introduces hierarchy and social calculus to cognition: where holistic modelling is unable to identify the role of B in the sets A-Relationship-B, B-Relationship-A and B-Relationship-C, segmented modelling makes it clear, letting the modeller compute A's likely relationship to C. The more known about relationships between individuals, the more accurate are the models made about unknown relationships. In a highly socialised context, social knowledge is power; so accumulating social knowledge makes an individual, in evolutionary terms, fitter. Although A-Relationship-B calculus has a cognitive overhead, its advantages for social interaction and social manipulation more than offset the cost.

Social Modelling after Sharing

The hypothesis of this book is that communicating two-argument A-Relationship-B social models became advantageous, creating Protolanguage 2. How this came about is still somewhat vague, but several key steps have be identified. Early humans, probably *Homo erectus*, lived in groups of 100 or more individuals (Aiello & Dunbar, 1993), which made social knowledge particularly valuable; and, because of the size of these groups, the level of genetic kinship was probably relatively low. There would have been a high level of cooperation in hunting and in most other activities (which Protolanguage 2 would have enhanced), and this would have encouraged both a reduced level of interpersonal conflict and a significant level of altruistic punishment. Fehr & Gächter (2002) show that humans have a natural aversion to freeloading, and we cooperatively punish it without close attention to personal cost.

Early humans are likely to have had individual specialisations, with those skilled at one thing able to swap their surplus product for the surpluses of other specialists. A primitive economy, policed by the altruistic punishment of economic freeloaders and renegades, emerges from this culture (Henrich & Boyd, 2008). It is unlikely to have been a barter economy involving negotiation, compromise, agreement and exchange, it more likely followed the principles of gifting described by Mauss (1950): a network of gifting which creates obligations for reciprocity. Each gift contains within it the suggestion of *your turn next*, which the self-contained micro-transactions of modern economics tries to ignore – perhaps foolishly. To see this gifting culture as primitive communism or primitive capitalism is to misunderstand it: it is not transactional, it is a two-way flow of moral debt and repayment.

Erdal & Whiten (1994) offer one source for this gifting culture. Their Vigilant Sharing model proposes that early humans, living cooperatively in large groups, developed a belief in their entitlement to a fair share. Acting on this belief helped to suppress greed and encouraged the low-cost altruistic punishment of withdrawing cooperation; which, if enough individuals punished perpetrators that way, became a costly group ostracism for the greedy. Boehm (1999) offers another source for the gifting culture: Reverse

Dominance, or collective action against controlling individuals. Here, group members work together to suppress alpha behaviours, creating a cultural environment where individual modesty and group achievement are valued. Once again, altruistic punishment keeps renegades in line, but it is administered cooperatively by the group rather than individually. Reverse Dominance was supported by the Political Singularity, which provided the group with cheap methods of altruistic punishment (Dessalles, 2014a): basically, tools for killing prey work equally well against bullies. Vigilant Sharing and Reverse Dominance are not mutually exclusive, and both must have occurred in early human culture.

The A-Relationship-B model requires segmentation and differentiation, and it enables hierarchy; so, when uttered, it would become the first grammatical form to be communicated. Initially, signalling this message-form would have been laborious, probably involving both gesture and vocalisation: the receiver must understand that the sender is attempting to share a two-argument cognitive model, and careful negotiation would be required by both parties to ensure the intended message was received (Scott-Phillips, 2010). However, once the initial negotiation toward an agreed form was completed, subsequent utterances could be quicker and simpler – it would advantage both parties to negotiate toward reduced signal cost by emphasising salience over detail. Both sender and receiver have an interest in ensuring faithful transmission, which creates a strong pressure toward simplification.

Two-argument A-Relationship-B forms would have been shared in large and socially equal groups where individual specialisation and gifting were practiced. This environment would make social knowledge valuable; but the group size would mean each individual must either spend time gathering the information needed to fully map their social group, or they must settle for a partial map. Individuals who shared their own social map honestly would enhance the social knowledge of others, creating better group consensus and cohesion.

Would this sharing enhance the sharer's fitness, though? It would certainly make them an ally worth cultivating; and altruistically sharing information would, in a reciprocal gift-driven environment, create the need for reciprocity. Also, sharing social knowledge is relatively cheap for the sharer but valuable to the receiver; so sharing should show a *profit*, creating obligations which are cheap for the receiver to fulfil but valuable to the sharer. Whether a genetic change or communication of a pre-existing cognitive system, sharing two-argument social models enhances group fitness; so it would spread through the group where it began, and then beyond, by inheritance or behavioural adoption. This model-sharing only works if what is shared is sufficiently honest; and reputation is what keeps it honest (Fitch & Hauser, 2003). Offering false information leads, at a minimum, to the sender's signalling being discounted, creating no obligation for the receiver; so long-term loss of credibility offsets any short-term advantages in lying. Basically, lying reduces individual fitness, being honest improves it.

Runaway Complexity

As well as bringing grammar into signalling, the two-argument form allowed sharing of noncurrent events: the A-Relationship-B message is, essentially, *telling-about* rather than instructing, because the receiver cannot immediately verify the signalled relationship between A and B, they rely on a pre-existing level of trust. The receiver accepts the signal because the sender, not the signal, has a trustworthy reputation. The signalling process involves both informational and interpersonal interaction, as Halliday's Systemic-Functional metafunctions describe (Halliday & Matthiessen, 2004): the signal is no longer a product of just the text (textual metafunction), it involves the sender and receiver as people (interpersonal metafunction) and highlights the contextual connotations surrounding the production of the utterance (ideational metafunction). Telling-about turns a signalling environment into a pervasive communication system linking individual utterances into narratives.

Introducing narrative into signalling also changes the modelling environment. As well as two-argument forms, it becomes advantageous to model three-argument forms, allowing received A-Relationship-B models to be tagged with the identity of the sender and generating hierarchical [A-

Relationship-B]-by-C models. This allows the author's trustworthiness to be factored into received two-argument forms; but it requires social modelling to become iterative, modelling the individuals in the message (A and B) within a model of the sender (C). This creates an extra cognitive overhead, so is unlikely to have emerged with the first utterances of A-Relationship-B messages. The iteration is, however, not recursive at this stage: while the modelled sender contains models of others, those models of others are not themselves model-makers.

When the cognitive [A-Relationship-B]-by-C model is communicated, it becomes useful for the receiver to tag the received model with the identity of the sender, creating nested hierarchies of [[A-Relationship-B]-by-C]-by-D. This four-argument form contains true recursion: the receiver's model contains D modelling C modelling A and B, an iteration of a modelled model-maker (D) modelling a model-maker (C) (Hofstadter, 2007). This recursive iteration is theoretically infinite; but in practice it is quite limited, with average humans able to effectively compute only three or four levels. Yet even this limited recursive capacity has not been observed in any other primate (Dunbar, 2004, ch3).

The telling-about of A-Relationship-B modelling is likely to have initiated various new grammatical tools. Differentiating the people represented in the shared social model can be done with simple labelling, or naming, and cognitively tagging people and things seems to precede the hominid clade; but sharing those tags as names would have required (and still does require) careful negotiation toward meaning. The individuals in the models being shared are often not available for direct reference, such as pointing; and, even where they are present, discretion may require a more indirect form of reference. Telling-about therefore benefits from ways of describing both the individuals being talked about and their relationships, introducing grammatical roles such as adjectives, adverbs, determiners and negators into communication. These descriptors, which probably developed slowly over many generations, vastly expanded the range of narratives possible, and eased the path to Complex Language 1 and Shared Storytelling and Fiction.

Grammar Universals in the Toolbox, Not the Rulebook

Generative linguists believe that grammar has a single set of universal principles which are innate in all humans. All languages therefore rely on, and contain, mechanisms that are generated by a language organ which provides a unified structure to human language; and this language organ is a product of a specific, language-related genetic mutation which occurred only in *Homo sapiens* (Chomsky, 2002, 64). There certainly seem to be universal features underlying human languages, but are they the product of a single system dedicated to language, or are they emergent from other capacities devised primarily for cognitive analysis and only secondarily used for communication? If universal features are emergent then we would not expect language to be a self-contained system designed for communication; instead, we can envisage it as more of a Heath-Robinson affair, with capacities designed for cognition being pressed into use for communication. The first grammatical utterances had to communicate A-Relationship-B models; but the act of generating them produced unexpected outcomes which, when communicated, created more outcomes; for instance, Protolanguage 2 led to awareness of self, which led to a dispassionate self, then awareness of selfness, and eventually to recursion.

In any signalling system, the nature of the messages to be communicated imposes constraints on the message forms. The A-Relationship-B model requires a three-component message; it requires recognition of objects and actions as two different classes; it requires tagging individuals with labels or names; and it requires negotiation toward meaning to ensure sender and receiver agree on meaning. Each grammatical innovation relies on cognitive processes evolved for other, non-communicative purposes, and which therefore have their own functions. A specialised language organ is not needed to explain the universals of language, they can be imposed on language just as easily by the non-communicative cognitive processes that language relies on. We would expect to see one-, two- and three-argument forms in every language, as well as object-action distinction, abstract reference, a way to share temporality, iteration, and Halliday's metafunctions. Indeed, at this level of comparison, language differences should seem quite trivial – as Chomsky (1957) originally proposed.

The origins of language grammar proposed here also suggests a solution to the problem of language complexity: did language start simple and become complex (Burling, 2005, ch9), or start complex and become simpler (Wray, 2002a)? The answer is both: the original set of communicable grammatical forms included only A-Relationship-B, with more complex grammatical structures developing out of this. Communication itself, however, would initially have been laborious, with each form being negotiated toward shared meaning; there were no pre-existing semantic givens shared by sender and receiver. As language spread through a community, and each new utterance joined the already-existing discourse, consensus would begin to appear through negotiation toward meaning and negotiation toward form. So, instead of a single change toward complexity or simplicity, utterance simplified while meaning and form complexified: as simple meanings and forms became easier to share, room was made for sharing more complex meanings and forms. Language started complex and remains complex, but in different ways.

The Socialisation-Cognition-Communication Braid

Is there a unified process behind language grammar? The EAORC Routes to Language map shows sixty-one capacities implicated in language origins, which together produce the complexities of modern human language (http://www.martinedwardes.me.uk/eaorc/eaorc_languageroute.html); but is there a single mechanism, a single explanation, behind the evolution of all these capacities? The simple answer is evolution itself: each capacity, when present, produces a fitter individual than when it is absent; and what makes individuals fitter is that they get more genes into the future; so individuals with a new and advantageous capacity will, over generations, become ubiquitous in the population. Another way of expressing this is that evolution has a positive bias toward fitness, it is not a random walk through the space of possibilities. However, this evolutionary explanation just restates the original question in a new form: is there a unified process behind the fitness of the capacities used in language grammar?

We have already seen in chapter 12 that capacities can be grouped into developmental pathways, and that three developmental pathways (modelling others, self-modelling, and abstract reference) form a ratcheting process which leads from Protolanguage 2 to Complex Language 1. These three pathways are in turn parts of longer pathways through the Routes to Language model, which can be described as socialisation, cognition and communication. Socialisation consists of the customs and culture the group imposes on the individual, and changes in socialisation inevitably affect cognition. Cognition covers how the individual deals with the societal impositions, and changes in cognition affect what is communicable between individuals in the group. Finally, communication is how the individual negotiates with other members of the group to mitigate the impositions society places on them; the outcomes of this negotiation then feed back into the group, affecting socialisation. To be an evolutionarily fit human you need to be socialised, clever and eloquent.

For instance, if I join a chess club then I must adopt the club's socialisation norms; I must learn the rules of chess, the social rules of the club, tournament chess rules, the pragmatic rules of being a chess person – and the taboos; I must absorb the semantics of chess (*e.g.*, addressing a new player), chess grammar (*e.g.*, in a tournament, say "check" before pressing the timer button), and chess phonology (*e.g.*, when to be silent). joining the club means learning new cognitive forms, which let me communicate in *chess club language* and thus assimilate. Assimilation introduces me to affiliative knowledge which I can use to signal my membership of the group; and which, in turn, helps me establish a social role in the club. This happens in every social group: the group imposes its rules on the joining individual; the individual adopts the rules; and, as an initiated member of the group, the individual becomes part of the group's social structure.

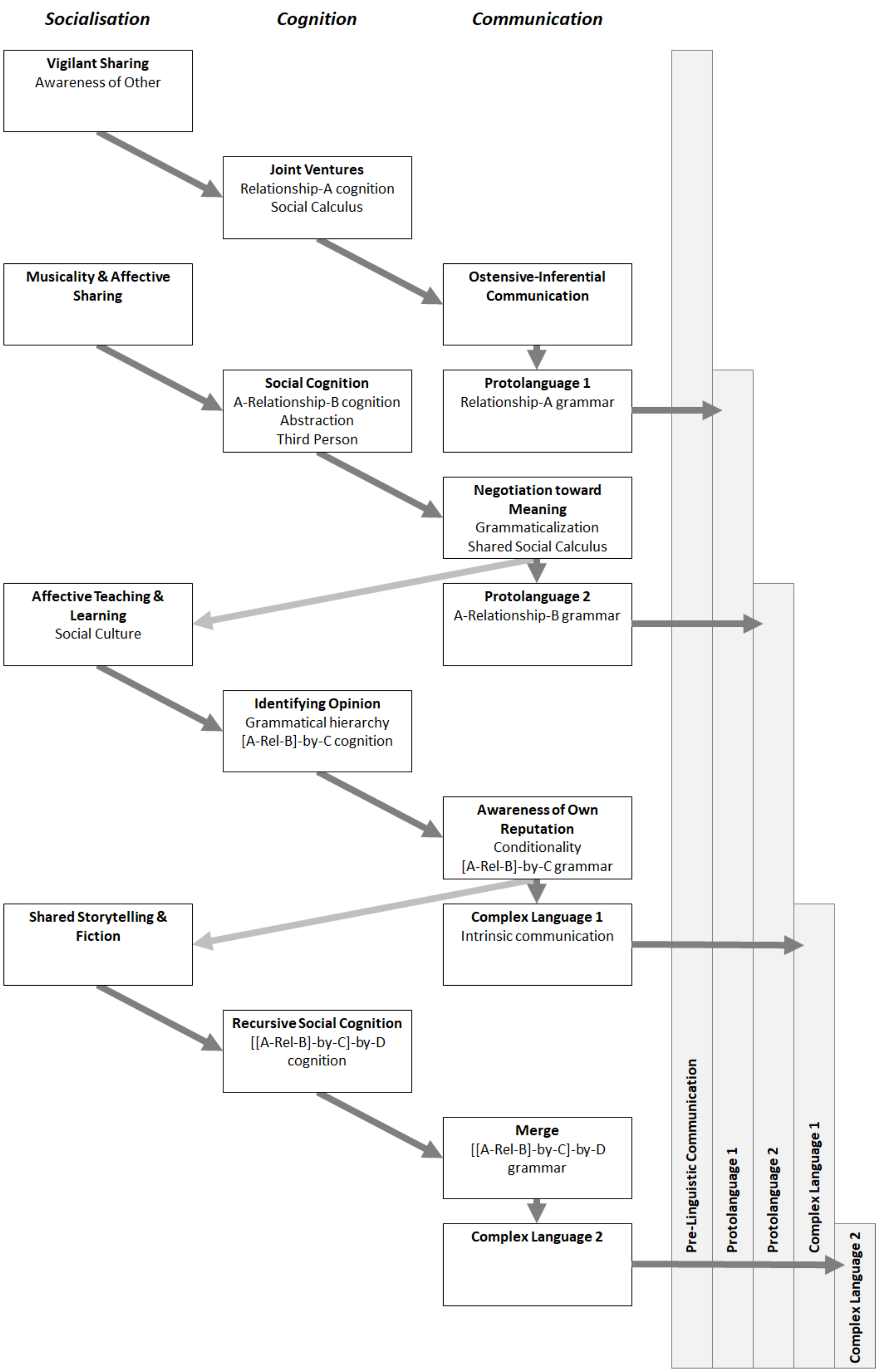

Figure 13.1: The Socialisation-Cognition-Communication Braid

Socialisation, Cognition and Communication form a braid linking prehumans to modern Homo sapiens. Social innovations drive cognitive complexity; which, when shared, require more complex utterance forms and more sophisticated ways of negotiating toward meaning; which then generate more social innovations; and so on. In terms of grammar (see figure 13.1), we can see the Vigilant Sharing capacity of Australopithecines leading to the capacity in early *Homo* to engage in Joint Ventures, which in turn leads to the capacity for Ostensive-Inferential Communication, generating Protolanguage 1. Similarly, we see the social capacity for Musicality and Affective Sharing (sharing emotional as well as environmental information) leading to Social Cognition and A-Relationship-B modelling, and on to the communicative capacity for Negotiation toward Meaning, which generates Protolanguage 2. Negotiation toward Meaning also leads to the Social Culture skill of Affective Teaching and Learning, which leads to the cognitive capacity for Identifying Opinion and [A-Relationship-B]-by-C modelling, which can be shared because the modelling creates an Awareness of Own Reputation. Awareness of Own Reputation generates Complex Language 1, and it also leads to the social capacity for Shared Storytelling and Fiction. This leads to Recursive Social Cognition, which leads to the communication of recursive forms, or Merge, and generates Complex Language 2. There is almost an inevitability in this process: in each case, adopting the next capacity gives the individual a fitness advantage over others – at least, until the others catch up and a new advantage emerges.

Each new communication method does not replace previous iterations; everything available in Protolanguages 1 and 2 and Complex Language 1 is available in Complex Language 2, or modern language. We still use one-word signals like *yes* and *no*, an inheritance from pre-linguistic communication; we still use one-argument forms from Protolanguage 1, with the imperative as an example; and we regularly use two-argument grammar from Protolanguage 2, and non-iterative utterances from Complex Language 1.

This leaves an intriguing question: is another iteration of human communication likely to emerge, or have we reached the ultimate communicative complexity? Since Complex Language 2 emerged, writing has been invented, which means speaker and listener need not be in the same place and time; so it enhances socialisation by disseminating information quicker and easier. Recently, as an outcome of writing, the computing revolution has enhanced our cognition by outsourcing some of our cerebral load to machines. So, is a new way of communicating coming … or is it already here?

Epilogue Language Grammar: Being Human

What I mind in modern society very much is the awful lack of grammar.
Ruth Rendell

This book has looked at the development of human language grammar out of the cognitive mechanisms for social modelling. Modern humans are a highly socialised species, so mapping our social interactions enhances our individual evolutionary fitness. Access to other people's social maps would also be valuable, and reciprocal sharing would make it valuable to both sender and receiver. Sharing social maps is not the whole story of grammar; but it opens the way for other social communication; and the rich social discourse it engenders generates a need for complex grammar. Modern language is still largely a mechanism for sharing social models: it is a product of sociality, a measure of sociality, and a tool for social exchange.

Grammar was initially communicatively complex, requiring negotiation toward meaning; but structural complexity replaced communicative complexity as the information to be signalled became more elaborate. Where Ruth Rendell sees an "awful lack of grammar" in modern usage, I see an increasing tangle of linguistic formality surrounding the information we want to share, much of it created by a need to communicate cultural convention rather than factual information.

The socialisation of language does not mean differences in grammar are wholly produced by differences in culture; social functionality in languages can be essentially universal – as we see is the case for norms of morality. The cultural differences between modern societies are often viewed as not just different ways of being human but as ways of being alien; but, as Tomasello (2018) shows, there are common cultural norms in all successful human societies. He suggests that modern human morality evolved out of pre-human self-interest in two stages: joint intentionality and collective intentionality. Joint intentionality involved collaboration in food procurement and consumption, creating the understanding that "we is greater than me"; while collective intentionality involved division of labour and task specialisation, creating the understanding that *I* must be part of *we*. We are all tied to our group by the need to exchange our personal surpluses for the different surpluses of others.

Alacorta & Sosis (2005) see universals even in religion, traditionally a great definer of *us* versus *them*. They identify four universal traits in belief structures: first is the magical element of counterintuition, seeing the unseen, giving existence to the non-existent, and, as Carroll (1872 [1982], ch5: Wool and Water) put it, "believing six impossible things before breakfast"; second is mystery, rituals shared with others initiated into the group but which seem arcane and unnecessary to outsiders; third is the binary division of reality into the sacred and the profane, with exaggerated offense being taken for affronts to the rules of the sacred; and fourth is a preference for initiation and induction rituals during adolescence. All major religions have these four traits, although they view their own synthesis of the four as coherent, sensible and necessary, while those of others are bewildering, irrational and haphazard. Indeed, this arbitrary separation of in-group and out-group could almost be a fifth universal trait in belief structures. Ganapini (2023) sees this willingness to acquiesce in the agreed truth of outlandish stories as conformity to group norms – six impossible things before breakfast, indeed.

Grammar, like morality and religion, may seem an arbitrary symbolic response to a need for a social mechanism, but there are regularities behind grammar systems. The symbolic response itself can vary quite widely, but the social need for and purpose of the symbolic response are universal. In the case of grammar, the social need is for communicative mechanisms which can coordinate cooperation; and the symbolic responses are tools for negotiating toward meaning, alongside shared markers of in-group membership which can be used to identify out-group individuals. The tools include rules dictated by necessary natural conditions, useful rules which make utterances easier to produce, and arbitrary rules because we can; and the in-group markers are expressed in language as idioms, dialects and individual languages, made possible by the wide range of phonological, prosodic, semantic and grammatical cues that can be built into utterances.

We should not treat grammar as just the rules dictated by necessary natural conditions, any more than we should treat morality as just joint intentionality, or religion as just mysticism.

Differences and Similarities

How distant is human language from other natural systems of communication? In chapter 1, two questions were posed: What is special about being human? And what is special about the specialities that make us human? The first question allowed us to look at language grammar in terms of the differences between modern humans and other species, while the second question asked whether those differences matter – and whether, despite the differences, we are neither special nor unusual in nature.

