the Church to be a gift from God. The theory is therefore presented in the form of a fable, which makes interesting reading.

Gordon W. Hewes deserves much of the credit for reviving the theory, and his 1973 article in *Current Anthropology* provides the platform for modern accounts, including my own, and the commentary following that article provides a useful critique. Armstrong et al.'s (1995) book *Gesture and the Nature of Language* explores the implications of sign language for an understanding of language in terms of gesture, and also explains how syntax might have evolved from simple manual gestures. This theme is further developed in Armstrong (1999).

The article by Rizzolatti and Arbib (1998) provides an accessible account of the nature of 'mirror neurons' and their possible role in the evolution of language, and my own recent book *From Hand to Mouth* (Corballis 2002) expands on all of the themes developed in this chapter.

12

The Origin and Subsequent Evolution of Language

Robin I. M. Dunbar

Language has two remarkable properties. First, it allows us to communicate ideas with each other; second, languages evolve and diversify with a speed and facility that is quite unique within biological evolution. The first has been the focus of much of the research on language over the past half century: ever since Chomsky's formative statement of the programme for linguistics, the mechanism that makes it possible for language to transmit information (i.e. grammar) has been the focus of a concerted attempt to unravel the processes that allow language to encode information. This emphasis on the formal aspects of grammatical structure has, however, tended to obscure a number of interesting questions that lie at the very roots of language as a biological phenomenon, namely whence it came and why it mutates so readily, spawning dialects and ultimately new languages in profusion.

The reasons why language evolved exclusively in the hominid (and possibly only human) lineage have been dealt with at length elsewhere: the burden of these considerations has been to suggest that language evolved for an essentially social function to facilitate the bonding of large social groups (see e.g. Dunbar 1996; Deacon 1997). In this chapter, I want to explore the other two issues (the precursors of language and the ease with which dialects form) in a little more detail. My concern here is to highlight questions rather than to provide answers. First, however, let me sketch out briefly the arguments for why language evolved.

Why Did Language Evolve?

Perhaps the fundamental problem associated with the evolution of language is why it evolved at all. Conventional wisdom has always assumed that it did so to facilitate the exchange of information—after all, that is what grammar is basically designed to do. This claim is, I guess, not in serious contention. A more exacting issue is what kind of information language was intended to convey. It would, I think, be fair to say that most people would, at least until very recently, have supposed that this was related to information about hunting or the manufacture of tools. 'There were bison down at the lake yesterday when I was passing there' or 'If you want to make an arrowhead, you need to hit the flint nodule right here to strike off a suitable flake'.

What is unsatisfactory about such claims is that (a) these kinds of technological activities take up a relatively small proportion of our time and (b) when we do engage in them, we actually rarely use language when doing so. Hunting is often best done in silence, and tool-making is best done by demonstration rather than instruction. In addition, the size of human social groups gives rise to a serious problem: grooming is the mechanism that is used to bond social groups among primates, but human groups are so large that it would be impossible to invest enough time in grooming to bond groups of this size effectively. The alternative suggestion, then, is that language evolved as a device for bonding large social groups—in other words, as a form of grooming-at-a-distance. The kind of information that language was designed to carry was not about the physical world, but rather about the social world. Note that the issue here is not the evolution of grammar as such, but the evolution of language. Grammar would have been equally useful whether language evolved to subserve a social or a technological function.

Testing between alternative hypotheses is not easy when, as in the case of language, there is only a single instance of it. However, the issue essentially boils down to a choice between two possible pathways to language as we have it. Both assume that, in the here and now, language can be used for both social and technological information exchange: the real question is which came first. The traditional view is that language evolved to subserve a technological function and, once we had it in place, its intrinsic capacities allowed it to be used for other more trivial purposes (such as social information exchange); the alternative view is that language evolved to subserve the social function, but its intrinsic properties allowed it eventually to be used for other purposes (e.g. technology) that were in themselves valuable. The issue, then, is which was cause and which consequence?

Two lines of evidence have been adduced (see Dunbar 1993; 1996) to support the second view. One is the factual observation that about two-thirds of speaking time in informal conversations is devoted to social topics (comments about myself (as speaker), the listener or a third party, planning social activities and the like). Only a very small proportion is devoted to technical topics or descriptions of the physical world. Indeed, we tend to find the latter topics rather boring and hard going, even when we are explicitly engaged in learning about them. In other words, most of what we devote our conversations to is servicing our social relationships. In some cases, this involves direct investment in a particular relationship (by spending time talking to that person), in others it involves finding out what has been happening in the community to which I belong. Language allows us to keep track of events in the community in a way that our monkey and ape cousins simply cannot do: what they do not see, they can never know about.

The second line of evidence is a purely logical one: if language does not act as bonding agent, what does? We cannot ignore that question, because the claim that language's primary function is the exchange of technical information assumes that large, stable, coherent social groups actually exist. Dismissing that question as irrelevant to language leaves the technological version of language evolution founded on sand, because the conventional primate bonding devices (grooming) are not good enough to maintain groups of the size typical of modern humans.

Finally, we may note that there are at present three versions of the social language hypothesis: Dunbar's (1993) social gossip (or social bonding) hypothesis, Deacon's (1997) social contract hypothesis and Miller's (1999) Scheherazade Effect hypothesis. Deacon's argument, in a nutshell, is that language became necessary in order to allow us to make the contracts on which sociality depends (in particular, the contracts that prevent us from running off with someone else's spouse when wives get left behind while husbands are off hunting). Miller's argument is that language evolved to allow us to advertise our suitability as mates and/or to allow mates to keep each other entertained (and so interested in each other) once a relationship has been established. In fact, these three hypotheses can be unified fairly easily, since the social contract and Scheherazade effect hypotheses both require the pre-existence of large bonded groups (otherwise they run into the same problem as above) while at the same time providing valuable additional benefits to languages in different domains once language (in its gossip/bonding form) is in place. In effect, we can see them as valuable consequences of language that reinforce its further rapid evolution rather than its primary selective advantage as such.

Having sketched in the background, let me now turn to the two issues that I want to discuss in more detail: how language evolved and why languages diversify so readily.

The Precursors of Language

The origins of language (and speech) are inevitably shrouded in mystery, buried in an inaccessible fossil past that leaves only the most indirect traces of behaviour. Although there is strong anatomical (Kay et al. 1998; MacLarnon and Hewitt 1999) and comparative (Aiello and Dunbar 1993) evidence to suggest that speech first appeared with the earliest *Homo sapiens* at around half a million years ago, the process by which speech evolved out of an organism with an essentially primate communication system remains uncertain. There have been two general positions: one claims that speech evolved out of an essentially primate vocalization system (Dunbar 1993; Aiello and Dunbar 1993; Burling 1993; MacNeilage 1998*a*), the other that it evolved via a stage of gestural communication (presumably from an origin in primate gestural communication) (Hewes 1973; Calvin 1983; Corballis 1983; Kimura 1993).

Although given very wide credence, the latter position has been based on what is essentially indirect evidence. Typical examples include the observation that apes appear to be better at gestural languages like ASL than vocal languages, or that deaf and dumb human children seem to develop signing languages complete with grammar spontaneously (i.e. without need of teaching), or that we use gestures as a part of everyday speech-based communication. The claim that aimed throwing would have provided a natural substrate off which to build fine motor control for hand signals has been a particularly potent symbol of this view.