Two physical differences between us and our closest extant relatives, the two *Pan* species, chimpanzees and bonobos, were discussed: bipedalism and brain construction. Bipedalism is important because it freed our forelimbs: where the *Pan* species have four legs, two of which serve in a second role for brachiation and in a third role as hands, adult humans have two legs and two hands, with each set of limbs being dedicated to its specific role: hind limbs for mobility, forelimbs for manipulation. With bipedalism, humans became more gracile, largely hairless, altricial (infants have a greater dependence on adult support), and longer-lived; and this longer life gives females a long fertility-free period after menopause to support both their later children and their grandchildren. Above all, however, bipedalism made possible intricate gestural control, a signature feature of the *Homo* clade, which eventually produced modern human civilisation – itself a masterclass in niche construction (and destruction). Yet bipedalism did not produce civilisation by itself: gestural control and niche construction both rely on increases in brainpower.

There are several significant differences between the *Pan* and *Homo sapiens* brain, including a three-fold difference in volume, increased connectivity between parts of the brain, and increased complexity in the prefrontal cortex, associated with planning, imagination, selfhood, awareness of other, working memory, space-time cognition, and social modelling. However, while these physical differences reflect the cognitive differences between *Pan* and *Homo sapiens*, they are both causes and products of cognitive change. A ratchet effect makes it difficult to definitively identify which happened first: cognitive changes led to physical brain changes led to cognitive changes, and so on. However, we know that the physical brain is a costly organ, so increasing that cost without a pre-existing purpose does not seem an evolutionarily fit strategy; evolution should produce brains only as costly as needed.

Fortunately, two key skills leave traces in the archaeological record of the ratchet at work. First is toolmaking, a peculiarly human skill (although not exclusively so): the history of human toolmaking traditions, from the Lomekwian to the Levallois-Mousterian, corresponds closely with the increasing cognitive complexity of the *Homo* clade. The fossil record since 3mya is rich enough to allow us to recreate the cognitive skills behind the stone tool manufacturing processes, and we can compare this to the less rich evidence of brain size increases in the *Homo* clade to infer brain complexity from tool complexity.

The second skill, hunting, has left less fossil evidence; but we can compare the hunting methods of modern *Pan* species, particularly chimpanzees, with those of modern human gatherer-hunter tribes. A chimpanzee hunt is usually opportunistic, initiated by a small number of individuals, with others joining in when they notice what is happening. In contrast, human hunts are often carefully planned cooperative ventures in pursuit of much larger prey. Chimpanzee hunts usually last only a few minutes, while humans carry out pursuit hunts lasting hours or even days. The spoils of a chimpanzee hunt usually go to the one who kills the prey, with others begging or bullying for a share; while human hunts often end symbolically, not when the prey is killed but when the kill is returned to the campsite for cooking. As cooking seems to have been largely under the supervision of the females, distribution of the cooked meat was also under their supervision; and, with the female commitment to offspring and allies, they were able to ensure egalitarian sharing.

Other cognitive differences between *Pan* species and *Homo sapiens* which have a role in the evolution of language grammar include:

Working memory: *Pan* species can be aware of between one and three things simultaneously; modern humans can manage between five and nine things simultaneously.

Mental modelling: both *Pan* individuals and modern humans can maintain a map of their emotional relationships with other individuals in their group. Modern humans can also maintain, manipulate and share a map of the relationships between other individuals; while Pan individuals may be able to maintain and manipulate a map of relationships between others, there is no indication that they share it.

Selfness: a modern human can objectively recognise themself and model themself dispassionately as a third-person node in their map of the relationships between others.

Curiosity: Humans retain a capacity for play into adulthood, and therefore a capacity for experimentation: they retain a lifelong curiosity about the world around them. Adult *Pan* species retain an interest in *what*, *where* and *who* questions, but have less interest in solving *when*, *how* and *why* puzzles.

Symbolic representation: an arbitrary sound can stand in place of an object or event (a vocable); an arbitrary gesture can stand in place of a vocable (a word); one arbitrary gesture can stand in place of another arbitrary gesture (a metaphor); and a network of metaphors can stand in place of a network of facts (a story). Many primates use vocables; some animals exposed to human language use words, and some even use metaphors; but only humans tell stories.

Temporality: Humans can share information which is relevant in the present but which is about events in the past or the future – telling-about. While other species may have awareness of past and future events, they do not appear to share that awareness.

Connectivity: Humans can link information together to create sharable ordered narratives; we have no evidence that other species do this.

Recursion: Humans can share information as a set of hierarchical facts, with one fact dependent on another. There is no indication of other species doing this.

There are also several key differences in the socialisation of modern humans and modern *Pan* species. Humans can organise themselves into much larger groups than chimpanzees and bonobos because human groups are richly hierarchical, with many levels of groups within groups; currently, the largest grouping is the culturally homogenous nation state, which seems to be effectively unlimited in size; but loyalty to this size of group is also more abstract and provisional than to smaller groups within the nation state. In contrast, chimpanzee and bonobo societies consist of groups of between 20 and 150 individuals, which can be broken down into alliance groups of adult males, mixed adults of both sexes, adult females with their offspring, or a single female with her offspring (Wrangham, 1975). Group organisation is more restrictive for chimpanzees than for humans, which means that the range of possible relationships is also restricted. Non-human apes tend to be patrilocal and, with the exception of bonobos, patriarchal; modern human groups can be patrilocal, matrilocal or nonlocal, and patriarchal, matriarchal or egalitarian.

However, the key socialisation differences between humans and other apes involve interpersonal relationships. Like bonobos, we treat sex and sexuality as primarily recreational rather than primarily reproductive, and we also have a high level of empathy with other members of our group. Where nonhuman apes can feel sympathy for others and offer consolation, humans can intuitively feel the pain those others are feeling. We have more than a theory of mind to give us an understanding of what others are thinking, we can engage with their beliefs and desires; and this empathy makes interactions between humans more benevolently altruistic than those between other primates. While it is true that some humans are unable to feel this empathy, displaying what, to others, appears to be gratuitous cruelty (*e.g.*, Giroux, 2017; Wehner, 2020), deliberately cruel (or psychopathic) individuals are in a small minority – although heavily over-represented in our leader cadres, even in democracies.

Despite all these differences making us the outlier species in the primate clade, we are nonetheless primates; and many of our *speci-al* capacities can be linked to analogous capacities in other primates. Even the attempts to teach human language to nonhumans have shown that, while full language seems to be impossible without a human brain, language-like communication is possible for several nonhuman species, including some outside the primate family. There may be no all-or-nothing difference between humans and other primates, only differences of degree. Language grammar does not need to be an exclusively human cognitive system which made complex communication possible; instead, it could have been exapted from complex social cognition for sharing social information.

And Finally ...

While I have tried to make this book comprehensible and comprehensive, it still provides an incomplete explanation for the origins of language grammar. For instance, how sharing social models become general sharing of subject-verb-object forms, regardless of social significance (such as *the couch had a blue cover*), is not addressed. Somehow a specific communication system became general, expanding labelling to cover any labelling need; and the Relationship link between labels has expanded to include any activity or relationship involving labelled objects. From our current position in language complexification we can say that generalisation must have occurred, and we can posit several ways in which it could have occurred. The limited purpose of this book allows me to draw a discrete veil over this part of the history of language development; but it is not because there is nothing to see, it only means I have told as much of the story as I want to tell, and the rest is a tale for another time – and probably another person.

Honesty has also been somewhat cursorily addressed, treated as a necessity for language without properly investigating why; it is more excuse than explanation for the lack of factuality in most modern language. The honesty problem needs to be analysed more closely, along with the question of why truth can be non-factual for humans; but the issues involved are too elaborate to be effectively tackled here. Several other aspects of language grammar need addressing in greater detail than given here; but they are also tasks for other days and other people.

The puzzle of how modern language grammar came to be is not for the faint-hearted: we may never have a full explanation for the sources of language grammar. The Linguistic Society of Paris banned discussion of language origins in 1866 for good reason (Aitchison, 1996, 5): evidence was sparse, wild speculation rampant. Yet this *what-iffery* is a vital part of the scientific process – and an important part of being human. This book has speculated on the sources of language grammar, but hopefully in a scientific way: proposing a case and testing it against evidence.

While this book does not definitively answer the question, "what caused language grammar", it hopefully tells a consistent and effective story, explaining more than it obscures. Regardless of the hypothesis argued here the search for the origins of language and grammar will continue to be pursued energetically. Understanding our species' reliance on grammatical language is crucial to understanding our humanity.

Bibliography

Ackermann, H., Hage, S.R. & Ziegler, W. 2014. 'Brain mechanisms of acoustic communication in humans and nonhuman primates: An evolutionary perspective'. *Behavioral and Brain Sciences* 37, 529-604.

Addessi, E., Crescimbene L. & Visalberghi, E. 2007. 'Do capuchin monkeys (Cebus apella) use tokens as symbols?' *Proceedings of the Royal Society B* 274, 2579-2585.

Adler, D.S., Wilkinson, K.N., Blockley, S., Mark, D.F., Pinhasi, R., Schmidt-Magee, B.A., Nahapetyan, S., Mallol, C., Berna, F., Glauberman, P.J., Raczynski-Henk, Y., Wales, N., Frahm, E., Jöris, O., MacLeod, A., Smith, V.C., Cullen, V.L. & Gasparian, B. 2014. 'Early Levallois technology and the Lower to Middle Paleolithic transition in the Southern Caucasus'. *Science* 345:6204, 1609-1613.

Aiello, L.C. & Dean, C. 1990. *An Introduction to Human Evolutionary Anatomy*. Academic Press Ltd.

Aiello, L.C. & Dunbar, R.I.M. 1993. 'Neocortex Size, Group Size, and the Evolution of Language'. *Current Anthropology* 34:2, 184-193.

Aiello, L.C. & Wheeler, P. 1995. 'The Expensive Tissue Hypothesis: the brain and the digestive system in human and primate evolution'. *Current Anthropology* 36:2, 199-221.

Aitchison, J. 1998. *The Articulate Mammal: An Introduction to Psycholinguistics (4th edition)*. Routledge.

Aitchison, J. 1996. *The Seeds of Speech: Language Origin and Evolution*. Cambridge University Press.

Alcorta, C.S. & Sosis, R. 2005. 'Ritual, Emotion, and Sacred Symbols: The Evolution of Religion as an Adaptive Complex'. *Human Nature*, 16:4, 323-359.

Allen, C. & Saidel, E. 1998. 'The evolution of reference'. In D.D. Cummins & C. Allen (eds.), *The Evolution of Mind*. Oxford University Press.

Allen, J. & Seidenberg, M.S. 1999. 'The Emergence of Grammaticality in Connectionist networks'. In B. MacWhinney (ed.), *The Emergence of Language*. Lawrence Erlbaum Associates.

Almécija, S., Smaers, J.B. & Jungers, W.L. 2015. 'The evolution of human and ape hand proportions'. *Nature Communications* 6, 7717.

Altschul, D.M., Hopkins, W.D., Herrelko, E.S., Inoue-Murayama, M., Matsuzawa, T., King, J.E., Ross, S.R. & Weiss, A. 2018. 'Personality links with lifespan in chimpanzees'. *eLife* 7, e33781.

Alvarez, G., Quinteiro, C. & Ceballos, F.C. (2011). 'Inbreeding and Genetic Disorder'. In K. Ikehara (ed.), *Advances in the Study of Genetic Disorders*, 21-44. IntechOpen.

Ambrose, S.H. 2001. ‘Paleolithic Technology and Human Evolution’. *Science* 291, 1748-1753.

Anderson, A., Chilczuk, S., Nelson, K., Ruther, R. & Wall-Scheffler, C. 2023. ‘The Myth of Man the Hunter: Women’s contribution to the hunt across ethnographic contexts’. *PLoS ONE* 18:6, e0287101.

Andrade, M.C.B. 2003. ‘Risky mate search and male self-sacrifice in redback spiders’. *Behavioral Ecology* 14:4, 531-538.

Anon, Various. [2000]. *Holy Bible: New International Version*. Hodder Stoughton.

Anthony, D.W. 2007. ‘The Promise and Politics of the Mother Tongue’. In *The Horse, the Wheel, and Language*, 3-20. Princeton University Press.

Arbib, M.A. 2021. ‘The aboutness of language and the evolution of the construction-ready brain’ In A. Lock, C. Sinha & N. Gonthier (eds.), *The Oxford Handbook of Human Symbolic Evolution*. Oxford University Press.

Arbib, M.A. 2008. ‘From grasp to language: Embodied concepts and the challenge of abstraction’. *Journal of Physiology – Paris* 102, 4-20.

Arbib, M. A. 2005. ‘From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics’. *Behavioral and Brain Sciences* 28, 105-167.

Arbib, M.A., Liebal, K. & Pika, S. 2008. ‘Primate Vocalization, Gesture, and the Evolution of Human Language’. *Current Anthropology*, 49:6, 1053-1076.

Areces, C. & Blackburn, P. 2005. ‘Reichenbach, Prior and Montague: a semantic get-together’. In *Articles in Honor of Dov Gabbay’s 60th Birthday*. Kings College Press.

Arias-Carrión, O., Stamelou, M., Murillo-Rodríguez, E., Menéndez-González, M. & Pöppel, E. 2010. ‘Dopaminergic reward system: a short integrative review’. *International Archives of Medicine* 3:24.

Aristotle. 350 BCE [2004]. *Nichomachean Ethics*. Penguin.

Asano, R. & Boeckx, C. 2015. ‘Syntax in language and music: what is the right level of comparison?’ *Frontiers in Psychology* 6, 942.

Atkinson, M., Mills, G.J. & Smith, K. 2019. ‘Social Group Effects on the Emergence of Communicative Conventions and Language Complexity’. *Journal of Language Evolution* 4:1, 1-18.

Atran, S. 2005. ‘Adaptationism for Human Cognition: Strong, Spurious or Weak?’ *Mind & Language* 20:1, 39-67.

Austin, J.L. 1962. *How to Do Things with Words*. Oxford University Press.

Baan, C., Bergmüller, R., Smith, D.W. & Molnar, B. 2014. ‘Conflict management in free-ranging wolves, Canis lupus’. *Animal Behaviour* 90, 327-334.

Baker, M.C. 2001. *The Atoms of Language: The Mind’s Hidden Rules of Grammar*. Basicbooks.

Baker, R.R. & Bellis, M.A. 1995. *Human sperm competition: Copulation, Masturbation, and Infidelity*. Chapman & Hall.

Balme, J. & Bowdler, S. 2006. 'Spear and Digging Stick: The origin of gender and its implications for the colonisation of new continents'. *Journal of Social Archaeology* 6: 379-401.

Balter, M. 2009. 'On the Origin of Art and Symbolism'. *Science* 323, 709-711.

Barclay, P. & Willer, R. 2007. 'Partner choice creates competitive altruism in humans'. *Proceedings of the Royal Society B* 274, 749-753.

Barker, J.L., Power, E.A., Heap, S., Puurtinen M. & Sosis, R. 2019. 'Content, cost, and context: A framework for understanding human signaling systems'. *Evolutionary Anthropology*, 28:2, 86-99.

Barnes, D. M. 2005. 'Possible Tool Use by Beavers, Castor canadensis, in a Northern Ontario Watershed'. *The Canadian Field-Naturalist* 119:3, 441-443.

Baron-Cohen, S. 1995. *Mindblindness: An Essay on Autism and Theory of Mind*. MIT Press.

Barone, P., Wenzel, L., Proft, M. & Rakoczy, H. 2022. 'Do young children track other's beliefs, or merely their perceptual access? An interactive, anticipatory measure of early theory of mind'. *Royal Society Open Science* 9, 211278.

Barton, D. & Papen, U. (2010). 'What is the Anthropology of Writing?' In D. Barton & U. Papen (eds.), *The Anthropology of Writing: Understanding Textually Mediated Worlds*, 3-32. Bloomsbury.

Bassett, D.S. & Gazzaniga, M.S. 2011. 'Understanding complexity in the human brain'. *Trends in Cognitive Science* 15:5, 200-209.

Bates, E., Dale, P.S. & Thal, D. 1995. 'Individual Differences and their Implications for Theories of Language Development'. In P. Fletcher & B. MacWhinney (eds.), *The Handbook of Child Language*. Blackwell Publishers.

Bateson, G. 1985. 'A Theory of Play and Fantasy'. In R.E. Innis (ed.), *Semiotics: An Introductory Anthology*. Indiana University Press.

Battu, B. & Rahwan, T. 2023. 'Cooperation without punishment'. *Nature Scientific Reports* 13, 1213.

Bauer, E.B. & Smuts, B.B. 2007.' Cooperation and competition during dyadic play in domestic dogs, Canis familiaris'. *Animal Behaviour* 73, 489-499.

Baumeister, R.F., Ainsworth, S.E. & Vohs, K.D. 2016. 'Are groups more or less than the sum of their members? The moderating role of individual identification'. *Behavioral and Brain Sciences* 39, e137.

Baumeister, R.F., Campbell, J.D., Krueger, J.I. & Vohs, K.D. 2005. 'Exploding the Self-esteem Myth'. *Scientific American Mind* 16-4, 50-57.

Beaulieu-Laroche, L., Brown, N.J., Hansen, M., Toloza, E.H.S., Sharma, J., Williams, Z.M., Frosch, M.P., Cosgrove, G.R., Cash, S.S. & Harnett, M.T. 2021. 'Allometric rules for mammalian cortical layer 5 neuron biophysics'. *Nature* 600, 274-278.

Beckerman, S. & Valentine, P. 2002. 'Introduction: The concept of partible paternity among Native South Americans'. In S. Beckerman & P. Valentine (eds.), *Cultures of Multiple Fathers: The Theory and Practice of Partible Paternity in Lowland South America*. University Press of Florida.

Beecher, M.D. 2021. 'Why Are No Animal Communication Systems Simple Languages?' *Frontiers in Psychology* 12, 602635.

Behrens, H. 2021. 'Constructivist Approaches to First Language Acquisition'. *Journal of Child Language* 48, 959-983.

Belin, P. 2006. 'Voice processing in human and non-human primates'. *Philosophical Transactions of the Royal Society B* 361, 2091-2107.

Belyk, M., Brown, R., Beal, D.S., Roebroeck, A., McGettigan, C., Guldner, S. & Kotz, S.A. 2021. 'Human larynx motor cortices coordinate respiration for vocal-motor control'. *NeuroImage* 239, 118326.

Benítez-Burraco, A. & Longa, V.M. 2010. 'Evo-Devo – Of Course, But Which One? Some Comments on Chomsky's Analogies between the Biolinguistic Approach and Evo-Devo'. *Biolinguistics* 4:4, 308-323.

Benson, J.D. & Greaves, W.S. 2005. *Functional Dimensions of Ape-Human Discourse*. Equinox Publishing Ltd.

Benson, M.S. 1993. 'The structure of four- and five-year-olds' narratives in pretend play and storytelling'. *First Language* 13, 203-223.

Benveniste, É. 1970. 'The Nature of Pronouns'. In P. Cobley (ed.), *The Communication Theory Reader*. Routledge.

Berg, L. 1972. *Look at Kids*. Penguin Books.

Berger, L.R., Hawks, J., de Ruiter, D.J., Churchill, S.E., Schmid, P., Delezene, L.K., Kivell, T.L., Garvin, H.M., Williams, S.A., DeSilva, J.M., Skinner, M.M., Musiba, C.M., Cameron, N., Holliday, T.W., Harcourt-Smith, W., Ackermann, R.R., Bastir, M., Bogin, B., Bolter, D., Brophy, J., Cofran, Z.D., Congdon, K.A., Deane, A.S., Dembo, M., Drapeau, M., Elliott, M.C., Feuerriegel, E.M., Garcia-Martinez, D., Green, D.J., Gurtov, A., Irish, J.D., Kruger, A., Laird, M.F., Marchi, D., Meyer, M.R., Nalla, S., Negash, E.W., Orr, C.M., Radovcic, D., Schroeder, L., Scott, J.E., Throckmorton, Z., Tocheri, M.W., VanSickle, C., Walker, C.S., Wei, P. & Zipfel, B. 2015. 'Homo naledi, a new species of the genus Homo from the Dinaledi Chamber, South Africa'. *eLife* 4, e09560.

Bergstrom, C.T. & Lachmann, M. 1997. 'Signalling among relatives I. Is costly signalling too costly?' *Philosophical Transactions of the Royal Society of London B*, 352, 609-617.

Bergstrom, T.C. 1994. 'On the Economics of Polygyny'. *eScholarship Repository, University of California*. (http://repositories.cdlib.org/ucsbecon/bergstrom/1994A)

Berkeley, G. 1710 [2003]. *A Treatise Concerning the Principles of Human Knowledge*. Project Gutenberg, https://www.gutenberg.org/ebooks/4723.

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

Berwick, R.C. & Chomsky, N. 2016. *Why Only Us: Language and Evolution*. MIT Press.

Bessa, J., Hockings, K. & Biro, D. 2021. 'First Evidence of Chimpanzee Extractive Tool Use in Cantanhez, Guinea-Bissau: Cross-Community Variation in Honey Dipping'. *Frontiers in Ecology & Evolution* 9, 625303.

Bican, A. & Rastall, P.R. 2013. *Axiomatic Functionalism: Theory and Application*. Peter Lang AG.

Bickerton, D. 2014. 'Some Problems for Biolinguistics'. *Biolinguistics* 8, 73-96.

Bickerton, D. 2002. 'Foraging Versus Social Intelligence in the Evolution of Protolanguage'. In A. Wray (ed.), *The Transition to Language*. Oxford University Press.