Most of these arguments do not really provide more than circumstantial (and far from conclusive) evidence in support of a gestural origin for language. If aimed throwing provided fine motor control for gestures, for example, it might well have provided a template for the genuinely fine motor control that is crucial for speech, but this does not imply that language necessarily passed through a gestural phase before becoming based in speech. Similarly, gestural languages may develop effortlessly in the deaf and dumb children in today's environment, but in the past the deaf and dumb were notoriously regarded as imbeciles *precisely* because they were unable to

communicate *in any form* with the rest of social community; the gestural languages of the deaf and dumb seem to develop naturally off the back of *spoken* language (or at least the cognitive substrates required for spoken language—the so-called 'language of thought') when members of the speaking community encourage the deaf and dumb to communicate using a gestural form of transmission. Gestural languages do not seem to develop spontaneously all that easily.

Gestural languages, however, have two genuinely serious disadvantages. One is the fact that they are entirely useless at night. Second, they require line-of-sight contact between communicators (and thus gain no advantage over the social grooming of primates: see Dunbar 1993)—a matter of very serious concern in a natural environment where a great deal of time has to be devoted to foraging. This is likely to be especially problematic if (as would seem more likely) the ancestral environment was more like scrub woodland (with its dense under-storey and poor visibility) than open, short-grass savannah. The merit of speech in this latter context is precisely that it allows time-sharing with other activities without reducing the efficiency with which the latter can be carried out. In other words, we can talk while we eat, dig, and walk in a way that we simply cannot while gesturing—in the last case, not least because gestural communication needs to be done more or less face to face, and walking face to face is not the most efficient way of progressing.

More recently, Rizzolatti and Arbib (1998) have attempted to provide a neural underpinning for the gestural origins theory. They point to the fact that watching an experimenter pick up and manipulate an object triggers responses in two key areas in the human neocortex, the superior temporal sulcus and Broca's area. The latter turns out to be anatomically homologous with area F5 in the macaque brain, where so-called 'mirror neurons' (neurons that fire both when performing an action and when watching another individual perform the same action) were first discovered in the macaque. Rizzolatti and Arbib suggested that mirror neurons might have provided a bridge between performing actions and communicating with others. They suggested that the very slight muscle movements that occur when we observe the actions of others might have provided the basis for a gestural phase during language evolution.

Interesting though this suggestion is, it does not provide unequivocal evidence of a gestural origin for speech: if primitive language principally involves statements about actions, then we might well expect mirror neurons to be involved, irrespective of whether communication is gestural or vocal. An even stronger claim might be made if it is the case (as has been argued) that the real function of Broca's area is the management of fine breath control for speech: in this event, it is hardly surprising that its neurons should fire when upper limb action is involved, since bracing the chest is necessarily involved in any such movements. There is, however, another reason why the Rizzolatti and Arbib claim may be misleading: there is evidence that mirror neurons may also be involved in the cognitive phenomenon known as 'theory of mind', the process by which we simulate the mind states of others (Gallese and Goldman 1998). Theory of mind is probably essential for language, not so much because it is involved in the production of speech *per se* but because it provides the mechanism that both enables speakers to ensure that their message has got through and allows hearers to figure out just what the speaker's message actually is (subtext and all) (Worden 1998; Dunbar 1996). In short, the evidence for mirror neurons does not allow us to differentiate clearly between the two alternative hypotheses.

All this notwithstanding, the weakness of the gestural theory for language origins is perhaps that it is difficult to see what real advantage a gestural phase would have offered. Much hangs here on what is envisaged in any such gestural stage. If the gestural stage is assumed to involve full conceptual language as we have it now (or even some more modest precursor to what we have now), then the hypothesis raises some puzzles. The gestures that punctuate everyday speech do not seem to have the same illocutionary force as the speech acts they accompany. Nonverbal aspects of speech seem to have more to do with emphasis and command. They elaborate on what is said, but they almost never replace what is said. In this respect, human gestural communication does not seem to differ from the gestural communication systems of our primate cousins. Because of this, it is difficult to see what would have been gained by moving from one already efficient communication channel (sound) to another (gesture) and then back again, when a direct route from primate vocal communication to human verbal communication is clearly possible. In addition, one must ask why, if gestural languages were so well developed as to allow the exchange of conceptual information, we should have lost so efficient a mode of communication since the rise of speech. Atrophy is not an adequate answer in biological terms: if gestures can supplement emphasis and instruction as effectively as they do, then it does not make sense to suggest that they cannot do the same for the conceptual content of speech. Conversely, if the gestural stage of language involved gestural communication as we have it now (i.e. essentially emotional rather than conceptual in content), then there is not much to choose between the vocal and the visual channels. Nonverbal vocal cues can be (and are) just as effective as gestural cues. Moreover, that channel already exists and is used to good effect by all primates (and, indeed, by most other mammals too).