Bickerton, D. 2000. 'How Protolanguage Became Language'. In C. Knight, M. Studdert-Kennedy & J.R. Hurford (eds.), *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form*. Cambridge University Press.

Bickerton, D. 1998. 'Catastrophic Evolution: the case for a single step from protolanguage to full human language'. In J.R. Hurford, M. Studdert-Kennedy & C. Knight (eds.), *Approaches to the Evolution of Language: Social and Cognitive Bases*. Cambridge University Press.

Bickerton, D. 1990. *Language and Species*. University of Chicago Press.

Biduła, S.P., Przybylski, L., Pawlak, M.A. & Króliczak, G. 2017. 'Unique Neural Characteristics of Atypical Lateralization of Language in Healthy Individuals'. *Frontiers in Neuroscience* 11, 525.

Binet, A. 1903. *L'Etude Experimentale de l'Intelligence (Experimental Studies of Intelligence)*. Schleicher Frères & Cie.

Bishop, D.V.M. & Snowling, M.J. 2004. 'Developmental Dyslexia and Specific Language Impairment: Same or Different?' *Psychological Bulletin* 130:6, 858-886.

Block, D. 2008. 'On the appropriateness of the Metaphor of LOSS'. In R. Rubdy & P. Tan (eds.), *Language as Commodity: Global Structures, Local Marketplaces*. Continuum.

Bloom, P. 2002. *How Children Learn the Meanings of Words*. MIT Press.

Bloomfield, L. 1933. *Language*. Holt, Rinehart & Winston.

Bloor, T. & Bloor, M. 1995. *The Functional Analysis of English: A Hallidayan Approach*. Arnold.

Blyth, C. 2009. *The Art of Conversation: A Guided Tour of a Neglected Pleasure*. Gotham Books.

Boas, F. 1938. 'Introduction'. In F. Boas(ed.), *General Anthropology*. D.C. Heath and Company.

Bod, R. 1998. *Beyond Grammar: An Experience-Based Theory of Language*. CSLI Publications.

Boehm, C. 1999. *Hierarchy in the Forest: The Evolution of Egalitarian Behaviour*. Harvard University Press.

Boesch-Achermann, H. & Boesch, C. 1993. 'Tool Use in Wild Chimpanzees: New Light from Dark Forests'. *Current Directions in Psychological Science*, 2-1, 18-21.

Bogin, B. 2015. 'Human Growth and Development'. In M.P. Meuhlenbein (ed.), *Basics in Human Evolution*. Academic Press, 285-293.

Bohnemeyer, J. 2009. 'Temporal anaphora in a tenseless language'. In W. Klein & P. Li (eds.), *The Expression of Time*, 83-127. De Gruyter Mouton.

Bollmann, S. & Barth, M. 2021. 'New acquisition techniques and their prospects for the achievable resolution of fMRI☆'. *Progress in Neurobiology* 207, 101936.

Bosch, L., Figueras, M., Teixidó M. & Ramon-Casas, M. 2013. 'Rapid gains in segmenting fluent speech when words match the rhythmic unit: evidence from infants acquiring syllable-timed languages'. *Frontiers in Psychology* 4:106, 91-102.

Bourke, A.F.G. & Franks, N.R. 1995. *Social Evolution in Ants*. Princeton University Press.

Boyd, R., Gintis, H., Bowles, S. & Richerson, P.J. 2003. 'The evolution of altruistic punishment'. *Proceedings of the National Academy of Science*, 100-6, 3531-3535.

Boyd, R. & Richerson, P.J. 1996. 'Why culture is common, but cultural evolution is rare'. *Proceedings of the British Academy* 88, 77-93.

Bozic, M., Tyler, L.K., Ives, D.T., Randall, B. & Marslen-Wilson, W.D. 2010. 'Bihemispheric foundations for human speech comprehension'. *PNAS* 107:40, 17439-17444.

Bradbury, J.W. & Vehrencamp, S.L. 1998. *Principles of Animal Communication*. Sinauer Associates inc.

Bradshaw, I.G.A. 2004. 'Not by Bread Alone: Symbolic Loss, Trauma, and Recovery in Elephant Communities'. Society & Animals 12:2, 143-158.

Brewer, S. 2001. *A Child's World: A Unique Insight into How Children Think*. Headline Book Publishing.

Brewer, S.M. & McGrew, W.C. 1990. 'Chimpanzee use of a tool set to get honey'. *Folia Primatologica* 54, 100-104.

Brighton, H., Kirby, S. & Smith, K. 2005. 'Cultural Selection for Learnability: three principles underlying the view that language adapts to be learnable'. In M. Tallerman (ed.), *Language Origins: Perspectives on Evolution*. Oxford University Press.

Broom, D.M., Senaa, H. & Moynihan, K.L. 2009. 'Pigs learn what a mirror image represents and use it to obtain information'. *Animal Behaviour*, 78-5, 1037-1041.

Brown, P., Sutikna, T., Morwood, M.J., Soejono, R.P., Jatmiko, Saptomo, E.W. & Due, R.A. 2004. 'A new small-bodied hominin from the Late Pleistocene of Flores, Indonesia'. *Nature* 431, 1055-1061.

Brown, R. 1970. ‘Three Processes in the Child’s Acquisition of Syntax’. *Psycholinguistics: Selected Papers by Roger Brown*. Macmillan Press.

Brozzoli, C., Roy, A.C., Lidborg, L.H., & Lövdén, M. 2019. ‘Language as a Tool: Motor Proficiency Using a Tool Predicts Individual Linguistic Abilities’. *Frontiers in Psychology* 10, 1639.

Bruner, J. 1986. *Actual Minds, Possible Worlds*. Harvard University Press.

Budiansky, S. 1998. *If a Lion Could Talk: How Animals Think*. Weidenfeld & Nicolson.

Burdukiewicz, J.M. 2014. ‘The origin of symbolic behavior of Middle Palaeolithic humans: Recent controversies’. *Quaternary International* 326-327, 398-405.

Burland, T.M., Bennett, N.C., Jarvis, J.U.M., & Faulkes, C.G. 2002. ‘Eusociality in African mole-rats: new insights from patterns of genetic relatedness in the Damaraland mole-rat (Cryptomys damarensis)’. *Proceedings of the Royal Society of London B,* 269:1495, 1025-1030.

Burling, R. 2005. *The Talking Ape: How Language Evolved*. Oxford University Press.

Burrows, A.M., Waller, B.M., Parr, L.A. & Bonar, C.J. 2006.’Muscles of facial expression in the chimpanzee (Pan troglodytes): descriptive, comparative and phylogenetic contexts’. *Journal of Anatomy*, 208:2, 153-167.

Buzing, P.C., Eiben, A.E. & Schut, M.C. 2005. ‘Emerging Communication and Cooperation in Evolving Agent Societies’. *Journal of Artificial Societies and Social Simulation*, 8-1.

Byrne, R.W. 2004. ‘The manual skills and cognition that lie behind hominid tool use’. In A.E. Russon & D.R. Begun (eds.), *The Evolution of Thought: Evolutionary Origins of Great Ape Intelligence*. Cambridge University Press.

Byrne, R.W. 2000. ‘Evolution of Primate Cognition’. *Cognitive Science* 24-3, 543-570.

Byrne, R.W. 1995. *The Thinking Ape: Evolutionary Origins of Intelligence*. Oxford University Press.

Caillois, R. 1958 [1961]. *Man, Play and Games*. University of Illinois Press.

Call, J. & Tomasello, M. 2008. ‘Does the chimpanzee have a theory of mind? 30 years later’. *Trends in Cognitive Sciences* 12:5, 187-192.

Calvin, W.H. & Bickerton, D. 2000. *Deus ex Machina*. MIT Press.

Cameron, D. 1998a. ‘Performing Gender Identity: young men’s talk and the construction of heterosexual masculinity’. In J. Coates (ed.), *Language and Gender: A Reader*. Blackwell Publishers Ltd.

Cameron, D. 1998b. ‘What Has Gender Got to Do with Sex?’ In R. Harris & G. Wolf (eds.), *Integrational Linguistics: A First Reader*. Elsevier Science Ltd.

Campbell, A. 1999. ‘Staying alive: Evolution, culture, and women’s intrasexual aggression’. *Behavioral and Brain Sciences* 22-02, 203-252.

Camus, A. 1951. *The Rebel*. Librairie Gallimard.

Carbonell, E., Barsky, D., Sala, R. & Celiberti, V. 2016. ‘Structural continuity and technological change in Lower Pleistocene toolkits’. *Quaternary International* 393, 6-18.

Carroll, L. 1872 [1982]. *Through the Looking Glass and What Alice Found There*. In *The Complete Illustrated Works of Lewis Carroll*. Chancellor Press.

Carruthers, P. 2009. ‘How we know our own minds: The relationship between mindreading and metacognition’. *Behavioral and Brain Sciences* 32:2, 121-138.

Carstairs-McCarthy, A. 1999. *The Origins of Complex Language: An Inquiry into the Evolutionary Beginnings of Sentences, Syllables and Truth*. Oxford University Press.

Carter, R. 1998. *Mapping the Mind*. Weidenfeld & Nicolson.

Cartmill, E.A., Beilock, S. & Goldin-Meadow, S. 2012. ‘A word in the hand: action, gesture and mental representation in humans and non-human primates’. *Philosophical Transactions of the Royal Society B* 367, 129-143.

Cäsar, C., Zuberbühler, K., Young, R.J. & Byrne, R.W. 2013. ‘Titi monkey call sequences vary with predator location and type’. *Biology Letters* 9, 20130535.

Cebeiro, A. & Key, A. 2023. ‘Captive bonobos (Pan paniscus) apply precision grips when using flaked stone tools’. *American Journal of Biological Anthropology* 2023: 1-14.

Chafe, W. 1970. *Meaning and the Structure of Language*. University of Chicago Press.

Charpentier, M.J.E., van Horn, R.C., Altmann, J. & Alberts, S.C. 2008. ‘Paternal effects on offspring fitness in a multimale primate society’. *Proceedings of the National Academy of Science* 105:6, 1988-1992.

Cheney, D.L. 2011. Extent and limits of cooperation in animals. *Proceedings of the National Academy of Science* 108:S2, 10902-10909.

Cheney, D.L. & Seyfarth, R.M. 1990. *How Monkeys See the World: Inside the Mind of Another Species*. University of Chicago Press.

Chiappe, D., Brown, A., Dow, B., Koontz, J., Rodriguez, M. & McCulloch, K. 2004. ‘Cheaters Are Looked At Longer and Remembered Better Than Cooperators in Social Exchange Situations’. *Evolutionary Psychology* 2, 108-120.

Chiat, S. 2000. *Understanding Children with Language Problems*. Cambridge University Press.

Chomsky, N. 2007. ‘Of Minds and Language’. *Biolinguistics* 1, 9-27.

Chomsky, N. 2006. *Language and Mind (3rd edition)*. Cambridge University Press.

Chomsky, N. 2005. ‘Three Factors in Language Design’. *Linguistic Inquiry* 36:1, 1-22.

Chomsky, N. 2002. *On Nature and Language*. Cambridge University Press.

Chomsky, N. 2000. *New Horizons in the Study of Language and Mind*. Cambridge University Press.

Chomsky, N. 1995. *The Minimalist Program*. MIT Press.

Chomsky, N. 1988. *Language and Problems of Knowledge: The Managua lectures*. MIT Press.

Chomsky, N. 1986. *Knowledge of Language – Its Nature, Origin, and Use*. Preager Publishers.

Chomsky, N. 1982. *Some Concepts and Consequences of the Theory of Government and Binding*. MIT Press.

Chomsky, N. 1980. *Rules and Representations*. Columbia University Press.

Chomsky, N. 1979 [1998]. *Language and Responsibility*. In *On Language: Chomsky's Classic Works Language and Responsibility and Reflections on Language*. The New Press.

Chomsky, N. 1975 [1998]. *Reflections on Language*. In *On Language: Chomsky's Classic Works Language and Responsibility and Reflections on Language*. The New Press.

Chomsky, N. 1965. *Aspects of the Theory of Syntax*. MIT Press.

Chomsky, N. 1957. *Syntactic Structures*. Mouton de Gruyter.

Christiansen, M.H. & Chater, N. 2022. *The Language Game: How Improvisation Created Language and Changed the World*. Bantam Press.

Cissewski, J. & Luncz, L.V. 2021. 'Symbolic Signal Use in Wild Chimpanzee Gestural Communication? A Theoretical Framework'. *Frontiers in Psychology* 12, 718414.

Clark, A.G., Glanowski, S., Nielsen, R., Thomas, P.D., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X., White, T.J., Sninsky, J.J., Adams, M.D. & Cargill, M. 2003. 'Inferring Nonneutral Evolution from Human-Chimp-Mouse Orthologous Gene Trios'. *Science* 302, 1960-1963.

Clark, E.V. 2003. *First Language Acquisition*. Cambridge University Press.

Clark, G. & Henneberg, M. 2017. 'Ardipithecus ramidus and the evolution of language and singing: An early origin for hominin vocal capability'. *HOMO – Journal of Comparative Human Biology* 68:2, 101-121.

Clark, R. & Kimbrough, S.O. 2017. 'Social structure, opportunistic punishment and the evolution of honest signaling'. *PLoS ONE* 12:12, e0188249.

Clutton-Brock, T. H. & Isvaran, K. 2006. 'Paternity loss in contrasting mammalian societies'. *Biology Letters* 2: pp513-516.

Comrie, B. 1976. *Aspect*. Cambridge University Press.

Conard, N.J. 2005. 'An overview of the patterns of behavioural change in Africa and Eurasia during the Middle and Late Pleistocene'. In F. d'Errico & L. Backwell (eds.), *From Tools to Symbols: From Early Hominids to Modern Humans*. Wits University Press.

Conard, N.J. 2003. 'Palaeolithic ivory sculptures from southwestern Germany and the origins of figurative art'. *Nature*, 426-6968, 830-832.

Connor, R.C., Krützen, M., Allen, S.J., Sherwin, W.B. & King, S.L. 2022. 'Strategic intergroup alliances increase access to a contested resource in male bottlenose dolphins'. *Proceedings of the National Academy of Science* 119:36, e2121723119.

Cook, G. 2000. *Language Play, Language Learning*. Oxford University Press.

Cook, V.J. & Newson, M. 1996. *Chomsky's Universal Grammar: An Introduction (2nd edition)*. Blackwell Publishers Ltd.

Cooper, B.G. & Goller, F. 2004. 'Multimodal Signals: Enhancement and Constraint of Song Motor Patterns by Visual Display'. *Science* 303, 544-546.

Corballis, M.C. 2007. 'The Uniqueness of Human Recursive Thinking'. *American Scientist* 95:3, 240-248.

Corballis, M.C. 2003. *From Hand to Mouth: The Origins of Language*. Princeton University Press.

Corballis, M.C. 2002. 'Did Language Evolve from Manual Gestures?' In A. Wray (ed.), *The Transition to Language: Studies in the Evolution of Language*, 161-179. Oxford University Press.

Corning, P.A. 2000. 'Biological Adaptation in Human Societies: a "basic needs" approach'. *Journal of Bioeconomics* 2, 41-86.

Cosmides, L. & Tooby, J. 1994. 'Beyond intuition and instinct blindness: towards an evolutionary rigorous cognitive science'. *Cognition*, 50, 41-77.

Craig, H.K. 1995. 'Pragmatic Impairments'. In P. Fletcher & B. MacWhinney (eds.), *The Handbook of Child Language*. Blackwell Publishers.

Creider, C. 2000. 'Mixed Categories in Word Grammar: Swahili infinitival Nouns'. *Linguistica Atlantica* 21, 53-68.

Croft, W. 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford University Press.

Crompton, R.H., Vereecke, E.E. & Thorpe, S.K.S. 2008. 'Locomotion and posture from the common hominoid ancestor to fully modern hominins, with special reference to the last common panin/hominin ancestor'. *Journal of Anatomy* 212, 501-543.

Cronin, H. 2005. 'Adaptation: A Critique of some Current Evolutionary Thought'. *Quarterly Review of Biology* 80:1, 19-26.

Currie, T.E., Campenni, M., Flitton, A., Njagi, T., Ontiri, E., Perret, C. & Walker, L. 2021. 'The cultural evolution and ecology of institutions'. *Philosophical Transactions of the Royal Society B* 376, 20200047.

Dąbrowska, E. 2016. 'Cognitive Linguistics' seven deadly sins'. *Cognitive Linguistics* 27:4, 479-491.

Dąbrowska, E. 2004. *Language, Mind and Brain: Some Psychological and Neurological Constraints on Theories of Grammar*. Edinburgh University Press.

Darwin, C.R. 1897. *The Expression of the Emotions in Man and Animals*. D. Appleton & Company.

Darwin, C.R. 1859. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. John Murray.

Kerstin Dautenhahn 2003. 'Stories of lemurs and robots: The social origin of story-telling'. In M. Mateas & P. Sengers (eds.), *Narrative Intelligence*, 63-90. John Benjamins Publishing Company.

Daver, G., Guy, F., Mackaye, H.T., Likius, A., Boisserie, J-R., Moussa, A., Pallas, L., Vignaud, P. & Clarisse, N.D. 2022. ‘Postcranial evidence of late Miocene hominin bipedalism in Chad’. *Nature,* s41586-022-04901-z.

David-Barrett, T. 2023. ‘Human group size puzzle: why it is odd that we live in large societies’. *Royal Society Open Science* 10, 230559.

Davidson, I. 2016. ‘Stone Tools: Evidence of Something in Between Culture and Cumulative Culture?’ In M.N. Haidle, N.J. Conard & M. Bolus (eds.), *The Nature of Culture,* 99-120. Springer.

Davidson, I. & McGrew, W.C. 2005. ‘Stone Tools and the Uniqueness of Human Culture’. *Journal of the Royal Anthropological Institute* (NS) 11, 793-817.

Dawkins, R. 1996. *Climbing Mount Improbable*. Penguin Books.

Dawkins, R. 1989. *The Selfish Gene (2nd edition)*. Oxford University Press.

Day, M.H. 1986. ‘Bipedalism: pressures, origins, and modes’. In B. Wood, L. Martin & P. Andrews (eds.), *Major Topics in Primate and Human Evolution*. Cambridge University Press.

Deacon, T. 1997. *The Symbolic Species: The Co-Evolution of Language and the Human Brain*. Penguin Books.

Deamer, F. 2020. ‘Why Do We Talk to Ourselves?’ *Review of Philosophy and Psychology* 12, 425-433.

Dediu, D. & Levinson, S.C. 2013. ‘On the antiquity of language: the reinterpretation of Neandertal linguistic capacities and its consequences’. *Frontiers in Psychology* 4, 397.

Dehaene, S. 2014. *Consciousness and the Brain: Deciphering How the Brain Codes Our Thoughts*. Viking.

Dennett, D. 2009. ‘Darwin’s strange inversion of reasoning’. *Proceedings of the National Academy of Science* 106:s1, 10,061-10,065.

Denton, D. K. 2006. ‘Keep it Simple: business decisions don’t have to be complicated to be good’. *Industrial Engineer,* May 2006, 30-35.

Descartes, R. 1641 [1998]. ‘Meditations on First Philosophy: replies to second objections’. In *Meditations and Other Metaphysical Writings*. Penguin Books.

DeSilva, J.M., Traniello, J.F.A., Claxton, A.G. & Fannin, L.D. 2021. ‘When and Why Did Human Brains Decrease in Size? A New Change-Point Analysis and Insights from Brain Evolution in Ants’. *Frontiers in Ecology and Evolution* 9, 742639.

Dessalles, J.L. 2014a. ‘The Role of the Human Political Singularity in the Emergence of Language’. In E.A. Cartmill, S. Roberts, H. Lyn & H. Cornish (eds.), *The Evolution of Language: Proceedings of the 10th International Conference (Evolang-X - Vienna),* 423-424. World Scientific.

Dessalles, J.L. 2014b. ‘Optimal Investment in Social Signals’. *Evolution: International Journal of Organic Evolution* 68:6, 1640-1650.

Dessalles, J.L. 2007. *Why We Talk: the evolutionary origins of language*. Oxford University Press.

Dessalles, J.L. 2000. ‘Language and Hominid Politics’. In C. Knight, M. Studdert-Kennedy & J.R. Hurford (eds.), *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form*. Cambridge University Press.

Détroit, F., Mijares, A.S., Corny, J., Daver, G., Zanolli, C., Dizon, E., Robles, E., Grün, R. & Piper, P.J. 2019. ‘A new species of Homo from the Late Pleistocene of the Philippines’. *Nature* 568, 181-186.

Deutscher, G. 2005. *The Unfolding of Language: The Evolution of Mankind’s Greatest Invention*. Arrow Books.

Diamond, J. 2011. ‘Deep Relationships between Languages’. *Nature* 476, 291-292.

Dickins, J. 2009. ‘Extended axiomatic functionalism: postulates’. *Linguistica Online*, 1-52. http://www.phil.muni.cz/linguistica/art/dickins/dic-001.pdf

Dickins, T.E. 2003. ‘General Symbol Machines: the first stage in the evolution of symbolic communication’. *Evolutionary Psychology* 1, 192-209.

Dieckmann, U. & Doebeli, M. 1999. ‘On the origin of species by sympatric speciation’. *Nature* 400, 354-357.

Dielenberg, R. 2013. ‘The comparative psychology of human uniqueness: A cognitive behavioral review’. In R.G. Bednarik (ed.), *The Psychology of Human Behavior*, 111-182. Nova Science Publishers.

Dik, S.C. 1981. *Functional Grammar*. De Gruyter.

Dingemanse, M., Torreira, F. & Enfield, N.J. 2013. ‘Is ‘‘Huh?’’ a Universal Word? Conversational Infrastructure and the Convergent Evolution of Linguistic Items’. *PLoS ONE* 8:11, e78273.