The burden of the argument, then, is that a vocal origin to human language seems a more plausible starting point than a gestural one. Although Owren and Rendell (2001) have argued vehemently against the claim that any aspect of non-human primate vocal communication pre-shadows human speech, their argument seems to focus on the assumption that what is being claimed is that non-human primates exhibit some (or even all) of the psycho-acoustic foundations of language (e.g. the ability to form consonants and vowel sounds). This is not the issue at all: such abilities probably depend on some uniquely human features, namely the fine control of breathing (to enable the phrasing of sound elements) and the peculiar physical structure of the human laryngeal and palatal spaces. The substantive issue is whether non-human primate vocal communication shows those features of human speech whereby *meaning* is attached to sounds. In this respect, the answer must be yes if any credence is to be given to the work of Cheney and Seyfarth (1990). As with the original chimpanzee language projects, the issue is not whether monkeys and apes have the physical apparatus to produce human speech sounds (after all, if they did, then they surely would be able to speak), but rather whether they have the cognitive components that allow meaning to be attached to arbitrary signals in order to transfer information from one mind to another. In other words, is the non-human primate mind a kind of paralinguistic mind waiting in the wings for the acquisition of an appropriate speech apparatus? Hovering behind this is the more fundamental question of whether other primates possess the kind of *social* mind that seems to be essential for human language: although there are dissenters (Povinelli 1999), many would argue that they do (Call 2001; Sudendorff and Whiten 2001). The question of whether or not non-human primates can produce vowel sounds is beside the point: that is merely a consequence of minor adjustments to the size and shape of the vocal apparatus.

The point here, of course, is that there is a world of difference between speech per se and language. The one is a (relatively) uninteresting channel of communication; the other is what the whole process of human communication is all about—the mental gymnastics involved in formulating and, subsequently, decoding utterances that have real meaning over and above the sign value of their basic constituents (words).

The issue, then, is whether non-human primates (or, if it comes to that, any other animals) share with us the ability to express conceptual information in a vocal channel. If the answer is in the affirmative, then the business of getting language off the ground is very easy. Indeed, Nowak et al. (Nowak, Krakauer, and Dress 1999; Nowak and Krakauer 1999; Nowak and Komarova 2001) have used a modelling approach to demonstrate that if the initial function of language is to allow a speaker to name 'objects' (things or events) and both speaker and listener gain a fitness pay-off if the information is successfully communicated between them (i.e. they cooperate in the exchange of information such that both gain in fitness as a result), language can easily evolve out of naming.

The key to their model is the risk of misunderstanding (i.e. the listener being mistaken about which object or event is being named), with the risk of misunderstanding being a function of how acoustically similar any two sounds (names) actually are. Nowak et al. (1999) showed that, if a language's fitness (indexed by its capacity to transfer information) is maximized by having a small number of simple sounds to describe a few valuable concepts or objects, then increasing the number of signals does not necessarily increase the fitness of the language. This suggests that non-human communication systems may be error-limited. However, this error limit can be overcome by combining sounds (phonemes) into words, since this has the effect of reducing the number of sounds that can be mistaken for each other. Since fitness increases exponentially with the length of the words used, word formation must have been a crucial step in the evolution of human language (the equivalent of moving from an analogue to a digital channel). Note that there have been claims that at least some non-human primates form utterances that are structured in this way (though these regularities in phonemic structure have usually been interpreted as syntactical: Cleveland and Snowdon 1981).