Diogo, R., Molnar, J.L. & Wood, B. 2017. ‘Bonobo anatomy reveals stasis and mosaicism in chimpanzee evolution, and supports bonobos as the most appropriate extant model for the common ancestor of chimpanzees and humans’. *Scientific Reports* 7, 608.

Dixon, R.M.W. 1997. *The Rise and Fall of Languages*. Cambridge University Press.

Domínguez-Rodrigo, M., Pickering, T.R., Semaw, S. & Rogers, M.J. 2005. ‘Cutmarked bones from Pliocene archaeological sites at Gona, Afar, Ethiopia: implications for the function of the world’s oldest stone tools’. *Journal of Human Evolution* 48, 109-121.

Donald, M. 2001. *A Mind So Rare: The Evolution of Human Consciousness*. WW Norton & co.

Dor, D. 2017a. ‘From experience to imagination: Language and its evolution as a social communication technology’. *Journal of Neurolinguistics* 43, 107-119.

Dor, D. 2017b. ‘The role of the lie in the evolution of human language’. *Language Sciences* 63, 44-59.

Dores Cruz, T.D., Thielmann, I., Columbus, S., Molho, C., Wu, J., Righetti, F., de Vries, R.E., Koutsoumpis, A., van Lange, P.A.M., Beersma, B. & Balliet, D. 2021. ‘The language of cooperation: reputation and honest signalling’. *Philosophical Transactions of the Royal Society B* 376, 20200301.

Draper, H. & Ives, J. 2009. 'Paternity testing: a poor test of fatherhood'. *Journal of Social Welfare and Family Law* 31:4, 407-418.

Dufour, V., Pasquaretta, C., Gayet, P. & Sterck, E.H.M. 2017. 'The Extraordinary Nature of Barney's Drumming: A Complementary Study of Ordinary Noise Making in Chimpanzees'. *Frontiers in Neuroscience* 11, 2.

Dunbar, R.I.M. 2004. *The Human Story: A New History of Mankind's Evolution*. Faber & Faber Ltd.

Dunbar, R.I.M. 1996. *Grooming, Gossip and the Evolution of Language*. Faber & Faber Ltd.

Dunbar, R.I.M. 1993. 'Co-Evolution of Neocortex Size, Group Size and Language in Humans'. *Behavioral and Brain Sciences* 16:4, 681-735.

Dunbar, R.I.M. 1992. 'Why gossip is good for you'. *New Scientist* 136:1848, 28.

Dunning, D., Heath, C. & Suls, J.M. 2004. 'Flawed Self-Assessment: Implications for Health, Education, and the Workplace'. *Psychological Science in the Public Interest* 5:3, 69-106.

Dunsworth, H.M. 2010. 'Origin of the Genus Homo'. *Evolution: Education and Outreach* 3, 353-366.

Duval, M., Sahnouni, M., Josep M. Parés, J.M., van der Made, J., Abdessadok, S., Harichane, Z., Cheheb, R.C., Boulaghraif, K. & Pérez-González, A. 2021. 'The Plio-Pleistocene sequence of Oued Boucherit (Algeria): A unique chronologically-constrained archaeological and palaeontological record in North Africa'. *Quaternary Science Reviews* 271, 107116.

Eckert, P. & McConnell-Ginet, S. 1998. 'Communities of Practice: where language, gender, and power all live'. In J. Coates (ed.), *Language and Gender: A Reader*. Blackwells Publishing.

Edwardes, M.P.J. 2019. *The Origins of Self: An Anthropological Perspective*. UCL Press.

Edwardes, M.P.J. 2018. 'Things Do Things to Things: Social calculus, agentive grammar and the beginnings of language'. *Presented at The Personification across Disciplines conference, Durham University, 17-19 September 2018*. http://martinedwardes.me.uk/academia_conferences_papers/thingsdothingstothings.pdf.

Edwardes, M.P.J. 2014a. 'Awareness of Self and Awareness of Selfness: Why the Capacity to Self-Model Represents a Novel Level of Cognition in Humans'. In G. Rundblad, A. Tytus, O. Knapton & C. Tang (eds.), *Selected Papers from the 4th UK Cognitive Linguistics Conference*, 68-83. UK Cognitive Linguistics Association.

Edwardes, M.P.J. 2014b. 'What were we talking about? Exchanging social models as a route to language'. In E.A. Cartmill, S. Roberts, H. Lyn & H. Cornish (eds.), *The*

Evolution of Language: Proceedings of the 10th International Conference (EVOLANG10), Vienna Austria, 14-17 April 2014, 58-65. World Scientific.

Edwardes, M.P.J. 2011. 'Are Complex Tenses Really Real?' In R. McColl Millar & M. Durham (eds.), *Applied Linguistics, Global and Local: Proceedings of the 43rd Annual Meeting of the British Association for Applied Linguistics, 9-11 September 2010, University of Aberdeen*, 69-78. Scitsiugnil Press.

Edwardes, M.P.J. 2010. *The Origins of Grammar: An Anthropological Perspective*. Continuum.

Edwardes, M.P.J. 2005. 'The Ideational Metafunction and Nonhuman Signalling'. *Presented at the 17th European Systemic Functional Linguistics Conference, 1-4 August 2005*. http://martinedwardes.me.uk/academia_conferences_papers/ideationandnonhumansignalling.pdf

Edwardes, M.P.J. 2003. 'I like Both Myself and Me'. In D. Hall, T. Markopoulos, A. Salamoura & S. Skoufaki (eds.), *Camling 2003: Proceedings of the University of Cambridge first postgraduate conference in language research*. CILR.

Edwardes, M.P.J. 2002. 'Playing Other Games with Language'. *Personal papers*, http://martinedwardes.me.uk/academia_sundry_essays/playing_language_games.pdf (25 April 2022).

Egas, M. & Riedl, R. 2005. 'The Economics of Altruistic Punishment and the Demise of Cooperation'. *Forschungsinstitut zur Zukunft der Arbeit*, discussion paper 1646, July 2005.

Eleuteri, V., Henderson, M., Soldati, A., Badihi, G., Zuberbühler, K. & Hobaiter, C. 2022. 'The form and function of chimpanzee buttress drumming'. *Animal Behaviour* 192, 189-205.

Ellis, E.M., Borovsky, A., Elman, J.L. & Evans, J.L. 2021. 'Toddlers' Ability to Leverage Statistical Information to Support Word Learning'. Frontiers in Psychology 12, 600694.

Embick, D., Marantz, A., Miyashita, Y., O'Neil, W. & Sakai, K.L. 2000. 'A syntactic specialization for Broca's area'. *Proceedings of the National Academy of Science* 97:11, 6150-6154.

Engelmann, J.M. & Tomasello, M. 2019. 'Children's Sense of Fairness as Equal Respect'. *Trends in Cognitive Sciences* 23:6, 454-463.

Engh, A.L., Hoffmeier, R.R., Cheney, D.L. & Seyfarth, R.M. 2006. 'Who, me? Can baboons infer the target of vocalizations?' *Animal Behaviour* 71, 381-387.

Epley, N. & Whitchurch, E. 2008. 'Mirror, Mirror on the Wall: Enhancement in Self-Recognition'. *Personality and Social Psychology Bulletin*, 34:9, 1159-1170.

Erdal, D. & Whiten, A. 1994. 'On Human Egalitarianism: an evolutionary product of Machiavellian status escalation?' *Current Anthropology*, 35:2, 175-183.

Etcoff, N.L., Ekman, P., Magee, J.J. & Frank, M.G. 2000. 'Lie detection and language comprehension'. *Nature* 405, 139.

Ettlinger, M., Margulis, E.H. & Wong, P.C.M. 2011. 'Implicit memory in music and language'. *Frontiers in Psychology* 2, 211.

Evans, N. & Levinson, S.C. 2009. 'The myth of language universals: Language diversity and its importance for cognitive science'. *Behavioral and Brain Sciences* 32, 429-492.

Evans, V. 2014. 'Real Talk'. *Aeon*, https://aeon.co/essays/the-evidence-is-in-there-is-no-language-instinct.

Evans, V. 2007.'How We Conceptualise Time: language, meaning and temporal cognition'. In V. Evans, B.K. Bergen & J. Zinken (eds.), *The Cognitive Linguistics Reader*. Equinox Publishing.

Evans, V. 2005. *The Structure of Time*. John Benjamins.

Evans, V., Bergen, B.K. & Zinken, J. 2007. 'The Cognitive Linguistics Enterprise: An Overview'. In V. Evans, B.K. Bergen & J. Zinken (eds.), *The Cognitive Linguistics Reader*. Equinox Publishing.

Evans, V. & Green, M. 2006. *Cognitive Linguistics: an introduction*. Edinburgh University Press.

Eysenck, M.W. 1993. *Principles of Cognitive Psychology*. Psychology Press.

Eysenck, M.W. & Keane, M.T. 1995. *Cognitive Psychology: a student's handbook (3rd edition)*. Psychology Press.

Faisal, A., Stout, D., Apel, J. & Bradley, B. 2010. 'The Manipulative Complexity of Lower Paleolithic Stone Toolmaking'. *PLoS ONE* 5:11, e13718.

Fedurek, P., Slocombe, K.E., Hartel, J.A. & Zuberbühler, K. 2015. 'Chimpanzee lip-smacking facilitates cooperative behaviour'. *Nature Scientific Reports* 5, 13460.

Feeney, A. & Edwards, R. 2021. 'Language evolution, narrative and the nature of cognition'. *Journal of Cognitive Science* 22:4, 487-539.

Fehr, E. & Renninger, S.V. 2004. 'The Samaritan Paradox'. *Scientific American Mind*, November 2004, 15-21.

Fehr, E. & Fischbacher, U. 2005. 'Human Altruism – Proximate Patterns and Evolutionary Origins'. *Analyse & Kritik* 27, 6-47.

Fehr, E. & Gächter, S. 2002. 'Altruistic punishment in humans'. *Nature* 415:6868, 137-140.

Fenici, M. 2012. 'Embodied Social Cognition and Embedded Theory of Mind'. *Biolinguistics* 6.3-4, 276-307.

Fillmore, C. 1976. 'Frame semantics and the nature of language'. *Annals of the New York Academy of Sciences: Conference on the Origin and Development of Language and Speech* 280, 20-32.

Fillmore, C. 1971. 'Types of Lexical Information'. In D.D. Steinberg & L.A. Jakobovits (eds.), *Semantics: an interdisciplinary reader in Philosophy, Linguistics and Psychology*. Cambridge University Press.

Fischer, G., Azorsa, F., Garcia, F.H., Mikheyev, A.S. & Economo, E.P. 2015. 'Two new phragmotic ant species from Africa: morphology and next-generation sequencing solve a caste association problem in the genus Carebara Westwood'. *ZooKeys* 525, 77-105.

Fisher, A.M., Holwell, G.I. & Price, T.A.R. 2020. 'Behavioural correlations and aggression in praying mantids'. *Behavioral Ecology and Sociobiology* 74, 61.

Fitch, W.T. 2011. 'The evolution of syntax: an exaptationist perspective'. *Frontiers in Evolutionary Neuroscience* 3:9.

Fitch, W.T. 2010. 'How language evolved: Hearing lost words'. *New Scientist* 2789.

Fitch, W.T. & Hauser, M.D. 2003. 'Unpacking "Honesty": Vertebrate Vocal Production and the Evolution of Acoustic Signals'. In A. Simmons, R.R. Fay & A.N. Popper (eds.), *Acoustic Communication*. Springer.

FitzSimons, G. E. 2019. 'Adults learning mathematics: transcending boundaries and barriers in an uncertain world'. *Adults Learning Mathematics: An International Journal*, 14:1, 41-52.

Flack, J.C., Girvan, M., de Waal, F.B.M. & Krakauer, D.C. 2006. 'Policing stabilizes construction of social niches in primates'. *Nature* 439-7075, 426-429.

Flack, J.C., Jeannotte, L.A. & de Waal, F.B.M. 2004. 'Play Signaling and the Perception of Social Rules by Juvenile Chimpanzees (Pan Troglodytes)'. *Journal of Comparative Psychology*, 118:2, 149-159.

Flower, T. 2011. 'Fork-tailed drongos use deceptive mimicked alarm calls to steal food'. *Proceedings of the Royal Society B* 278, 1548-1555.

Fodor, J.A. 2008. *LOT2: The Language of Thought Revisited*. Oxford University Press.

Fodor, J.A. 1983. *The Modularity of Mind*. MIT Press.

Foitzik, S. & Fritsche, O. 2021. *Empire of Ants: The hidden Worlds and extraordinary lives of Earth's tiny conquerors*. Rowohlt Verlag GmbH.

Foley, J. & Thompson, L. 2003. *Language Learning: a lifelong process*. Arnold.

Fossey, D. 1983. *Gorillas in the Mist*. Phoenix.

Fouts, R. with Mills, S.T. 1997. *Next of Kin: my conversations with chimpanzees*. Avon Books Inc.

Fowler, J.H. 2005. 'Altruistic Punishment and the Origin of Cooperation'. *Proceedings of the National Academy of Science*, 102:19, 7047-7049.

Fowler, J.H., Dawes, C.T. & Christakis, N.A. 2009. 'Model of genetic variation in human social networks'. *Proceedings of the National Academy of Science* 106:6, 1720-1724.

Fraser, P., Solé, R. & de las Cuevas, G. 2021. 'Why Can the Brain (and Not a Computer) Make Sense of the Liar Paradox?' *Frontiers in Ecology and Evolution* 9, 802300.

Freud, S. 1923 [1961]. 'The Ego and the Id'. In *The Standard Edition of the Complete Psychological Works of Sigmund Freud, Volume XIX*. Hogarth Press.

Frey, S.H. 2008. ‘Tool use, communicative gesture and cerebral asymmetries in the modern human brain’. *Philosophical Transactions of the Royal Society B* 363, 1951-1957.

Friedman, W.J. 2000. ‘The Development of Children’s Knowledge of the Times of Future Events’. *Child Development* 71:4, 913-932.

von Frisch, K. 1973. ‘Decoding the Language of the Bee’. *Nobel Lecture*, December 12, 1973.

Frith, U. 2003. *Autism: Explaining the Enigma (2nd edition)*. Blackwell Publishing.

Frith, U. 1993 [1999]. ‘Autism’. *The Scientific American Book of the Brain*. The Lyons Press.

Fröhlich, M., Sievers, C., Townsend, S.W., Gruber, T. & van Schaik, C.P. 2019. ‘Multimodal communication and language origins: integrating gestures and vocalizations’. *Biological Reviews* 94:5, 1809-1829.

Fruth, B. & Hohmann, G. 2018. ‘Food Sharing across Borders: First Observation of Intercommunity Meat Sharing by Bonobos at LuiKotale, DRC’. *Human Nature* 29, 91-103.

Fruth, B. & Hohmann, G. 1993. ‘Ecological and Behavioral Aspects of Nest Building in Wild Bonobos (Pan paniscus)’. *Ethology* 94:2, 113-126.

Fudge, E.C. 1995. ‘The Glossematic School of Linguistics’. In E.F.K. Koerner & R.E. Asher (eds.), *Concise History of the Language Sciences: From the Sumerians to the Cognitivists*, 262-268. Pergamon.

Fuentes, A., Wyczalkowski, M.A. & MacKinnon, K.C. 2010. ‘Niche Construction through Cooperation: A Nonlinear Dynamics Contribution to Modeling Facets of the Evolutionary History in the Genus Homo’. *Current Anthropology* 51:3, 435-444.

Futuyma, D.J. 2015. ‘Can Modern Evolutionary Theory Explain Macroevolution?’ In E. Serrelli & N. Gontier (eds.), *Macroevolution*, 29-85. Springer International Publishing.

Gabrić, P., Banda, M. & Karavanić, I. 2018. ‘Palaeolithic toolmaking and the evolution of cognition and language’. In T. Strle & O. Markič (eds.), *Proceedings of the 21st Information Society International Multiconference IS 2018, Volume B, 8-12 October 2018*. Institut Jožef Stefan, Ljubljana.

Gabunia, L. & Vekua, A. 1995. ‘A Plio-Pleistocene hominid from Dmanisi, East Georgia, Caucasus’. *Nature* 373, 509-512.

Gagliardi, C., Arrigoni, F., Nordio, A., de Luca, A., Peruzzo, D., Decio, A., Leemans, A. & Borgatti, R. 2018. ‘A Different Brain: Anomalies of Functional and Structural Connections in Williams Syndrome’. *Frontiers in Neurology* 9, 721.

Gallup, A.C. 2007. ‘Yawning as a Brain Cooling Mechanism: Nasal Breathing and Forehead Cooling Diminish the Incidence of Contagious Yawning’. *Evolutionary Psychology* 5:1, 92-101.

Gallup, G.G. 1970. ‘Chimpanzees: self-recognition’. *Science*, 167-3194, 86-87.
Ganapini, M.B. 2023. ‘The signaling function of sharing fake stories’. *Mind & Language* 38:1, 64-80.
Ganea, P.A. 2005. ‘Contextual Factors Affect Absent Reference Comprehension in 14-Month-Olds’. *Child Development*, 76:5, 989-998.
Gardner, H. 1983. *Frames of Mind: The Theory of Multiple Intelligences*. Basic Books.
Gardner, R.A. & Gardner, B.T. 1969. ‘Teaching Sign Language to a Chimpanzee’. *Science* 165:3894, 664-672.
Gardner, A. & West, S.A. 2004. ‘Cooperation and Punishment, Especially in Humans’. *The American Naturalist* 164:6, 753-764.
Gavrilets, S. & Vose, A. 2006. ‘The dynamics of Machiavellian intelligence’. *Proceedings of the National Academy of Science*, 103:45, 16823-16828.
Gaylard, H.L. 1995. *Phrase Structure in a Computational Model of Child Language Acquisition*. PhD thesis, School of Computer Science, University of Birmingham.
Geeraerts, D. 2010. ‘Recontextualizing Grammar: Underlying trends in thirty years of Cognitive Linguistics’. In E. Tabakowska, M. Choinski & L. Wiraszka (eds.), *Cognitive Linguistics in Action: From Theory to Application and Back*. De Gruyter Mouton.
Geeraerts, D. 2006. ‘Introduction: a rough guide to Cognitive Linguistics’. In D. Geeraerts (ed.), *Cognitive Linguistics: Basic Readings*. Mouton de Gruyter.
Geissmann, T. 2000. ‘Gibbon songs and human music from an evolutionary perspective’. In N.L. Wallin, B. Merker, & S. Brown (eds.), *The Origins of Music*, 103-123. MIT Press.
Gendron, M., Roberson, D., van der Vyver, J.M. & Barrett, L.F. 2014. ‘Perceptions of Emotion from Facial Expressions Are Not Culturally Universal: Evidence from a Remote Culture’. *Emotion*, 14:2, 251-262.
Gibbs, R.W. & Perlman, M. 2006. ‘The Contested Impact of Cognitive Linguistic Research on the Psycholinguistics of Metaphor Understanding’. In G. Kristiansen, M. Achard, R. Dirven & F.J. Ruiz de Mendoza Ibáñez (eds.), *Cognitive Linguistics: Current Applications and Future Perspectives*. Mouton de Gruyter.
Gibson, K.R. 2002. ‘Evolution of Human Intelligence: The Roles of Brain Size and Mental Construction’. *Brain, Behavior and Evolution* 59: pp10-20.
Gilboa, A. & Moscovitch, M. 2021. ‘No consolidation without representation: Correspondence between neural and psychological representations in recent and remote memory’. *Neuron* 109, 2239-2255.
Gillespie-Lynch, K., Greenfield, P.M., Lyn, H. & Savage-Rumbaugh, S. 2014. ‘Gestural and symbolic development among apes and humans: support for a multimodal theory of language evolution’. *Frontiers in Psychology* 5, 1228.
Gintis, H. 2008. ‘Punishment and Cooperation’. *Science* 319:5868, 1345-1346.
Giomi, R. 2020. ‘Headedness and modification in Functional Discourse Grammar’. *Glossa: a Journal of General Linguistics* 5:1, 118.

Giroux, H. 2017. ‘Donald Trump’s passion for cruelty’. *The Conversation*, https://theconversation.com/donald-trumps-passion-for-cruelty-84819.

Godfrey-Smith, P. 2016. *Other Minds: The Octopus and the Evolution of Intelligent Life*. William Collins.

Goldin-Meadow, S., & Alibali, M.W. 2013. ‘Gesture’s role in speaking, learning, and creating language’. *Annual Review of Psychology* 64, 257-283.

Goldstein, M.C., Jiao, B., Beall, C.M. & Tsering, P. 2002. ‘Fertility and Family Planning in Rural Tibet’. *The China Journal* 47, 19-39.

Goleman, D. 1995. *Emotional Intelligence: Why It Can Matter More Than IQ*. Bloomsbury.

Gómez-Olivencia, A., Crevecoeur I. & Balzeau, A. 2015. ‘La Ferrassie 8 Neandertal child reloaded: New remains and re-assessment of the original collection’. *Journal of Human Evolution* 82, 107-126.

Goodall, J. 1990. *Through a Window: Thirty Years with the Chimpanzees of Gombe*. Phoenix.

Goodall, J. 1988. *In the Shadow of Man*. Weidenfeld & Nicolson.

Goodman, J.R. 2022. ‘Evolution of language can help us sift truth from lies in modern world’. *New Scientist* 3371, https://www.newscientist.com/article/mg25333713-000-evolution-of-language-can-help-us-sift-truth-from-lies-in-modern-world/.

Goodman, J.R., Caines, A. & Foley, R.A. 2023. ‘Shibboleth: An agent-based model of signalling mimicry’. *PLoS ONE* 18:7, e0289333.

Gopnik, A., Meltzoff, A. & Kuhl, P. 1999. *How Babies Think*. Weidenfeld & Nicolson.

Gothard, K.M. 2014. ‘The amygdalo-motor pathways and the control of facial expressions’. *Frontiers in Neuroscience* 8, 43.

Gotts, S.J., Job, H.J., Wallace, G.L., Saad, Z.S., Cox, R.W. & Martin, A. 2013. ‘Two distinct forms of functional lateralization in the human brain’. *Proceedings of the National Academy of Science* 110:36, E3435-E3444.

Gould, J.L. & Gould, C.G. 1995. *The Honey Bee*. Scientific American Library.

Gould, S.J. 1981. *The Mismeasure of Man*. Penguin.

Gould, S.J. & Lewontin, R.C. 1979 [1997]. ‘The Spandrels of San Marco and the Panglossian Paradigm: a critique of the adaptationist programme’. In M. Ridley (ed.), *Evolution*. Oxford University Press.

Grammer, K., Schiefenhövel, W., Schleidt, M., Lorenz, B. & Eibl-Eibesfeldt, I. 1988. ‘Patterns on the Face: The Eyebrow Flash in Crosscultural Comparison’. *Ethology* 77:4, 279-299.

Granberg, Å., Armelius, K. & Armelius, B-A. 2001. ‘Self-Image Pattern and Treatment Outcome in Severely Disturbed Psychiatric Patients’. *Umeå Psychology Reports* 1, 1-13.

Gravina, B., Mellars, P. & Ramsey, C.B. 2005. ‘Radiocarbon dating of interstratified Neanderthal and early modern human occupations at the Chatelperronian type-site’. *Nature* 438:7064, 51-56.

Gredebäck, G., Lindskog, M., Juvrud, J.C., Green, D. & Marciszko, C. 2018. 'Action Prediction Allows Hypothesis Testing via Internal Forward Models at 6 Months of Age'. *Frontiers in Psychology* 9, 290.

Green, R.E., Krause, J., Ptak, S.E., Briggs, A.W., Ronan, M.T., Simon, J.F., Du, L., Egholm, M., Rothberg, J.M., Paunovic, M. & Pääbo, S. 2006. 'Analysis of one million base pairs of Neanderthal DNA'. *Nature* 444, 330-336.

Greenfield, P.M. 1991. 'Language, tools and brain: The ontogeny and phylogeny of hierarchically organized sequential behavior'. *Behavioral and Brain Sciences* 14, 531-595.

Greenfield, S. 2000. *Brain Story: Unlocking our Inner World of Emotions, Memories, Ideas and Desires*. BBC Worldwide.

Greenspan, S.I. 2001. 'The Affect Diathesis Hypothesis: The Role of Emotions in the Core Deficit in Autism and in the Development of Intelligence and Social Skills'. *Journal of Developmental and Learning Disorders* 5:1, 1-45.

Grice, P. 1989. *Studies in the Way of Words*. Harvard University Press.

Grueter, C.C., Qi, X., Zinner, D., Bergman, T., Li, M., Xiang, Z., Zhu, P., Migliano, A.B., Miller, A., Krützen, M., Fischer, J., Rubenstein, D.I., Vidya, T.N.C., Li, B., Cantor, M. & Swedell, L. 2020. 'Multilevel Organisation of Animal Sociality'. *Trends in Ecology & Evolution* 35:9, 834-847.

Güntürkün, O., Ströckens, F. & Ocklenburg, S. 2020. 'Brain Lateralization: A Comparative Perspective'. *Physiology Review* 100: 1019-1063.

Haeckel, E. 1874 [1912]. *The Evolution of Man: A Popular Scientific Study*. Watts & Co.

Halliday, M.A.K. 2004. *The Language of Early Childhood*. Continuum.

Halliday, M.A.K. 1994. *An Introduction to Functional Grammar (2nd edition)*. Edward Arnold.

Halliday, M.A.K. 1985. *An Introduction to Functional Grammar*. Arnold.

Halliday, M.A.K. 1969 [2003]. 'A Brief Sketch of Systemic Grammar'. In J.J. Webster (ed.), *On Language and Linguistics*, ch7. Continuum.

Halliday, M.A.K. & Matthiessen, C.M.I.M. 2014. *Halliday's Introduction to Functional Grammar*. Routledge.

Halliday, M.A.K. & Matthiessen, C.M.I.M. 2004. *An Introduction to Functional Grammar (3rd edition)*. Routledge.

Hamilton, W.D. 1964. 'The Genetical Evolution of Social Behaviour II'. *The Journal of Theoretical Biology*, 7, 17-52.

Hanks, W.F. 2005. 'Explorations in the Deictic Field'. *Current Anthropology* 46-2, 191-220.

Harcourt-Smith, W. 2013. 'The Origins of Bipedal Locomotion'. In W. Henke & I. Tattersall (eds.), *Handbook of Paleoanthropology*, 1483-1518. Springer-Verlag.

Harder, P. 1996. *Functional Semantics: A Theory of Meaning, Structure and Tense in English*. Mouton de Gruyter.

Hare, B. & Tomasello, M. 2004. 'Chimpanzees are more skilful in competitive than in cooperative cognitive tasks'. *Animal Behaviour* 68, 571-581.

Harman, G.H. 1968 [1971]. 'Three Levels of Meaning'. In D.D. Steinberg & L.A. Jakobovits (eds.), *Semantics: An Interdisciplinary Reader in Philosophy, Linguistics and Psychology*. Cambridge University Press.

Harmand, S., Lewis, J.E., Feibel, C.S., Lepre, C.J., Prat, S., Lenoble, A., Boës, X., Quinn, R.L., Brenet, M., Arroyo, A., Taylor, N., Clément, S., Daver, G., Brugal, J-P., Leakey, L., Mortlock, R.A., Wright, J.D., Lokorodi, S., Kirwa, C., Kent, D.V. & Roche, H. 2015. '3.3-million-year-old stone tools from Lomekwi 3, West Turkana, Kenya'. *Nature* 521, 310-315.

Harris, R.A. 1993. *The Linguistics Wars*. Oxford University Press.

Harrod, J.B. 2014. 'The case for chimpanzee religion'. *Journal for the Study of Religion, Nature and Culture* 8:1, 8-45.

Hartwigsen, G., Bengio, Y. & Bzdok, D. 2021. 'How does hemispheric specialization contribute to human-defining cognition?' *Neuron* 04, 024.

Hauser, M.D., Chomsky, N. & Fitch, W.T. 2002. 'The Faculty of Language: what is it, who has it, and how did it evolve?' *Science* 298:5598, 1569-1579.

Hawkes, K. 2004. 'The Grandmother effect'. *Nature* 428, 128-129.

Hawkes, K., O'Connell, J.F., Blurton Jones, N.G., Alvarez, H. & Charnov, E.L. 1998. 'Grandmothering, menopause, and the evolution of human life histories'. *Proceedings of the National Academy of Science* 95, 1336-1339.

Hawking, S. 1996. *The Illustrated a Brief History of Time*. Bantam Press.

Hayes, K.J. & Nissen, C.H. 1971. 'Higher Mental Functions of a Home-Raised Chimpanzee'. In A.M. Schrier & F. Stollnitz (eds.), *Behavior of Nonhuman Primates*, 59-115. Academic Press.

Hayes, M.P. 2015. 'The biology and ecology of the large blue butterfly Phengaris (Maculinea) arion: a review'. *Journal of Insect Conservation* 19, 1037-1051.

Hayne, H., Imuta, K. & Scarf, D. 2015. Memory Development During Infancy and Early Childhood across Cultures. In J.D. Wright (ed.-in-chief), *International Encyclopedia of the Social & Behavioral Sciences*, 2nd edition, Vol. 15, 147-154. Elsevier.

Heald, J.B., Lengyel, M. & Wolpert, D.M. 2023. 'Contextual inference in learning and memory'. *Trends in Cognitive Sciences* 27:1, 43-64.

Healy, J.M. 1994. *Your Child's Growing Mind: A Practical Guide to Brain Development and Learning from Birth to Adolescence*. Doubleday.

Heine, B. & Kuteva, T. 2012. 'Grammaticalization Theory as a Tool for Reconstructing Language Evolution'. In M. Tallerman & K.R. Gibson (eds.), *The Oxford Handbook of Language Evolution*, ch54. Oxford University Press.

Heine, B. & Kuteva, T. 2007. *The Genesis of Grammar: A Reconstruction*. Oxford University Press.

Heine, B. & Song, K.A. 2010. 'On the genesis of personal pronouns: Some conceptual sources'. *Language and Cognition* 2:1, 117-147.

Heintz, C. & Scott-Phillips, T.C. 2022. 'Expression unleashed: The evolutionary & cognitive foundations of human communication'. *Behavioral and Brain Sciences,* Accepted manuscript, 1-46.

Henderson, I. & Robertson, P. 2007. 'Control and Eradication of the North American Ruddy Duck in Europe'. In G.W. Witmer, W.C. Pitt & K.A. Fagerstone (eds.), *Managing Vertebrate Invasive Species: Proceedings of an International Symposium*, 387-392. USDA/APHIS/WS, National Wildlife Research Center.

Hengeveld, K. & Mackenzie, J.L. 2008. *Functional Discourse Grammar: A Typologically-Based Theory of Language Structure*. Oxford University Press.

Henneberg, M. 1988. 'Decrease of Human Skull Size in the Holocene'. *Human Biology* 60:3, 395-405.

Henrich, J. & Boyd, R. 2008. 'Division of Labor, Economic Specialization, and the Evolution of Social Stratification'. *Current Anthropology* 49:4, 715-724.

Henrich, J. & Gil-White, F.J. 2001. 'The Evolution of Prestige: Freely conferred deference as a mechanism for enhancing the benefits of cultural transmission'. *Evolution and Human Behavior* 22, 165-196.

Henshilwood, C.S., d'Errico, F., Vanhaeren, M., van Niekerk, K. & Jacobs, Z. 2004. 'Middle Stone Age Shell Beads from South Africa'. *Science* 304:5669, 404-404.

Henshilwood, C.S. & Marean, C.W. 2003. 'The Origin of Modern Human Behavior: Critique of the Models and Their Test Implications'. *Current Anthropology* 44:5, 627-651.

Herculano-Houzel, S. 2010. 'Coordinated scaling of cortical and cerebellar numbers of neurons'. *Frontiers in Neuranatomy* 4, 12.

Herman, L.M. & Uyeyama, R.K. 1999. 'The dolphin's grammatical competency: Comments on Kako (1999)'. *Animal Learning & Behavior*, 27:1, 18-23.

Hervey, S.G.J. 1979. *Axiomatic Semantics: A Theory of Linguistic Semantics*. Scottish Academic Press.

Hess, E. 2008. *Nim Chimpsky: The Chimp Who Would Be Human*. Bantam Books.

Hewes, G.W. 1961. 'Food transport and the origin of hominid bipedalism'. *American Anthropologist* 63, 687-710.

Hirsh-Pasek, K. & Golinkoff, R.M. 1996. *The Origins of Grammar: Evidence from Early Language Comprehension*. MIT Press.

Hjelmslev, L. 1963. *Language: An Introduction (Prolegomena to a Theory of Language)*. University of Wisconsin Press.

Hobaiter, C. & Byrne, R.W. 2014. 'The Meanings of Chimpanzee Gestures'. *Current Biology* 24, 1596-1600.

Hobaiter, C., Graham, K.E. & Byrne, R.W. 2022. 'Are ape gestures like words? Outstanding issues in detecting similarities and differences between human language and ape gesture'. *Philosophical Transactions of the Royal Society B* 377, 20210301.

Hoey, M. 2005. *Lexical Priming: A New Theory of Words and Language*. Routledge.
Hofmann, W., Wisneski, D.C., Brandt, M.J. & Skitka, L.J. 2014. 'Morality in everyday life'. *Science* 345:6202, 1340-1343.
Hofstadter, D. 2007. *I Am a Strange Loop*. Basic Books.
Holen, Ø.H. & Johnstone, R.A. 2018. 'Reciprocal mimicry: Kin selection can drive defended prey to resemble their Batesian mimics'. *Proceedings of the Royal Society B* 285, 20181149.
Hölldobler, B. & Wilson, E.O. 2009. *The Super-Organism: The Beauty, Elegance, and Strangeness of Insect Societies*. W.W. Norton & co.
Hood, L. & Rowen, L. 2013. 'The Human Genome Project: Big science transforms biology and medicine'. *Genome Medicine* 5, 79.
Hopkins, W.D. & Pilcher, D.L. 2001. 'Neuroanatomical Localization of the Motor Hand Area With Magnetic Resonance Imaging: The Left Hemisphere Is Larger in Great Apes'. *Behavioral Neuroscience* 115:5, 1159-1164.
Hopkins, W.D., Marino, L., Rilling, J.K. & MacGregor, L.A. 1998. 'Planum temporale asymmetries in great apes as revealed by magnetic resonance imaging (MRI)'. *NeuroReport* 9, 2913-2918.
Hopper, P.J. & Traugott, E.C. 1993. *Grammaticalization*. Cambridge University Press.
Hornstein, N. 2001. *Move! A Minimalist Theory of Construal*. Blackwell Publishers Ltd.
Hornstein, N. 1990. *As Time Goes By: Tense and Universal Grammar*. MIT Press.
Hovers, E. 2012. 'Invention, reinvention and innovation: The makings of Oldowan lithic technology'. In S. Elias (ed.), *Origins of Human Innovation and Creativity*, ch5, 51-68. Elsevier.
Hovers, E. & Belfer-Cohen, A. 2013. 'On Variability and Complexity: Lessons from the Levantine Middle Paleolithic Record'. *Current Anthropology*, 54:S8, S337-S357.
Hrdy, S.B. 2009. *Mothers and Others: The Evolutionary Origins of Mutual Understanding*. Belknap Press.
Hudson, R. 1998. *English Grammar*. Routledge.
Huizinga, J. 1938 [1950]. *Homo ludens: A Study of the Play Element in Culture*. Beacon Press.
von Humboldt, W. 1836 [1999]. *On Language: On the Diversity of Human Language Construction and Its Influence on the Mental Development of the Human Species*. Cambridge University Press.
Hunt, G.R. & Gray, R.D. 2004. 'Direct observations of pandanus-tool manufacture and use by a New Caledonian crow (Corvus moneduloides)'. *Animal Cognition* 7, 114-120.
Hurford, J.R. 2008. 'The evolution of human communication and language'. In P. D'Ettorre & D. Hughes (eds.), *Sociobiology of Communication: An Interdisciplinary Perspective*, 249-264. Oxford University Press.
Hurford, J.R. 2007. *The Origins of Meaning: Language in the Light of Evolution*. Oxford University Press.

Hurford, J.R. 2003. ‘The Neural Basis of Predicate-Argument Structure’. *Behavioural and Brain Sciences* 26:3, 283-284.

Hurford, J.R., Studdert-Kennedy, M. & Knight, C. (eds.) 1998. *Approaches to the Evolution of Language*. Cambridge University Press.

Islam, M.K., Lawag, I.L., Sostaric, T., Ulrich, E., Ulrich, D., Dewar, T., Lim, L.Y. & Locher, C. 2022. ‘Australian Honeypot Ant (Camponotus inflatus) Honey – A Comprehensive Analysis of the Physicochemical Characteristics, Bioactivity, and HPTLC Profile of a Traditional Indigenous Australian Food’. *Molecules*, 27:7, 2154.

Iverson, J.M. & Goldin-Meadow, S. 1998. ‘Why People Gesture when They Speak’. *Nature*, 396, 228.

Jack, R.E., Garrod, O.G.B., Yu, H., Caldara, R. & Schyns, P.G. 2012. ‘Facial expressions of emotion are not culturally universal’. *Proceedings of the National Academy of Science*, 109:19, 7241-7244.

Jackendoff, R. 2002. Foundations of Language: *Brain, Meaning, Grammar, Evolution*. Oxford University Press.

Jackendoff, R. 1972. *Semantic Interpretation in Generative Grammar*. MIT Press.

Jaeggi, A.V., Burkart, J.M. & van Schaik, C.P. 2010. ‘On the psychology of cooperation in humans and other primates: combining the natural history and experimental evidence of prosociality’. *Philosophical Transactions of the Royal Society B*, 365, 2723-2735.

Jakobson, R. 1987. *Language in Literature*. Harvard University Press.

Janmaat, K.R.L., Polansky, L., Ban, S.D. & Boesch, C. 2014. ‘Wild chimpanzees plan their breakfast time, type, and location’. *Proceedings of the National Academy of Science*, 111:46, 16343-16348.

Jaynes, J. 1990. *The Origin of Consciousness in the Breakdown of the Bicameral Mind*. Penguin.

Johanson, D. & Edgar, B. 1996. *From Lucy to Language*. Weidenfeld & Nicolson.

Johansson, S. 2005. *Origins of Language*. John Benjamins Publishing co.

Johnson, M. 1987. *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*. University of Chicago Press.

Johnson, S. 2003. ‘Emotions and the Brain’. *Discover Magazine*, April 2003.

Johnson, M.K., Raye, C.L., Mitchell, K.J. & Ankudowich, E. 2012. ‘The cognitive neuroscience of true and false memories’. *Nebraska Symposium on Motivation* 58, 15-52.

Johnstone, L.T., Karlsson, E.M. & Carey, D.P. 2021. ‘Left-Handers Are Less Lateralized Than Right-Handers for Both Left and Right Hemispheric Functions’. *Cerebral Cortex* 31, 3780-3787.

Jolly, C.J. 2009. ‘Fifty Years of Looking at Human Evolution: Backward, Forward, and Sideways’. *Current Anthropology* 50:2, 187-199.

Jordan, J.J. 2023. ‘A pull versus push framework for reputation’. *Trends in Cognitive Sciences* 1364-6613.

Jordan, J.S. 2003. ‘Emergence of self and other in perception and action: An event-control approach’. *Consciousness and Cognition* 12, 633-646.

Jordan, R. 1998. ‘Is Autism a pathology? Reflections from theory, research and practice’. In P. Shattock & G. Linfoot (eds.), *Psychobiology of Autism: Current Research and Practice*. Sunderland Autism Research Unit, Durham conference papers.

Jordan, K., Weiss, D., Hauser, M. & McMurray, B. 2004. ‘Antiphonal Responses to Loud Contact Calls Produced by Saguinus Oedipus’. *International Journal of Primatology* 25:2, 465-475.

Jung-Beeman, M. 2005. ‘Bilateral brain processes for comprehending natural language’. *Trends in Cognitive Sciences* 9:11, 512-518.

Jürgens, U. 2002. ‘Neural pathways underlying vocal control’. *Neuroscience and Biobehavioral Reviews* 26, 235-258.

Kahneman, D. 2011. *Thinking, Fast and Slow*. Farrar, Straus & Giroux.

Kaminski, J., Call, J. & Fischer, J. 2004. ‘Word Learning in a Domestic Dog: Evidence for “Fast Mapping”‘. *Science,* 304-5677: pp1682-1683.

Kang, S. & Tversky, B. 2016. ‘From hands to minds: Gestures promote understanding’. *Cognitive Research: Principles and Implications* 1, 4.

Karolis, V.R., Corbett, M. & de Schotten, M.T. 2019. ‘The architecture of functional lateralisation and its relationship to callosal connectivity in the human brain’. *Nature Communications* 10, 1417.

Kashdan, T.B. & Silvia, P.J. 2009. ‘Curiosity and Interest: The Benefits of Thriving on Novelty and Challenge’. In S.J. Lopez & C.R. Snyder (eds.), *Oxford Handbook of Positive Psychology*, 367-374. Oxford University Press.

Kathol, A. 2000. *Linear Syntax*. Oxford University Press.

Kaye, K. 1982. *The Mental and Social Life of Babies: How Parents Create Persons*. Methuen & co.

Keizer, E. 2015. *A Functional Discourse Grammar for English*. Oxford University Press.

Kellogg, W.N. & Kellogg, L.A. 1933 [1967]. *The Ape and the Child: A Study of Environmental Influence upon Early Behavior*. Hafner Publishing Company.

Kendon, A. 2009. ‘Kinesic Components of Multimodal Utterances’. *Berkeley Linguistics Society* 35:2, 36-53.

Kendon, A. 1991. ‘Some Considerations for a Theory of Language Origins’. *Man* 26:2, 199-221.

Key, C.A. & Aiello, L.C. 1999. 'The Evolution of Social Organisation'. In R. Dunbar, C. Knight & C. Power (eds.), *The Evolution of Culture*. Edinburgh University Press.

Key, J.M. & Dunmore, C.J. 2018. 'Manual restrictions on Palaeolithic technological behaviours'. *PeerJ* 6, e5399.

Kinsella, A.R. 2009. *Language Evolution and Syntactic Theory*. Cambridge University Press.

Kinzler, K.D., Dupoux, E. & Spelke, E.S. 2007. 'The native language of social cognition'. *Proceedings of the National Academy of Science* 104:30, 12577-12580.

Kirby, S. 1998. 'Language evolution without natural selection: From vocabulary to syntax in a population of learners'. *Edinburgh Occasional Papers in Linguistics* April 3, 1998.

Kivell, T.L., Kibii, J.M., Churchill, S.E., Schmid, P. & Berger, L.R. 2011. 'Australopithecus sediba Hand Demonstrates Mosaic Evolution of Locomotor and Manipulative Abilities'. *Science* 333, 1411-1417.

Klein, R.G. 2009. 'Darwin and the recent African origin of modern humans'. *Proceedings of the National Academy of Science* 106:38, 16007-16009.

Knight, C. 2010. 'The Origins of Symbolic Culture'. In U.J. Frey, C. Störmer & K.P. Willführ (eds.), *Homo Novus – A Human Without Illusions,* 193-211. Springer-Verlag.

Knight, C. 2008. 'Language co-evolved with the rule of law'. *Mind & Society* 7:1, 109-128.