Nowak et al. (2000) extended this approach to the evolution of grammatical structure. They considered a language that uses *noun+verb* complexes to describe events and showed that, when there is a large number of relevant events, syntactic organization of the stream of words conveys a significant fitness advantage (less risk of misunderstanding). The point at which syntax becomes advantageous in this respect depends on (a) the difficulty of memorizing the particular signals and (b) the fraction of all word combinations that describe real meaningful events (contrast *pigs fly* with *pigs walk*). When a noun+verb phrase is twice as difficult to memorize as a non-syntactic utterance (e.g. a noun on its own) and one third of all possible noun–verb combinations are real possible events, then the minimum object–event combination that will make syntax advantageous is surprisingly small: a matrix of eighteen nouns by eighteen verbs. The dynamics of this system is, however, complex: in smaller systems (say, a six by six matrix), the number of learning opportunities becomes crucial: the fitness of a syntactic language only exceeds that of a non-syntactic one when the number of exposures (i.e. learning opportunities) is below about 400. In other words, when the number of possible noun–verb combinations is small but the frequency of exposure is very large (> 400 occasions), a non-syntactic system is actually more efficient.

In other words, so long as there is a modest selective advantage in terms of more effective cooperation between individuals, both language (in the simple sense of naming objects) and, more importantly, grammar (the ability to describe events) can evolve quite easily by a conventional Darwinian process. This makes it all a great deal easier to move from a simple non-human primate vocal system with referential meaning (like that of the Cheney–Seyfarth vervets) to a more sophisticated human-like language with grammar. Essentially, all that is required is an increase in the number of relevant events that it is useful to be able to describe to another individual. It is difficult to see how anything in the physical world of the later hominids could have been so different from that of their forebears (or indeed, most other primates) to fulfil that condition, but the demands of an increasingly large social world would seem to do so admirably.

One other point of interest: Nowak and Komarova (2001) show that what is important here is the accuracy of language acquisition. When this is low, a variety of different grammars can emerge in a population, but when it is moderate to high the population rapidly converges on one standard grammar, and a coherent communication system emerges. The point is that any cognitive mechanism that facilitates accuracy of acquisition (such as a Universal Grammar) necessarily pushes the system towards a single language form. This would suggest that there would be a strong selection pressure favouring some kind of neural hardwiring for a predisposition to attend to and learn language—any language—which, of course, is exactly what we see in human infants.

While it is feasible to argue that language evolved directly out of non-human primate vocalizations as a device for exchanging information, there remains an alternative possibility that might suggest a more long-drawn-out process. This would view information exchange as a later phase responsible for the explosive evolution of the language capacity but not, perhaps, for its initial appearance. Aiello and Dunbar (1993) noted that we might see language evolution as just such a multi-stage process. We might envisage these as (1) an initial form much like conventional non-human primate contact calls that serve to keep track of other group members (or, perhaps, as in the case of the gelada, to act as a form of grooming-at-a-distance between pairs of friends), (2) a more developed form of this (chorusing?) designed to overcome the physical constraints on grooming that limit group sizes to around sixty individuals, (3) a more fully fledged language that uses grammatical structures to convey social information and, finally, (4) a fully developed modern human language capable of deeply abstract symbolic representation of concepts (with the last of these perhaps involving changes more at a software than a hardware level). Aiello and Dunbar (1993) envisaged these stages kicking in successively over the course of hominid evolution as the need to evolve increasingly large social groups to cope with key ecological problems became more intrusive.