Knight, C. 1999. 'Sex and Language as Pretend Play'. In R. Dunbar, C. Knight & C. Power (eds.), *The Evolution of Culture*. Edinburgh University Press.

Knight, C. 1998. 'Ritual/speech coevolution: a solution to the problem of deception'. In J.R. Hurford, M. Studdert-Kennedy & C. Knight (eds.), *Approaches to the Evolution of Language*. Cambridge University Press.

Knight, C. 1991. *Blood Relations: Menstruation and the Origins of Culture*. Yale University Press.

Knight, C. & Lewis, J. 2017. 'Wild Voices: Mimicry, Reversal, Metaphor, and the Emergence of Language'. *Current Anthropology* 58:4, 435-453.

Knight, C., Power, C. & Watts, I. 1995. The Human Symbolic Revolution: A Darwinian Account. *Cambridge Archaeological Journal* 5:1, 75-114.

Koelsch, S., Rohrmeier, M., Torrecuso, R. & Jentschke, S. 2013. 'Processing of hierarchical syntactic structure in music'. *Proceedings of the National Academy of Science* 110:38, 15443-15448.

Kövecses, Z. 2002. *Metaphor: A Practical Introduction*. Oxford University Press.

Kret, M.E., Prochazkova, E., Sterck, E.H.M. & Clay, Z. 2020. 'Emotional expressions in human and non-human great apes'. *Neuroscience and Biobehavioral Reviews* 115, 378-395.

Kret, M.E., & Tomonaga, M. 2016. 'Getting to the Bottom of Face Processing: Species-Specific Inversion Effects for Faces and Behinds in Humans and Chimpanzees (Pan Troglodytes)'. *PLoS ONE* 11, e0165357.

van Krieken, K. & Sanders, J. 2021. 'Storytelling on Oral Grounds: Viewpoint Alignment and Perspective Taking in Narrative Discourse'. *Frontiers in Psychology* 12, 634930.

Kringelbach, M.L., Stark, E.A., Alexander, C., Bornstein, M.H. & Stein, A. (2016). On cuteness: unlocking the parental brain and beyond. *Trends in Cognitive Science* 20:7, 545-558.

Krupp, D.B. & Maciejewski, W. 2022. 'The evolution of extraordinary self-sacrifice'. *Nature Scientific Reports* 12:1, 90.

Kyle, J.G. & Woll, B. 1985. *Sign Language: The Study of Deaf People and Their Language*. Cambridge University Press.

Lakoff, G. 2014. 'Mapping the brain's metaphor circuitry: metaphorical thought in everyday reason'. *Frontiers in Human Neuroscience* 8, 958.

Lakoff, G. 1993 [2006]. 'Conceptual Metaphor: the contemporary theory of metaphor'. In D. Geeraerts (ed.), *Cognitive Linguistics: Basic Readings*. Mouton de Gruyter.

Lakoff, G. 1987. *Women, Fire and Dangerous Things: What categories reveal about the mind*. University of Chicago Press.

Lakoff, G. 1971. 'On Generative Semantics'. In D.D. Steinberg & L.A. Jacobovits (eds.), *Semantics: An Interdisciplinary Reader in Philosophy, Linguistics and Psychology*. Cambridge University Press.

Lakoff, G. & Johnson, M. 1999. *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. Basic Books.

Lakoff, G. & Johnson, M. 1980. *Metaphors We Live By*. University of Chicago Press.

Laland, K.N., Odling-Smee, J. & Feldman, M.W. 1999. 'Niche Construction, Biological Evolution and Cultural Change'. *Behavioral and Brain Sciences* 23:1, 131-175.

Lamb, S. 2016. 'Linguistic structure: A plausible theory'. *Language Under Discussion* 4:1, 1-37.

Lameira, A.R. & Call, J. 2018. 'Time-space-displaced responses in the orangutan vocal system'. *Science Advances* 4:eaau3401.

Langacker, R.W. 1991. *Foundations of Cognitive Grammar: Volume II, Descriptive Application*. Stanford University Press.

Langacker, R.W. 1987. *Foundations of Cognitive Grammar: Volume I, Theoretical Prerequisites*. Stanford University Press.

Langley, M.C., Benítez-Burraco, A. & Kempe, V. 2020. 'Playing with language, creating complexity: Has play contributed to the evolution of complex language?' *Evolutionary Anthropology* 29:1, 29-40.

Le, Q.V., Le, Q.V., Nishimaru, H., Matsumoto, J., Takamura, Y., Hori, E., Maior, R.S., Tomaz, C., Ono, T. & Nishijo, H. 2020. 'A Prototypical Template for Rapid Face

Detection Is Embedded in the Monkey Superior Colliculus'. *Frontiers in Systems Neuroscience* 14:5.
Leder, D., Hermann, R., Hüls, M., Russo, G., Hoelzmann, P., Nielbock, R., Böhner, U., Lehmann, J., Meier, M., Schwalb, A., Tröller-Reimer, A., Koddenberg, T. & Terberger, T. 2021. 'A 51,000-year-old engraved bone reveals Neanderthals' capacity for symbolic behaviour'. *Nature Ecology & Evolution* 5, 1273-1282.
Lee, D. 1992. *Competing Discourses: Perspective and Ideology in Language*. Longman.
Lee, S.M., Murray, C.M., Lonsdorf, E.V., Fruth, B., Stanton, M.A., Nichols, J. & Hohmann, G. 2020. 'Wild bonobo and chimpanzee females exhibit broadly similar patterns of behavioral maturation but some evidence for divergence'. *American Journal of Physical Anthropology* 171:1, 100-109.
Leerssen, J. 2021. 'Culture, humanities, evolution: the complexity of meaning-making over time'. *Philosophical Transactions of the Royal Society B*, 376, 20200043.
Lehmann, C. 2015. *Thoughts on Grammaticalization (3rd edition)*. Language Science Press.
Lehmann, J., Korstjens, A.H. & Dunbar, R.I.M. 2007. 'Group size, grooming and social cohesion in primates'. *Animal Behaviour* 74, 1617-1629.
Publisher: Routledge Publication Date: 12/03/1998
Lemke, J.L. (1998). 'Multiplying Meaning: Visual and Verbal Semiotics in Scientific Text'. In J.R. Martin & R. Veel (eds.), *Reading Science*. Routledge.
Levinson, S.C. & Holler, J. 2014. 'The origin of human multi-modal communication'. *Philosophical Transactions of the Royal Society B* 369, 20130302.
Lewin, R. 2005. 'Unit 32: The Evolution of Language'. In *Human Evolution: An Illustrated Introduction (5th edition)*. Blackwell Publishing.
Lewis, H.M., Vinicius, L., Strods, J., Mace, R. & Migliano, A.B. 2014. 'High mobility explains demand sharing and enforced cooperation in egalitarian hunter-gatherers'. *Nature Communications* 5, 5789.
Lewis, J. 2009. 'As Well as Words: Congo Pygmy hunting, mimicry and play'. In R. Botha & C. Knight (eds.), *The Cradle of Language*. Oxford University Press.
Lewis, S.L. & Maslin, M.A. 2015. 'Defining the Anthropocene'. *Nature* 519, 171-180.
Lieberman, D.E., Bramble, D.M., Raichlen D.A. & Shea, J.J. 2009. 'Brains, Brawn, and the Evolution of Human Endurance Running Capabilities'. In F.E. Grine, J.G. Fleagle & R.E. Leakey (eds.), *The First Humans – Origin and Early Evolution of the Genus Homo*. Springer.
Liebers, D., de Knijff, P. & Helbig, A.J. 2004. 'The herring gull complex is not a ring species'. *Proceedings of the Royal Society of London B* 271, 893-901.
Ligon, R.A., Diaz, C.D., Morano, J.L., Troscianko, J., Stevens, M., Moskeland, A., Laman, T.G. & Scholes, E. III. 2018. 'Evolution of correlated complexity in the radically different courtship signals of birds-of-paradise'. *PLoS Biology* 16:11, e2006962.
Liti, G., Carter, D.M., Moses, A.M., Warringer, J., Parts, L., James, S.A., Davey, R.P., Roberts, I.N., Burt, A., Koufopanou, V., Tsai, I.J., Bergman, C.M., Bensasson, D.,

O'Kelly, M.J.T., van Oudenaarden, A., Barton, D.B.H., Bailes, E., Nguyen, A.N., Jones, M., Quail, M.A., Goodhead, I., Sims, S., Smith, F., Blomberg, A., Durbin, R. & Louis, E.J. 2009. 'Population genomics of domestic and wild yeasts'. *Nature*, 458:7236, 337-341.

Locke, J. 1689 [1998]. *An Essay Concerning Human Understanding*. Wordsworth Editions Ltd.

Locke, J.L. 2021. 'The Indexical Voice: Communication of Personal States and Traits in Humans and Other Primates'. *Frontiers in Psychology* 12, 651108.

Locke, J.L. 2017. 'Emancipation of the voice: Vocal complexity as a fitness indicator'. *Psychonomic Bulletin & Review* 24:1, 232-237.

Locke, J.L. 1998. 'Social Sound-Making as a Precursor to Spoken Language'. In J.R. Hurford, M. Studdert-Kennedy & C. Knight (eds.), *Approaches to the Evolution of Language*. Cambridge University Press.

Lockwood, D.G. 2002. 'Highlighting in Stratificational-Cognitive Linguistics'. In W. Spruiell, M. Cummings, D. Lockwood & P. Fries (eds.), *Relations and Functions Within and Around Language*. Continuum.

López-Barroso, D., Catanic, M., Ripollés, P., Dell'Acqua, F., Antoni Rodríguez-Fornells, A. & de Diego-Balaguera, R. 2013. 'Word learning is mediated by the left arcuate fasciculus'. *PNAS* 110:32, 13168-13173.

Losin, E.A.R., Russell, J.L., Freeman, H., Meguerditchian, A. & Hopkins, W.D. 2008. 'Left Hemisphere Specialization for Oro-Facial Movements of Learned Vocal Signals by Captive Chimpanzees'. *PLoS ONE*, 3:6, e2529.

Loued-Khenissi, L., Döll, O. & Preuschoff, K. 2019. 'An Overview of Functional Magnetic Resonance Imaging Techniques for Organizational Research'. *Organizational Research Methods*, 22:1, 17-45.

Luppi, A.I., Mediano, P.A.M., Rosas, F.E., Holland, N., Fryer, T.D., O'Brien, J.T., Rowe, J.B., Menon, D.K., Bor, D. & Stamatakis, E.A. 2022. 'A synergistic core for human brain evolution and cognition'. *Nature Neuroscience* 35618951.

Lupyan, G. & Dale, R. 2010. 'Language Structure Is Partly Determined by Social Structure'. *PLoS ONE* 5:1, e8559.

Lupyan, G. & Winter, B. 2018. 'Language is more abstract than you think, or, why aren't languages more iconic?' *Philosophical Transactions of the Royal Society B* 373, 20170137.

Lycett, S.J., Collard, M. & McGrew, W.C. 2007. 'Phylogenetic analyses of behavior support existence of culture among wild chimpanzees'. *Proceedings of the National Academy of Sciences*, 104:45, 17588-17592.

Mackenzie, J.L. 2016. 'Functional linguistics'. In K. Allan (ed.), *The Routledge Handbook of Linguistics*. Routledge.

MacNeilage, P.F. 1998. 'Evolution of the mechanism of language output: comparative neurobiology of vocal and manual communication'. In J.R. Hurford, M. Studdert-

Kennedy & C. Knight (eds.), *Approaches to the Evolution of Language*. Cambridge University Press.

Malik, K. 2000. *Man, Beast and Zombie: what science can and cannot tell us about human nature*. Phoenix.

Malik, K. 1998. 'The Darwinian Fallacy'. *Prospect* December 1998. https://www.prospectmagazine.co.uk/magazine/thedarwinianfallacy (8 June 2022).

Mann, W., Marshall, C.R., Mason, K. & Morgan, G. 2010. 'The Acquisition of Sign Language: The Impact of Phonetic Complexity on Phonology'. *Language Learning and Development* 6, 60-86.

Manning, A.G., Khakimov, R.I., Dall, R.G. & Truscott, A.G. (2015). "Wheeler's delayed-choice gedanken experiment with a single atom". *Nature Physics* 11, 539-543.

Marino, L., Connor, R.C., Fordyce, R.E., Herman, L.M., Hof, P.R., Lefebvre, L., Lusseau, D., McCowan, B., Nimchinsky, E.A., Pack, A.A., Rendell, L., Reidenberg, J.S., Reiss, D., Uhen, M.D., van der Gucht, E. & Whitehead, H. 2007. 'Cetaceans Have Complex Brains for Complex Cognition'. *PLoS Biology*, 5:5, e139, 966-972.

Marks, J. 2002. *What it Means to be 98% Chimpanzee: Apes, People, and Their Genes*. University of California Press.

Marshall, A.J., Wrangham, R.W. & Arcadi, A.C. 1999. 'Does learning affect the structure of vocalizations in chimpanzees?' *Animal Behaviour*, 58:4, 825-830.

Martínez, I.P. 2014. 'Variation, development and pragmatic uses of innit in the language of British adults and teenagers'. *English Language and Linguistics* 19:3, 383-405.

Martins, P.T. & Boeckx, C. 2019. 'Language evolution and complexity considerations: The no half-Merge fallacy'. *PLoS Biology* 17:11, e3000389.

Massey, D.S. 2002. 'A Brief History of Human Society: The Origin and Role of Emotion in Social Life'. *American Sociological Review*, 67-2: 1-29.

Mauss, M. 1950. *The Gift: The Form and Reason for Exchange in Archaic Societies*. Routledge.

McAuliffe, K., Blake, P.R., Steinbeis, N. & Warneken, F. 2017. 'The developmental foundations of human fairness'. *Nature Human Behaviour* 1, 0042.

McBrearty, S. 2007. 'Down with the Revolution'. In P. Mellars, K. Boyle, O. Bar-Yosef & C. Stringer (eds.), *Rethinking the human revolution*. McDonald Institute Monographs.

McGrew, W.C., Marchant, L.F., Scott, S.E. & Tutin, C.E.G. 2001. 'Intergroup Differences in a Social Custom of Wild Chimpanzees: The Grooming Hand-Clasp of the Mahale Mountains'. *Current Anthropology* 42:1, 148-153.

McManus, I.C. 2009. 'The history and geography of human handedness'. In I.E.C. Sommer & R.S. Kahn (eds.), *Language Lateralization and Psychosis*, 37-57. Cambridge University Press.

Meares, R. & Sullivan, G. 2004. 'Two Forms of Human Language'. In G. Williams & A. Lukin (eds.), *The Development of Language: Functional Perspectives on Species and Individuals*. Continuum.

W.H. Mearns, W.H. 1899 [1922]. *Antigonish. In The Conning Tower*, F.P. Adams (ed.). New York World, 27 March 1922.

Meir, N. & Novogrodsky, R. 2019. 'Prerequisites of Third-Person Pronoun Use in Monolingual and Bilingual Children with Autism and Typical Language Development'. *Frontiers in Psychology* 10, 2289.

Melis, A.P., Hare, B. & Tomasello, M. 2006. 'Engineering cooperation in chimpanzees: tolerance constraints on cooperation'. *Animal Behaviour* 72, 275-286.

Méndez, C.A., Celeghin, A., Diano, M., Orsenigo, D., Ocak, B. & Tamietto, M. 2022. 'A deep neural network model of the primate superior colliculus for emotion recognition'. *Philosophical Transactions of the Royal Society B* 377, 20210512.

Mesoudi, A. & Thornton, A. 2018. 'What is cumulative cultural evolution?' *Proceedings of the Royal Society B* 285, 20180712.

Meyer, M., Fu, Q., Aximu-Petri, A., Glocke, I., Nickel, B., Arsuaga, J-L., Martínez, I., Gracia, A., Bermúdez de Castro, J.M., Carbonell, E. & Pääbo, S. 2014. 'A mitochondrial genome sequence of a hominin from Sima de los Huesos'. *Nature* 505, 403-406.

Michaelis, L.A. 2006. 'Time and Tense'. In B. Aarts & A. McMahon (eds.), *The Handbook of English Linguistics*, 220-243. Blackwell.

Miller, G.A. 1956. 'The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information'. *Psychological Review* 63, 81-97.

Miller, M.B. & Gazzaniga, M.S. 1998. 'Creating false memories for visual scenes'. *Neuropsychologia* 36:6, 513-520.

Walter Mischel & Ebbesen, E.B. 1970. Attention in Delay of Gratification. *Journal of Personality and Social Psychology* 16:2, 329-337.

Mitani, J.C. 2009. 'Male chimpanzees form enduring and equitable social bonds'. *Animal Behaviour* 77, 633-640.

Mithen, S. 2005. *The Singing Neanderthals: The Origins of Music, Language, Mind and Body*. Weidenfeld & Nicolson.

Mithen, S. 1996. *The Prehistory of the Mind: A Search for the Origins of Art, Religion and Science*. Phoenix.

Molleman, L., van den Broek, E. & Egas, M. 2013. 'Personal experience and reputation interact in human decisions to help reciprocally'. *Proceedings of the Royal Society B* 280:1757, 20123044.

Moore, M.W. 2010. 'Grammars of Action and Stone Flaking Design Space'. In A. Nowell and I. Davidson (eds.), *Stone Tools and the Evolution of Human Cognition*, 13-43. University Press of Colorado.

Morley, G.D. 2000. *Syntax in Functional Grammar: An Introduction to Lexicogrammar in Systemic Linguistics*. Continuum.

Morrel, J., Singapuri, K., Landa, R.J. & Reetzke, R. 2023. ‘Neural correlates and predictors of speech and language development in infants at elevated likelihood for autism: a systematic review’. *Frontiers in Human Neuroscience* 17:1211676.

Morris, D. 1994. *Body Talk: The Meaning of Human Gestures*. Crown.

Morsella, E., Godwin, C.A., Jantz, T.K., Krieger, S.C. & Gazzaley, A. 2016. ‘Homing in on consciousness in the nervous system: An action-based synthesis’. *Behavioral and Brain Sciences* 39, e168.

Mosch, S.C., Max, J.E. & Tranel, D. 2005. ‘A Matched Lesion Analysis of Childhood versus Adult-Onset Brain Injury Due to Unilateral Stroke: Another Perspective on Neural Plasticity and Recovery of Social Functioning’. *Cognitive and Behavioral Neurology*, 18:1, 5-17.

Mufwene, S.S. 2016. ‘Language Evolution, by Exaptation, with the Mind Leading’. *In New Horizons in Evolutionary Linguistics: Journal of Chinese Linguistics Monograph Series* 27, 155-186.

Mufwene, S.S. 2013. ‘Language as Technology: Some questions that evolutionary linguistics should address’. In T. Lohndal (ed.), *In Search of Universal Grammar: From Norse to Zoque*, 327-358. John Benjamins.

Mufwene, S.S. 2008. ‘What Do Creoles and Pidgins Tell Us about the Evolution of Language?’ In *Language Evolution: Contact, Competition and Change*, ch5. Continuum Press.

Mulder, J.W.F. & Hervey, S.G.J. 1980. *The Strategy of Linguistics: Papers on the Theory and Methodology of Axiomatic Functionalism*. Scottish Academic Press.

Muñoz-Fuentes, V., Vilà, C., Green, A.J., Negro, J.J. & Sorenson, M.D. 2007. ‘Hybridization between white-headed ducks and introduced ruddy ducks in Spain’. *Molecular Ecology* 16:3, 629-638.

Murray, C.M., Mane, S.V. & Pusey, A.E. 2007. ‘Dominance rank influences female space use in wild chimpanzees, Pan troglodytes: Towards an ideal despotic distribution’. *Animal Behaviour* 74:6, 1795-1804.

Murray, C.M., Margaret A. Stanton, M.A., Lonsdorf, E.V., Wroblewski, E.E. & Pusey, A.E. 2016. ‘Chimpanzee fathers bias their behaviour towards their offspring’. *Royal Society Open Science* 3, 160441.

Napoli, D.J. & Sutton-Spence, R. 2014. ‘Order of the major constituents in sign languages: implications for all language’. *Frontiers in Psychology* 5, 376.

Nettle, D. 1999. *Linguistic Diversity*. Oxford University Press.

Nettle, D. & Pollet, T.V. 2008. ‘Natural Selection on Male Wealth in Humans’. *The American Naturalist*, 172:5, 658-666.

Newmeyer, F.J. 2000. *Language Form and Language Function*. MIT Press.

Newton, M. 2002. *Savage Girls and Wild Boys: A History of Feral Children*. Faber & Faber.

Newton-Fisher, N.E. 2014. 'The Hunting Behavior and Carnivory of Wild Chimpanzees'. In W. Henke & I. Tattersall (eds.), *Handbook of Paleoanthropology*, 1661-1691. Springer.

Ni, X., Ji, Q., Wu, W., Shao, Q., Ji, Y., Zhang, C., Liang, L.,Ge, J., Guo, Z., Li, J., Li, Q., Grün, R. & Stringer, C. 2021. 'Massive cranium from Harbin in northeastern China establishes a new Middle Pleistocene human lineage'. *The Innovation* 100130.

Nichols, S. & Stich, S. 2003. 'How to Read Your Own Mind: A Cognitive Theory of Self-Consciousness'. In Q. Smith and A. Jokic (eds.), *Consciousness: New Essays*. Oxford University Press.

Niekus, M.J.L.Th., Kozowyk, P.R.B., Langejans, G.H.J., Ngan-Tillarde, D., van Keulen, H., van der Plicht, J., Cohen, K.M., van Wingerden, W., van Os, B., Bjørn I. Smit, B.I., Amkreutz, L.W.S.W., Johansen, L., Verbaas, A. & Dusseldorp, G.L. 2019. 'Middle Paleolithic complex technology and a Neandertal tar-backed tool from the Dutch North Sea'. *PNAS* 116:44, 22081-22087.