I want to focus here on the second step—the development of a more communal form of contact calling as an intermediate step between conventional non-human primate contact calling and fully fledged grammatical (but still exclusively social) language. The issue is about bonding social groups. Among Old World monkeys and apes at least, this is done mainly through social grooming, a deeply intense activity that facilitates the release of endogenous opioids (see e.g. Keverne et al. 1989). If the effectiveness of social bonding (and a willingness to cooperate and exchange support) is a function of the opioid high generated by social grooming, and if the size of the group that can be supported off the back of such a form of social interaction is limited by the amount of time that animals can afford to devote to social interaction, then the limiting size of social groups that we should expect to see among primates is about sixty—much as seems to be the case. Part of the problem here is created by the fact that social grooming is very much a one-on-one activity: primates (including humans) do not groom more than one individual at a time (even though several animals may occasionally groom the same recipient). So intense and personalized is this activity that failure to groom with sufficient commitment and intensity is often the precursor to the end of a grooming bout.

If the time available for social interaction is limited, then the only way

that group size can be increased beyond the limit created by this time constraint is for the available social time to be used more efficiently. There are, in principle, two ways of doing this: one is cognitively expensive, the other cognitively cheap. The expensive way would seem to be to introduce a quantitative change in the kind of information that can be transmitted during interactions: language does this very effectively by allowing us to seek out information on who is doing what with whom and how often. This, however, requires the computational devices capable of supporting language (minimally, the capacity to encode and process grammar, plus advanced theory of mind). An alternative, and cheaper, option is to increase the number of individuals that can be 'groomed' simultaneously. Language actually allows us to do this, because we can talk to several individuals at once in a way that is just not possible with physical grooming (see Dunbar et al. 1995). But we also have a way of achieving the same effect without language, namely by means of communal singing (and possibly dance). Singing seems to be an exceptionally powerful stimulant of endogenous opioids, as evidenced by our enthusiasm for it and the way it makes us feel generally relaxed and positive towards those with whom we have been singing.

The significance of singing in this context is that it does not require words (language as we know it), merely rhythmically synchronized vocalizations—in fact, just the kind of thing we already find widely among other primates. Although it is clear that communal singing must increase the interaction group size (the number of people simultaneously subjected to an opioid release effect), it is not clear what the magnitude of this effect is likely to be or by how much it would increase social group size above the limiting size that can be sustained by conventional social grooming. What it might at least do is bridge the gap between this limiting size (somewhere around sixty individuals) and the size of social groups that eventually emerged off the back of conventional language (around 150) (see Dunbar 1993). An intermediate step in the process of language evolution that involved an intense communal activity like chorusing would provide a strong natural pressure in the direction of a verbal means of communication.

Why Do Languages Diversify?

In the previous section I focused on the question of how language came about. But given that language evolved to allow information to be exchanged (everyone agrees, for example, that this is what grammar is well designed to do), why on earth do languages diversify so rapidly that they very quickly become mutually unintelligible? After all, there are now some 6,000 living languages, plus an untold number that have already become extinct. Within languages, dialects are equally diverse: until as late as the 1970s, a native English speaker's dialect was sufficient to identify the speaker's natal location to within 40 km. of his or her place of birth (Trudgill 1999). Languages spawn dialects with unseemly speed, and dialects in turn eventually give rise to new languages. Ancestral Indo-European, the language spoken by a small group of (probably Anatolian) agro-pastoralists around 6000 BC, has given rise to around 150 descendent languages from Gaelic in the west to Bengali in the East over an 8,000-year period. Interpolating these values into the standard Gaussian logistic growth equation for biological population growth suggests that the Indo-European language family has evolved at a rate equivalent to the budding off of a new language from each existing language, on average, about once every 1,600 years—or about as long as it has taken modern English and Scots to evolve out of Anglo-Saxon. (We might, of course, expect geographical dispersion to have been important in addition, as in the case of the modern Latinate languages.)

It seems that the question as to why languages should be so labile in this respect has rarely been asked. From an evolutionary point of view, however, this is a singularly puzzling phenomenon. After all, if language evolved to allow individuals to exchange information, one might expect stability to be an important consideration, especially in a multi-generational community where the 'wisdom of the ancients' might provide an important resource for survival. Yet dialects seem to evolve with a speed that approximates the scale of a generation (Barrett et al. 2002). Since it is not beyond the wit of evolution to have produced language structures that are more resistant (if not totally resistant) to corruption in this way, the implication is that the corruptibility of language is precisely the whole point (and has been deliberately selected for).

One answer, of course, is simply drift: the gradual accumulation of accidental mutations (mispronunciations, unintended slippages of meaning) over long periods of time. Such a process is relatively slow, however, and would not be expected to lead to quite such a degree of differentiation in so short a space of time. If the process is not accidental, then it must be deliberate, and deliberate in this context means 'under the influence of selection'. What selection processes could promote such high rates of language change?

The most plausible selection pressure is likely to be the need to differentiate communities. The key problem faced by all intensely social organisms that depend on cooperation for successful survival and reproduction is the freerider—the individual who takes the benefits of cooperation but does not pay the costs (Enquist and Leimar 1993; Dunbar 1999). Nettle and Dunbar (1997; Nettle 1999*b*) have argued that dialects are particularly well designed to act as badges of group membership that allow everyone to identify members of their exchange group: dialects are difficult to learn well, generally have to be learned young and change sufficiently rapidly that it is possible to identify an individual not just with a locality but also within a generation within that locality. They used a simple spatial model (in which a dialect was a very simple six-digit barcode attached to each individual) to show that a rate of dialect change approaching 50 per cent per generation was required to ensure that individuals who had to exchange resources with each other in order to reproduce were not exploited by freeriders.

Nettle (1999*b*) analysed the world distribution of languages in relation to a number of climatic variables (including latitude, ambient temperature and the length of the growing season) and showed that there were more languages per unit area, each with a smaller number of speakers, in low-latitude habitats (with long growing seasons) than in high-latitude habitats (with short growing seasons). These results may reflect the influence of two (not entirely mutually exclusive) selection pressures. One is that, in more seasonal habitats (such as those at high latitudes), individuals need a wider network of social contacts in order to buffer themselves against the vagaries of the natural environment. Since famine is likely to be both unpredictable and more frequent, the trading network needs to be wide enough to ensure that at least someone you know well enough to trade with is living in an area of surplus during famine years. Reciprocal arrangements across this large set of individuals ensures that food supplies are always available. Being able to communicate well enough to trade is, of course, important; but being able to signal descent from a common ancestor by using the same language may be crucial in facilitating the willingness to engage in long-term reciprocal exchange with strangers.

In low-latitude habitats in the tropics, where the growing season is more or less continuous, famine is a rare event and individuals are more selfsufficient. Because such habitats are capable of sustaining very high population densities, individuals living there incur high costs from the effects of competition and crowding, not to say the perennial risk of individuals using sheer muscle to raid and steal from their more industrious neighbours (the classic producer/scrounger problem). To buffer themselves against these costs, they need to ensure that their social networks work extremely well. Small networks are more effective in this respect because it is possible to bond small groups more effectively than large ones; in addition, it is easier to identify members of the group through badges like dialects. So we might anticipate pressure to diversify languages in the equatorial regions.

Beyond this, however, we have little real understanding of the processes involved in either dialect change or language evolution, or for that matter in the functions that these processes subserve. We assume that these functions are largely social, and we have some understanding of the types of process that can precipitate language change (trade, colonization, emulation of culturally or economically superior groups, etc.), but by and large there is little other than conjecture to explain why these processes exist or why they should work in the way they do. Part of the problem is, of course, the temporal scale on which these changes occur (generations in the case of dialects, perhaps millennia in the case of languages). Inevitably, this makes it all but impossible for us to observe the process first-hand. These are, nonetheless, fruitful areas for future exploration.