Nietzsche, F. 1874 [1976]. 'On Truth and Lie in an Extra-Moral Sense'. In *The Portable Nietzsche*. Viking Press.

Nieuwburg, E.G.I., Ploeger, A. & Kret, M.E. 2021. 'Emotion recognition in nonhuman primates: How experimental research can contribute to a better understanding of underlying mechanisms'. *Neuroscience and Biobehavioral Reviews* 123, 24-47.

Niles, J.D. 1999. Homo Narrans: *The Poetics and Anthropology of Oral Literature*. University of Pennsylvania Press.

Nishimura, T., Mikami, A., Suzuki, J. & Matsuzawa, T. 2006. 'Descent of the hyoid in chimpanzees: evolution of face flattening and speech'. *Journal of Human Evolution* 51, 244-254.

Noonan, M. 1999. 'Non-structuralist syntax'. In M. Darnell, E.A. Moravcsik, M. Noonan, F.J. Newmeyer & K.M. Wheatley (eds.), *Functionalism and Formalism in Linguistics, Volume I: General Papers*. John Benjamins.

Noss, A.J. & Hewlett, B.S. 2001. 'The Contexts of Female Hunting in Central Africa'. *American Anthropologist* 103:2, 1024-1040.

Nsubuga, A.M., Robbins, M.M., Boesch, C. & Vigilant, L. 2008. 'Patterns of Paternity and Group Fission in Wild Multimale Mountain Gorilla Groups'. *American Journal of Physical Anthropology* 135, 263-274.

Núñez, R.E. & Sweetser, E. 2006. 'With the Future Behind Them: Convergent Evidence From Aymara Language and Gesture in the Crosslinguistic Comparison of Spatial Construals of Time'. *Cognitive Science* 30, 1-49.

Odling-Smee, J. & Laland, K.N. 2009. 'Cultural Niche Construction: evolution's cradle of language'. In R. Botha & C. Knight (eds.), *The Prehistory of Language*. Oxford University Press.

Olulade, O.A., Seydell-Greenwald, A., Chambers, C.E., Turkeltaub, P.E., Dromerick, A.W., Berl, M.M., Gaillard, W.D. & Newport, E.L. 2020. 'The neural basis of language development: Changes in lateralization over age'. *Proceedings of the National Academy of Science* 117:38, 23477-23483.

Opie, C. 2004. *Testing the Grandmothering Hypothesis: The Provisioning of Homo Erectus Infants and Juveniles*. Dissertation submitted in partial fulfilment of the requirements for the degree of M.Sc. in Human Evolution and Behaviour (UCL) of the University of London.

Osuna-Mascaró, A.J., O'Hara, M., Folkertsma, R., Tebbich, S., Beck, S.R. & Auersperg, A.M.I. 2023. 'Flexible tool set transport in Goffin's cockatoos'. *Current Biology* 33, 1-9.

Packer, C. & Pusey, A.E. 1983. 'Adaptations of Female Lions to Infanticide by Incoming Males'. *American Naturalist* 121:5, 716-728.

Pain, R. 2022. 'Stone tools, predictive processing and the evolution of language'. *Mind & Language* DOI: 10.1111/mila.12419.

Palmer, G. B. 2006. 'Energy Through Fusion at Last: Synergies in Cognitive Anthropology and Cognitive Linguistics'. In G. Kristiansen, M. Achard, R. Dirven, & F.J. Ruiz de Mendoza Ibáñez (eds.), *Cognitive Linguistics: Current Applications and Future Perspectives*. Mouton de Gruyter.

Papadatou-Pastou, M., Ntolka, E., Schmitz, J., Martin, M., Munafò, M.R., Ocklenburg, S. & Paracchini, S. 2020. 'Human handedness: A meta-analysis'. *Psychological Bulletin* 146:6, 481-524.

Paradis, J. & Crago, M. 2000. 'Tense and temporality: A comparison between children learning a second language and children with SLI'. *Journal of Speech, Language and Hearing Research*, 43:4, 834-848.

Parish, A.R. 1996. 'Female Relationships in Bonobos (Pan paniscus): Evidence for Bonding, Cooperation, and Female Dominance in a Male-Philopatric Species'. *Human Nature* 7:1, 61-96.

Parisi, D. & Schlesinger, M. 2002. 'Artificial Life and Piaget'. *Cognitive Development* 17, 1301-1321.

Parker, A.R. 2006. *Evolution as a Constraint on Theories of Syntax: The Case against Minimalism*. PhD thesis, Edinburgh University, March 2006.

Parr, L.A. 2011. 'The evolution of face processing in primates'. *Philosophical Transactions of the Royal Society B* 366, 1764-1777.

Parr, L.A. & Waller, B.M. 2006. 'Understanding chimpanzee facial expression: insights into the evolution of communication'. *SCAN* 1, 221-228.

Parr, L.A., Waller, B.M. & Fugate, J. 2005. 'Emotional communication in primates: implications for neurobiology'. *Current Opinion in Neurobiology* 15, 716-720.

Parten, M.B. 1932. 'Social participation among preschool children'. *Journal of Abnormal and Social Psychology* 27, 243-269.

Patterson, F.G.P. & Cohn, R.H. 1990. ‘Language acquisition by a lowland gorilla: Koko’s first ten years of vocabulary development’. *Word* 41:2, 97-143

Penn, D.C., Holyoak, K.J. & Povinelli, D.J. 2008. ‘Darwin’s mistake: Explaining the discontinuity between human and nonhuman minds’. *Behavioral and Brain Sciences* 31, 109-178.

Pepperberg, I.M. 2021. ‘Nonhuman and Nonhuman-Human Communication: Some Issues and Questions’. *Frontiers in Psychology* 12, 647841.

Pepperberg, I.M. 2005. ‘An avian parallel to primate mirror neurons and language evolution? Commentary to Michael A. Arbib, From Monkey-like Action Recognition to Human Language: An evolutionary framework for neurolinguistics’. *Behavioral and Brain Sciences* 28, 105-167.

Pepperberg, I.M. 1999. *The Alex Studies: Cognitive and Communicative Abilities of Grey Parrots*. Harvard University Press.

Perfors, A., Tenenbaum, J.B. & Regier, T. 2011. ‘The learnability of abstract syntactic principles’. *Cognition* 118, 306-338.

Persaud, R. 2005. ‘Deception special: The truth about lies’. *New Scientist,* 30 July 2005.

Peters, A.M. 1995. ‘Strategies in the Acquisition of Syntax’. In P. Fletcher & B. MacWhinney (eds.), *The Handbook of Child Language*. Blackwell Publishers.

Petraglia, M.D. & Groucutt, H.S. 2017. ‘Out of Africa: The Evolution and History of Human Populations in the Southern Dispersal Zone’. In J.H. Schwartz (ed.), *Rethinking Human Evolution*, 129-138. MIT Press.

von Petzinger, G. & Nowell, A. 2011. ‘A question of style: reconsidering the stylistic approach to dating Palaeolithic parietal art in France’. *Antiquity* 85, 1165-1183.

Piaget, J. 1947 [1950]. *The Psychology of Intelligence*. Routledge.

Piaget, J. 1923 [1959]. *The Language and Thought of the Child*. Routledge.

Pilley, J.W. & Reid, A.K. 2011. ‘Border collie comprehends object names as verbal referents’. *Behavioural Processes* 86, 184-195.

Pinker, S. 2002. *The Blank Slate: The Modern Denial of Human Nature*. Penguin.

Pinker, S. 1997. *How the Mind Works*. Penguin.

Pinker, S. 1994. *The Language Instinct*. Penguin.

Pinker, S. & Bloom, P. 1990. ‘Natural language and natural selection’. *Behavioral and Brain Sciences* 13:4, 707-784.

Pinker, S. & Jackendoff, R. 2005. ‘The faculty of language: what’s special about it?’ *Cognition* 95, 201-236.

Pitchford, I. 2001. ‘The Origins of Violence: Is Psychopathy an Adaptation?’ *The Human Nature Review* 1, 28-36.

Plotnik, J.M., de Waal, F.B.M. & Reiss, D. 2006. ‘Self-recognition in an Asian elephant’. *Proceedings of the National Academy of Science* 103:45, 17053-17057.

Plunkett, K. 1995. ‘Connectionist Approaches to Language Acquisition’. In P. Fletcher & B. MacWhinney (eds.), *The Handbook of Child Language*. Blackwell Publishers.

Pollick, A.S. & de Waal, F.B.M. 2007. ‘Ape gestures and language evolution’. *Proceedings of the National Academy of Science* 104:19, 8184-8189.

Port, R.F. & Leary, A.P. 2005. ‘Against Formal Phonology’. *Language* 81:4, 927-964.

Postal, P.M. 1968. ‘Underlying and Superficial Linguistic Structure’. In R.C. Oldfield & J.C. Marshall (eds.), *Language*. Penguin.

Povinelli, D.J. 2000. *Folk Physics for Apes: The Chimpanzee’s Theory of How the World Works*. Oxford University Press.

Powers, S.T., van Schaik, C.P. & Lehmann, L. 2021. ‘Cooperation in large-scale human societies – What, if anything, makes it unique, and how did it evolve?’ *Evolutionary Anthropology* 30:4, 280-293.

Premack, D. & Hauser, M.D. 2001. ‘A whale of a tale: Calling it culture doesn’t help’. *Behavioral and Brain Sciences* 24:2, 350-351.

Premack, D. & Premack, A.J. 1983. *The Mind of an Ape*. WW Norton & co.

Premack, D. & Woodruff, G. 1978. ‘Does the chimpanzee have a theory of mind?’ *Behavioral and Brain Sciences* 49, 515-526.

Premo, L.S. & Kuhn, S.L. 2010. ‘Modeling Effects of Local Extinctions on Culture Change and Diversity in the Paleolithic’. *PLoS ONE* 5:12, e15582.

Pressman, S.D., Cohen, S., Miller, G.E., Barkin, A., Rabin, B.S. & Treanor, J.J. 2005. ‘Loneliness, Social Network Size, and Immune Response to Influenza Vaccination in College Freshmen’. *Health Psychology* 24:3, 297-306.

Prior, H., Schwarz, A. & Güntürkün, O. 2008. ‘Mirror-Induced Behavior in the Magpie (Pica pica): Evidence of Self-Recognition’. *PLoS Biology* 6:8, 1642-1650.

Progovac, L. 2019. ‘Minimalism in the Light of Biology: What to Retain and What to Discard?’ *Frontiers in Psychology* 10, 1303.

Pruetz, J.D. & Bertolani, P. 2007. ‘Savanna Chimpanzees, Pan troglodytes verus, Hunt with Tools’. *Current Biology* 17, 412-417.

Prüfer, K., Posth, C., Yu, H., Stoessel, A., Spyrou, M.A., Deviese, T., Mattonai, M., Ribechini, E., Higham, T., Velemínský, P., Brůžek, J. & Krause, J. 2021. ‘A genome sequence from a modern human skull over 45,000 years old from Zlatý kůň in Czechia’. *Nature Ecology & Evolution* 5, 820-825.

Pullum, G.K. & Scholz, B.C. 2010. ‘Recursion and the Infinitude Claim’. In H. Van der Hulst (ed.), *Recursion and Human Language*, 113-137. Mouton de Gruyter.

Pusey, A.E. and Schroepfer-Walker, K. 2013. ‘Female competition in chimpanzees’. *Philosophical Transactions of the Royal Society B* 368, 20130077.

Qi, T., Schaadt, G. & Friederici, A.D. 2021. Associated functional network development and language abilities in children. *NeuroImage* 242, 118452.

Quam, R., Martínez, I., Lorenzo, C., Bonmatí, A., Rosa, M., Jarabo, P. & Arsuaga, J.L. 2012. ‘Studying Audition in Fossil Hominins: A New Approach to the Evolution of Language?’ In M.K. Jackson (ed.), *Psychology of Language: Psychology of Emotions, Motivations and Actions*, 47-95. Nova Science Publishers.

Queller, D.C. 1994. ‘Extended parental care and the origin of eusociality’. *Proceedings of the Royal Society of London B* 256, 105-111.

Radford, A. 2009. *Analysing English Sentences: A Minimalist Approach*. Cambridge University Press.

Raichle, M.E. & Gusnard, D.A. 2002. ‘Appraising the brain’s energy budget’. *Proceedings of the National Academy of Science* 99:16, 10237-10239.

Ramat, A.G. 1999. ‘Functional typology and strategies of clause connection in Second Language Acquisition’. *Linguistics* 37:3, 519-548.

Read, D.W., Manrique, H.M. & Walker, M.J. 2022. ‘On the working memory of humans and great apes: Strikingly similar or remarkably different?’ *Neuroscience and Biobehavioral Reviews* 134, 104496.

Reboul, A.C. 2015. ‘Why language really is not a communication system: a cognitive view of language evolution’. *Frontiers in Psychology* 6, 1434.

Reichenbach, H. 1927 [1956]. *The Philosophy of Space & Time*. Dover Publications inc.

Reiss, D. & Marino, L. 2001. ‘Mirror self-recognition in the bottlenose dolphin: A case of cognitive convergence’. *Proceedings of the National Academy of Science* 98:10, 5937-5942.

Rendall, D. 2021. ‘Aping Language: Historical Perspectives on the Quest for Semantics, Syntax, and Other Rarefied Properties of Human Language in the Communication of Primates and Other Animals’. *Frontiers in Psychology* 12, 675172.

Rendell, L. & Whitehead, H. 2001. ‘Culture in whales and dolphins’. *Behavioral and Brain Sciences* 24:2, 309-382.

Richmond, B.G. & Jungers, W.L. 2008. ‘Orrorin tugenensis Femoral Morphology and the Evolution of Hominin Bipedalism’. *Science* 319, 1662-1665.

Riehl, C. & Frederickson, M.E. 2016. ‘Cheating and punishment in cooperative animal societies’. *Philosophical Transactions of the Royal Society B* 371:1687, 20150090.

Rilling, J.K., Glasser, M.F., Preuss, T.M., Ma, X., Zhao, T., Hu, X. & Behrens, T.E.J. 2008. ‘The evolution of the arcuate fasciculus revealed with comparative DTI’. *Nature Neuroscience* 11:4, 426-428.

Roberts, P. 2015. Patrick Roberts (2016). ‘We have never been behaviourally modern: The implications of Material Engagement Theory and Metaplasticity for understanding the Late Pleistocene record of human behaviour’. *Quaternary International* 405:A, 8-20.

Roberts, S.G.B. & Roberts, A.I. 2018. ‘Social Brain Hypothesis: Vocal and Gesture Networks of Wild Chimpanzees’. *Frontiers in Psychology* 7, 1756.

Roebroeks, W., MacDonald, K., Scherjon, F., Bakels, C., Kindler, L., Nikulina, A., Pop, E. & Gaudzinski-Windheuser, S. 2021. ‘Landscape modification by Last Interglacial Neanderthals’. *Science Advances* 7, eabj5567.

Roffman, I., Peleg, G., Stadler, A. & Nevo, E. 2015. ‘Meaningful Informational Exchange and Pantomime in Chimpanzees and Bonobos: Implications for Proto-Language in Hominins’. *Human Evolution* 30:3-4, 141-174.

Rogers, L.J. 2021. ‘Lateralization and Cognitive Capacity’. *Animals* 11, 1996.

Roksandic, M., Radović, P., Wu, X-J. & Bae, C.J. 2021. ‘Resolving the “muddle in the middle”: The case for Homo bodoensis sp. nov’. *Evolutionary Anthropology* 2021, 1-10.

Rolian, C. & Carvalho, S. 2018. ‘Tool Use and Manufacture in the Last Common Ancestor of Pan and Homo’. In M.N. Muller (ed.), *Chimpanzees and Human Evolution,* 602-644. Harvard University Press.

Rose, D. 2006. ‘A Systemic Functional Approach to Language Evolution’. *Cambridge Archaeological Journal* 16:1, 73-96.

Rosenbaum, S., Hirwa, J.P., Silk, J.B., Vigilant, L. & Stoinski, T.S. 2015. ‘Male rank, not paternity, predicts male-immature relationships in mountain gorillas, Gorilla beringei beringei’. *Animal Behaviour* 104, 13-24.

Roth, G. & Dicke, U. 2005. ‘Evolution of the brain and intelligence’. *Trends in Cognitive Sciences* 9:5, 250-257.

Sabbagh, L. 2006. ‘The Teen Brain, Hard at Work – No, Really’. *Scientific American Mind* 17:4, 20-25.

Sag, I.A., Chaves, R.P., Abeillé, A., Estigarribia, B., Flickinger, D., Kay, P., Michaelis, L.A., Müller, S., Pullum, G.K., van Eynde, F. & Wasow, T. 2020. ‘Lessons from the English auxiliary system’. *Journal of Linguistics* 56, 87-155.

Sampson, G. 2005. *The ‘Language Instinct’ Debate (revised edition)*. Continuum.

Sampson, G. 1997. *Educating Eve: The ‘Language Instinct’ Debate*. Cassell.

Samuels, R. 2006. ‘Is the Human Mind Massively Modular?’ In R.J. Stainton (ed.), *Contemporary Debates in Cognitive Science*, 37-56. Blackwell Publishing.

Sandler, W. 2022. ‘Redefining Multimodality’. *Frontiers in Communication* 6, 758993.

Sankararaman, S., Mallick, S., Patterson, N. & Reich, D. 2016. ‘The combined landscape of Denisovan and Neanderthal ancestry in present-day humans’. *Current Biology* 26:9, 1241-1247.

Sanz, C., Morgan, D. & Gulick, S. 2004. ‘New Insights into Chimpanzees, Tools, and Termites from the Congo Basin’. *American Naturalist* 164, 567-581.

Sapir, E. 1921. *Language: An Introduction to the Study of Speech*. Harcourt Brace & co.

Satık, D. 2022. ‘The Strong Minimalist Thesis Is too Strong: Syntax Is More Than Just Merge’. *Biolinguistics* 16, e9861.

de Saussure, F. 1972. *Course in General Linguistics*. Duckworth.

Sauter, D.A., Eisner, F., Ekman, P. & Scott, S.K. 2010. ‘Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations’. *Proceedings of the National Academy of Science* 107:6, 2408-2412.

Savage-Rumbaugh, E.S. 1999. 'Ape Language: between a rock and a hard place'. In B. King (ed.), *The Origins of Language: What Non-human Primates Can Tell Us*. SAR Press.

Savage-Rumbaugh, E.S, Fields, W., Segerdahl, P., & Rumbaugh, D. 2005. 'Culture Prefigures Cognition in Pan/Homo Bonobos'. *Theoria: An International Journal for Theory, History and Foundations of Science* 54, 311-328.

Savage-Rumbaugh, E.S. & Lewin, R. 1994. *Kanzi: The Ape at the Brink of the Human Mind*. John Wiley & Sons Inc.

Savage-Rumbaugh, E.S., Murphy, J., Sevcik, R.A., Brakke, K.E., Williams, S.L., Rumbaugh, D.M. & Bates, E. 1993. 'Language Comprehension in Ape and Child'. *Monographs of the Society for Research in Child Development* 58:3/4.

Schacter, D.L., Addis, D.R. & Buckner, R.L. 2007. 'Remembering the past to imagine the future: the prospective brain'. *Nature Reviews: Neuroscience* 8, 657-661.

Schmitt, D. 2003. 'Insights into the evolution of human bipedalism from experimental studies of humans and other primates'. *The Journal of Experimental Biology* 206, 1437-1448.

Schoenemann, P.T. 2022. 'Evidence of Grammatical Knowledge in Apes: An Analysis of Kanzi's Performance on Reversible Sentences'. *Frontiers in Psychology* 13, 885605.

Schoenemann, P. T., Sheehan, M. J. & Glotzer, L. D. 2005. 'Prefrontal White Matter Volume Is Disproportionately Larger in Humans than in Other Primates'. Nature Neuroscience, 8:2, 242-252.

Schulkin, J. & Raglan, G.B. 2014. The evolution of music and human social capability. *Frontiers in Neuroscience* 8, 292.

Schwartz, B. 2004. 'The Tyranny of Choice'. *Scientific American* 290:4, 42-47.

Scott-Phillips, T.C. 2010. 'The Evolution of Communication: Humans may be exceptional'. In B. Galantucci & S. Garrod (eds.), *Experimental Semiotics: A New Approach for Studying the Emergence and the Evolution of Human Communication*, 78-99. John Benjamins Publishing Company.

Scott-Phillips, T.C. 2008. 'On the Correct Application of Animal Signalling Theory to Human Communication'. In A. D. M. Smith, K. Smith & R. Ferrer-i-Cancho (eds.), *Proceedings of the 7th International Conference on the Evolution of Language*. Singapore: World Scientific.

Sear, R. & Mace, R. 2008. 'Who keeps children alive? A review of the effects of kin on child survival'. *Evolution and Human Behavior*, 29:1, 1-18.

Searcy, W.A & Nowicki, S. 2008. 'Bird Song and the Problem of Honest Communication: How is the honesty of animal signals maintained when exaggeration and bluff are so tempting?' *American Scientist*, 96, 114-121.

Searle, J.R. 1999. *Mind, Language and Society: Philosophy in the Real World*. Weidenfeld & Nicolson.

Searle, J. R. 1993. 'Metaphor'. In A. Ortony (ed.), *Metaphor and Thought (2nd edition)*. Cambridge University Press.

Segerdahl, P., Fields, W. & Savage-Rumbaugh, E.S. 2005. *Kanzi's Primal Language: The Cultural Initiation of Primates into Language*. Palgrave Macmillan.

Seidl, A. & Cristia, A. 2012. 'Infants' learning of phonological status'. *Frontiers in Psychology* 3:448.

Seyfarth, R.M. & Cheney, D.L. 2013. 'Affiliation, empathy, and the origins of theory of mind'. *Proceedings of the National Academy of Science* 110:Suppl2, 10349-10356.

Shea, J.J. 2006. 'Child's Play: Reflections on the Invisibility of Children in the Paleolithic Record'. *Evolutionary Anthropology* 15, 212-216.

Shimelmitz, R., Weinstein-Evron, M., Ronen, A. & Kuhn, S.L. 2016. 'The Lower to Middle Paleolithic transition and the diversification of Levallois technology in the Southern Levant: Evidence from Tabun Cave, Israel'. *Quaternary International* 409:B, 23-40.

Shipton, C. 2019. 'Three Stages in the Evolution of Human Cognition: Normativity, recursion, and abstraction'. In T.B. Henley, M.J. Rossano & E.P. Kardas (eds.), *Handbook of Cognitive Archaeology: Psychology in Prehistory*, 153-173. Routledge.

Shipton, C., Clarkson, C. & Cobden, R. 2019. 'Were Acheulean Bifaces Deliberately Made Symmetrical? Archaeological and Experimental Evidence'. *Cambridge Archaeological Journal* 29:1, 65-79.

Shipton, C. Clarkson, C., Pal, J.N., Jones, S.C., Roberts, R.G., Harris, C., Gupta, M.C., Ditchfield, P.W. & Petraglia, M.D. 2013. 'Generativity, hierarchical action and recursion in the technology of the Acheulean to Middle Palaeolithic transition: A perspective from Patpara, the Son Valley, India'. *Journal of Human Evolution* 65:2, 93-108.

Shirai, Y. & Miyata, S. 2006. 'Does past tense marking indicate the acquisition of the concept of temporal displacement in children's cognitive development?' *First Language* 26:1, 45-66.

Shoshani, J., Kupsky, W.J. & Marchant, G.H. 2006. 'Elephant Brain Part I: Gross morphology, functions, comparative anatomy, and evolution'. *Brain Research Bulletin* 70, 124-157.

Singer, T., Seymour, B., O'Doherty, J.P., Stephan, K.E., Dolan, R.J. & Frith, C.D. 2006. 'Empathic neural responses are modulated by the perceived fairness of others'. *Nature* 439, 466-469.

Slocombe, K.E. & Zuberbühler, K. 2007. 'Chimpanzees modify recruitment screams as a function of audience composition'. *Proceedings of the National Academy of Science* 104:43, 17228-17233.

Sloman, A. 2008.'Putting the Pieces Together Again'. In R. Sun (ed.), *The Cambridge Handbook of Computational Psychology*. Cambridge University Press.

Smith, D., Schlaepfer, P., Major, K., Dyble, M., Page, A.E., Thompson, J., Chaudhary, N., Salali, G.D., Mace, R., Astete, L., Ngales, M., Vinicius, L. & Migliano, A.B.

2017. 'Cooperation and the evolution of hunter-gatherer storytelling'. *Nature Communications* 8, 1853.

Smith, E.A. & Bliege Bird, R.L. 2000. 'Turtle hunting and tombstone opening: Public generosity as costly signaling'. *Evolution and Human Behavior* 21, 245-261.

Smith, P.K., Cowie, H. & Blades, M. 2003. *Understanding Children's Development (4th edition)*. Blackwell Publishing.

Sober, E. & Wilson, D.S. 1999. *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Harvard University Press.

Solms, M. 2017. 'What is "the unconscious," and where is it located in the brain? A neuropsychoanalytic perspective'. *Annals of the New York Academy of Sciences* 1406, 90-97.

Spencer, H. 1898. *The Principles of Biology (volume 1)*. D. Appleton & Co.

Spocter, M.A., Hopkins, W.D., Garrison, A.R., Bauernfeind, A.L., Stimpson, C.D., Hof, P.R. & Sherwood, C.C. 2010. 'Wernicke's area homologue in chimpanzees (Pan troglodytes) and its relation to the appearance of modern human language'. *Proceedings of the Royal Society B* 277, 2165-2174.

Sporns, O. 2022. 'The complex brain: connectivity, dynamics, information'. *Trends in Cognitive Sciences* 26:12, 1066-1067.

St. Clair, R.N. 1982. 'Language and the Social Construction of Reality'. *Language Sciences* 4:2, 221-236.

Steels, L. 2003. 'Social Language Learning'. In M. Tokoro & L. Steels (eds.), *The Future of Learning*. IOS Press.

Steels, L. 1998. 'Synthesising the origins of language and meaning using co-evolution, self-organisation and level formation'. In J. Hurford, C. Knight & M. Studdert-Kennedy (eds.), *Approaches to the Evolution of Language: Social and Cognitive Bases*. Cambridge University Press.

Steklis, H.D. & Harnad, S.R. 1976. 'From hand to mouth: Some critical stages in the evolution of language'. In S. Harnad, H.D. Steklis & J. Lancaster (eds.), *Origins and Evolution of Language and Speech*. New York Academy of Sciences.

Sterelny, K. 2016. 'Cumulative Cultural Evolution and the Origins of Language'. *Biological Theory* 11, 173-186.

Sterelny, K. 2007. 'Social intelligence, human intelligence and niche construction'. *Philosophical Transactions of the Royal Society B* 362, 719-730.

Stern, W. 1914. *The Psychological Methods of Testing Intelligence, Volume 1*. Warwick & York.

Stewart, I. & Cohen, J. 1997. *Figments of Reality: The Evolution of the Curious Mind*. Cambridge University Press.

Stout, D. 2011. 'Stone toolmaking and the evolution of human culture and cognition'. *Philosophical Transactions of the Royal Society B* 366, 1050-1059.

Stout, D. & Chaminade, T. 2009. 'Making Tools and Making Sense: Complex, Intentional Behaviour in Human Evolution'. *Cambridge Archaeological Journal* 19:1, 85-96.

Stout, D., Chaminade, T., Apel, J., Shafti, A. & Faisal, A.A. 2021. 'The measurement, evolution, and neural representation of action grammars of human behavior'. *Nature Scientific Reports* 11:1, 13720.
Stringer, C. & Andrews, P. 2005. *The Complete World of Human Evolution*. Thames & Hudson.
Stringer, C. & Gamble, C. 1993. *In Search of the Neanderthals*. Thames & Hudson.
Subiaul, F., Vonk, J., Okamoto-Barth, S. & Barth, J. 2008. 'Do chimpanzees learn reputation by observation? Evidence from direct and indirect experience with generous and selfish strangers'. *Animal Cognition* 11:4, 611-623.
Suchak, M., Eppley, T.M., Campbell, M.W., Feldman, R.A., Quarles, L.F. & de Waal, F.B.M. 2016. 'How chimpanzees cooperate in a competitive world'. *Proceedings of the National Academy of Science* 113:36, 10215-10220.
Sudd, J.H. & Franks, N.R. 1987. *The Behavioural Ecology of Ants*. Blackie & Son Ltd.
Szaflarski, J.P., Holland, S.K., Schmithorst, V.J. & Byars, A.W. 2006. 'An fMRI study of language lateralization in children and adults'. *Human Brain Mapping* 27:3, 202-212.
Számadó, S., Samu, F. & Takács, K. 2022. 'Condition-dependent trade-offs maintain honest signalling'. *Royal Society Open Science* 9, 220335.
Szathmáry, E. 2006. 'Darwin for All Seasons'. *Science* 313, 306-307.
Szathmáry, E. & Számadó, S. 2008. 'Language: a social history of words'. *Nature* 456, 40-41.

Tan, J. & Hare, B. 2013. 'Bonobos Share with Strangers'. *PLOS ONE* 8:1, e51922.
Tannen, D. 1994. *Gender and Discourse*. Oxford University Press.
Tanner, J.E., Patterson, F.G. & Byrne, R.W. 2006. 'The development of spontaneous gestures in zoo-living gorillas and sign-taught gorillas: from action and location to object representation'. *Journal of Developmental Processes* 1, 69-102.
Tarski, A. 1936. 'The concept of truth in formalized languages'. In A. Tarski (ed.), *Logic, Semantics, Metamathematics*, 152-278. Oxford University Press.
Taylor, P.D. & Day, T. 2004. 'Cooperate with thy Neighbour?' *Nature* 428, 611-612.
Terrace, H.S., Petitto, L.A., Sanders, R.J. & Bever, T.G. 1979. 'Can an Ape Create a Sentence?' *Science* 206, 891-902.
Theissen, G. 2006. 'The proper place of hopeful monsters in evolutionary biology'. *Theory in Biosciences* 124, 349-369.
Thomas, J. 1995. *Meaning in Interaction: An Introduction to Pragmatics*. Longman Ltd.
Thompson, G. 1996. *Introducing Functional Grammar*. Arnold.
Thorndike, E. 1927. *The Measurement of Intelligence*. Columbia University.
Thulborn, K.R., Carpenter, P.A. & Just, M.A. 1999. 'Plasticity of Language-Related Brain Function During Recovery from Stroke'. *Stroke* 30, 749-754.

Tibbetts, P.E. 2014. ‘Where does cognition occur: In one’s head or in one’s Embodied/extended environment?’ *The Quarterly Review of Biology* 89:4, 359-368.

Tibbetts, E.A. & Dale, J. 2004. ‘A socially enforced signal of quality in a paper wasp’. *Nature* 432, 218-222.

Tibbetts, E.A., Sheehan, M.J. & Dale, J. 2008. ‘A testable definition of individual recognition’. *Trends in Ecology and Evolution* 23:7, 356.

Tokuyama, N., Sakamaki, T. & Furuichi, T. 2019. ‘Inter-group aggressive interaction patterns indicate male mate defense and female cooperation across bonobo groups at Wamba, Democratic Republic of the Congo’. *American Journal of Physical Anthropology* 170:4, 535-550.

Tokuyama, N., Toda, K., Poiret, M-L., Iyokango, B., Bakaa, B. & Ishizuka, S. 2021. ‘Two wild female bonobos adopted infants from a different social group at Wamba’. *Nature Scientific Reports* 11, 4967.

Tomalin, M. 2004. ‘Leonard Bloomfield: linguistics and mathematics’. *Historiographia Linguistica xxxi* 1, 105-136.

Tomasello, M. 2019. *Becoming Human: A Theory of Ontogeny*. Belknap Press.

Tomasello, M. 2018. ‘The Origins of Morality: How we learned to put our fate in one another’s hands’. *Scientific American* 319:3, 70-75.

Tomasello, M. 2008. *Origins of Human Communication*. MIT Press.

Tomasello, M. 2003a. *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Harvard University Press.

Tomasello, M. 2003b. ‘On the Different Origins of Symbols and Grammar’. In M.H. Christiansen & S. Kirby (eds.), *Language Evolution*. Oxford University Press.

Tomasello, M. 1999. *The Cultural Origins of Human Cognition*. Harvard University Press.

Tomasello, M. & Call, J. 1997. *Primate Cognition*. Oxford University Press.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. 2005. ‘Understanding and sharing intentions: The origins of cultural cognition’. *Behavioral and Brain Sciences* 28, 675-735.

Tomasello, M. & Gonzalez-Cabrera, I. 2017. ‘The Role of Ontogeny in the Evolution of Human Cooperation’. *Human Nature* 28, 274-288.

Tomasello, M., Kruger, A.C. & Ratner, H.H. 1993. ‘Cultural learning’. *Behavioral and Brain Sciences* 16, 495-552.

Toolan, M. 1996. *Total Speech: An Integrational Linguistic Approach to Language*. Duke University Press.

Torr, J. 1997. *From Child Tongue to Mother Tongue: A Case Study of Language Development in the First Two and a Half Years*. Monographs in Systemic Linguistics.

Toth, N. & Schick, K. 2009. ‘The Oldowan: The Tool Making of Early Hominins and Chimpanzees Compared’. *Annual Review of Anthropology* 38, 289-305.

Traulsen, A. and Nowak, M.A. 2006. 'Evolution of cooperation by multilevel selection'. *Proceedings of the National Academy of Science* 103:29, 10952-10955.
Trivers, R. 1971. 'The Evolution of Reciprocal Altruism'. *The Quarterly Review of Biology* 46, 35-57.

Uomini, N.T. 2009. 'The prehistory of handedness: Archaeological data and comparative ethology'. *Journal of Human Evolution* 57, 411-419.

Valian, V. 2006. 'Young Children's Understanding of Present and Past Tense'. *Language Learning and Development* 2:4, 251-276.
Vallar, G., Bisiach, E., Cerizzat, M., Rusconi, M.L. 1988. 'The Role of the Left Hemisphere in Decision-Making'. *Cortex* 24, 399-410.
Vehrencamp, S.L. 2000. 'Handicap, Index, and Conventional Signal Elements of Bird Song'. In Y. Espmark, T. Amundsen & G. Rosenqvist (eds.), *Animal Signals: Signalling and Signal Design in Animal Communication*, 277-300. Tapir Academic Press.
Vidal, C.M., Lane, C.S., Asrat, A., Barfod, D.N., Mark, D.F., Tomlinson, E.L., Tadesse, A.Z., Yirgu, G., Deino, A., Hutchison, W., Mounier, A. & Oppenheimer, C. 2022. 'Age of the oldest known Homo sapiens from eastern Africa'. *Nature* 601, 579-583.
Villar-Rodríguez, E., Cano-Melle, C., Marin-Marin, L., Parcet, M.A. & Avila, C. 2023. 'What happens to the inhibitory control functions of the right inferior frontal cortex when this area is dominant for language?' *eLife* 12:RP86797.
van Vugt, M. & Schaller, M. 2008. 'Evolutionary Approaches to Group Dynamics: An Introduction'. *Group Dynamics: Theory, Research, and Practice* 12:1, 1-6.
Vygotsky, L. 1935 [1978]. Interaction between Learning and Development. In *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press.
Vygotsky, L. 1934 [1986]. *Thought and Language*. MIT Press.

de Waal, F.B.M. 2011a. 'What is an animal emotion?' *Annals of the New York Academy of Sciences* 1224, 191-206.
de Waal, F.B.M. 2011b. 'Empathy in Primates and Other Mammals'. In J. Decety (ed.), *Empathy: From Bench to Bedside*, 88-106. MIT Press.
de Waal, F.B.M. 2009. 'Darwin's last laugh'. *Nature* 460, 175.
de Waal, F.B.M. 2006a. 'Bonobo Sex & Society'. *Scientific American Special Edition* 16:2, 14-21.
de Waal, F.B.M. 2006b. *Primates and Philosophers: How Morality Evolved*. Princeton University Press.

de Waal, F.B.M. 2004. 'Peace Lessons from an Unlikely Source'. *PLoS Biology* 2:4, 434-436.

de Waal, F.B.M. 1996. *Good Natured: The Origins of Right and Wrong in Humans and Other Animals*. Harvard University Press.

de Waal, F.B.M., Dindo, M., Freeman, C.A. & Hall, M.J. 2005. 'The monkey in the mirror: Hardly a stranger'. *Proceedings of the National Academy of Science* 102:32, 11140-11147.

de Waal, F.B.M. & Lanting, F. 1997. *Bonobo: The Forgotten Ape*. University of California Press.

Wacewicz, S. & Żywiczyński, P. 2018. 'Language origins: Fitness consequences, platform of trust, cooperation, and turn-taking'. *Interaction Studies* 19:1-2, 167-182.

Wang, E.T., Kodama, G., Baldi, P. & Moyzis, R.K. 2006. 'Global landscape of recent inferred Darwinian selection for Homo sapiens'. *Proceedings of the National Academy of Science* 103:1, 135-140.

Warneken, F. & Tomasello, M. 2006. 'Altruistic Helping in Human Infants and Young Chimpanzees'. *Science* 311, 1301-1303.

Watson, S.K., Townsend, S.W., Schel, A.M., Wilke, C., Wallace, E.K., Cheng, L., West, V. & Slocombe, K.E. 2015. 'Vocal Learning in the Functionally Referential Food Grunts of Chimpanzees'. *Current Biology* 25:4, 495-499.

Webb, C.E., Romero, T., Franks, B. & de Waal, F.B.M. 2013. 'Long-term consistency in chimpanzee consolation behaviour reflects empathetic personalities'. *Nature Communications* 8, 292.

Wehner, P. 2020. 'The Malignant Cruelty of Donald Trump'. *The Atlantic*, https://www.theatlantic.com/ideas/archive/2020/05/malignant-cruelty-donald-trump/612097/

Weir, A.A.S., Chappell, J. & Kacelnik, A. 2002. 'Shaping of Hooks in New Caledonian Crows'. *Science* 297, 981.

Weist, R.M. 1986. 'Tense and aspect: Temporal systems in child language'. In P. Fletcher & M. Garman (eds.), *Language acquisition: Studies in First Language Development*. Cambridge University Press.

White, F.J. & Wood, K.D. 2007. 'Female Feeding Priority in Bonobos, Pan paniscus, and the Question of Female Dominance'. *American Journal of Primatology* 69, 837-850.

White, M. & Ashton, N. 2003. 'Lower Palaeolithic Core Technology and the Origins of the Levallois Method in North-Western Europe'. *Current Anthropology* 44:4, 598-609.

Whiten, A. & Byrne, R.W. 1988. 'Tactical deception in primates'. *Behavioral and Brain Sciences* 11, 233-273.

Whorf, B.L. 1956. *Language, Thought and Reality*. MIT Press.

Wickler, W. 1965. 'Mimicry and the Evolution of Animal Communication'. *Nature* 208, 519-521.

Wiessner, P.W. 2014. 'Embers of society: Firelight talk among the Ju/'hoansi Bushmen'. *Proceedings of the National Academy of Science* 111:39, 14027-14035.

Wilkinson, G.S. 1984. 'Reciprocal food sharing in the vampire bat'. *Nature* 308, 181-184.

Willis, S.L., Schaie, K.W. & Martin, M. 2009. 'Cognitive Plasticity'. In V.L. Bengtson, M. Silverstein, N. Putney & D. Gans (eds.), *Handbook of Theories of Aging (2nd edition)*, 295-322. Springer.

Wilson, D. 2011. 'Parallels and differences in the treatment of metaphor in relevance theory and cognitive linguistics'. *Intercultural Pragmatics* 8:2, 177-196.

Wilson, E.O. & Tyson, P. 2008. 'Lord of the Ants: A conversation with E.O. Wilson'. *Nova Online*, https://www.pbs.org/wgbh/nova/eowilson/biophilia.html.

Wilson, R.A. 1937. *The Miraculous Birth of Language*. British Publisher's Guild.

van Wolkenten, M., Brosnan, S.F. & de Waal, F.B.M. 2007. 'Inequity responses of monkeys modified by effort'. *Proceedings of the National Academy of Science* 104:47, 18854-18859.

Wood, B. & Grabowski, M. 2015. 'Macroevolution in and Around the Hominin Clade'. In E. Serrelli & N. Gontier (eds.), *Macroevolution: Explanation, Interpretation and Evidence*, 345-376. Springer International Publishing.

Worden, R.P. 1998. 'The Evolution of Language from Social Intelligence'. In J.R. Hurford, M. Studdert-Kennedy & C. Knight (eds.), *Approaches to the Evolution of Language*. Cambridge University Press.

Wrangham, R.W. 1975. *The Behavioural Ecology of Chimpanzees in Gombe National Park, Tanzania*. Dissertation submitted to the University of Cambridge for the degree of Doctor of Philosophy, 5.51-5.60.

Wray, A. 2002a. 'Dual Processing in Protolanguage: performance without competence'. In A. Wray (ed.), *The Transition to Language*. Oxford University Press.

Wray, A. 2002b. *Formulaic Language and the Lexicon*. Cambridge University Press.

Wray, A. 2000. 'Holistic Utterances in Protolanguage: the link from primates to humans'. In C. Knight, M. Studdert-Kennedy & J. R. Hurford (eds.), *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form*. Cambridge University Press.

Xin, J., Huang, K., Yi, A., Feng, Z., Liu, H., Liu, X., Liang, L., Huang, Q. & Yaqiong Xiao, Y. 2023. 'Absence of associations with prefrontal cortex and cerebellum may link to early language and social deficits in preschool children with ASD'. *Frontiers in Psychiatry* 14:1144993.

Yang, C.D. 2002. *Knowledge and Learning in Natural Language*. Oxford University Press.

Yokose, J., Marks, W.D. & Kitamura, T. 2024. ‘Visuotactile integration facilitates mirror-induced self-directed behavior through activation of hippocampal neuronal ensembles in mice.’ *Neuron* 112:2, 306-318.e8

Zahavi, A. 2003. ‘Indirect selection and individual selection in sociobiology: my personal views on theories of social behaviour’. *Animal Behaviour* 65, 859-863.

Zahavi, A. & Zahavi, A. 1997. *The Handicap Principle: A Missing Piece of Darwin’s Puzzle*. Oxford University Press.

Zdrazilova, L., Sidhu, D.M. & Pexman, P.M. 2018. Communicating abstract meaning: Concepts revealed in words and gestures’. *Philosophical Transactions of the Royal Society B* 373, 20170138.

Ziff, P. 1960. *Semantic Analysis*. Cornell University Press.

Zlatev, J., Wacewicz, S., Zywiczynski, P. & van de Weijer, J. 2017. ‘Multimodal-first or pantomime-first? Communicating events through pantomime with and without vocalization’. *Interaction Studies* 18:3, 465-488.

Zuberbühler, K. 2000. ‘Referential labelling in Diana monkeys’. *Animal Behaviour* 59, 917-927.

Index

As an online publication, this book is fully searchable using standard software. An index list is therefore not provided.

The Sources of Language Grammar

An Anthropological Perspective

Martin P.J. Edwardes

SCITSIUGNIL PRESS

ISBN: 978-1-9999369-4-5

9 781999 936945