Notice that this explanation for language instability is based on the same argument as that for language evolution, namely the importance of bonding small communities or social groups. In terms of parsimony, this provides considerable support for the social bonding argument, since the alternative hypothesis for the evolution of language (technical information exchange) cannot explain why languages should diversify so easily. It requires some other unrelated explanation to account for this feature of language behaviour.

Conclusions

In this chapter I have focused on two questions about language that have received a great deal less attention than they deserve. One was how language evolved in the first place, and the other was why languages are so prone to diversification. My aim here has been to identify neglected issues that need more detailed study, rather than to provide conclusive answers. In offering some suggestions as to what those answers might be, I have endeavoured to offer what I think are plausible directions in which that research might go.

Although there has been some desultory discussion of the first issue (how language came to evolve at all), much of this has focused on a rather odd form of communication (gesture) and paid very little attention to what is actually involved in communication (e.g. its cognitive requirements). I have tried to suggest some reasons why gestural theories of language origins do not make sense. Whatever the origins of language, it is clear that some additional steps are needed between the basal primate communication system from which language must have evolved and language as we know it today. I have previously suggested that this intermediate step (or steps) is likely to have involved the use of primate-like vocalizations in chorusing. I presented some additional arguments favouring this suggestion, one of which is the way that communal singing seems to trigger endorphin release (thus mimicking social grooming). It is clear that once a vocal channel of this kind is in fluent operation, it requires very little (as Nowak's models have shown) to push it that one crucial step further, to the stage at which meaning is attached to sounds (especially given that primates seem already capable of doing that).

The second issue that I highlighted (the rapidity with which languages diversify, first into dialects and then into true languages) has received almost no attention. This is particularly surprising because this remarkable characteristic seems to run counter to everything that we commonly suppose language evolved to do (namely to facilitate the exchange of information). The implication is that language's corruptibility has been directly selected for. I suggested that this might have been to facilitate group bonding by allowing community members to identify each other more easily (the 'dialects as social badges' argument). This explanation has the distinct merit of allowing both the evolution of language and the evolution of language diversification to be derived from the same basic idea—the suggestion that language evolved to subserve an essentially social function (the bonding of what are, by primate standards, large communities).

FURTHER READING

The best general summary of my ideas about the evolution of language as a device for social bonding is my book *Grooming, Gossip and the Evolution of Language* (Dunbar 1996). Much of the empirical data on which this hypothesis depends is scattered in the primary research journals; since this is difficult to track down, the book provides an accessible non-technical summary. Deacon's *The Symbolic Species* (1997) provides an additional excellent review of the neuroanatomical and other aspects of language origins, while offering a particular spin on the social language theme. Nettle's *Language Diversity* (1999*b*) is a particularly important volume, summarizing a great deal of his work on how and why languages diversify so rapidly.

13

Launching Language: The Gestural Origin of Discrete Infinity

Michael Studdert-Kennedy and Louis Goldstein

Introduction

'Human language is based on an elementary property that also seems to be biologically isolated: the property of discrete infinity' (Chomsky 2000: 3). 'Discrete infinity' refers to the property by which language constructs from a few dozen discrete elements an infinite variety of expressions of thought, imagination, and feeling. The property 'seems to be biologically isolated', because it is unique among systems of animal communication. From another point of view, however, it is not isolated at all, but rather an instance of a general principle common to all natural systems that 'make infinite use of finite means' (Humboldt 1836/1999: 91), including physics, chemistry, genetics, and language, namely, 'the particulate principle of self-diversifying systems' (Abler 1989).

The Particulate Principle

According to the particulate principle, the only route to unbounded diversity of form and function is through a combinatorial hierarchy in which discrete elements, drawn from a finite set, are repeatedly permuted and combined to yield larger units higher in the hierarchy and more diverse in structure and function than their constituents. The particulate units in physical chemistry include atoms, ions, and molecules; in biological inheritance, chemical radicals, genes and proteins; and in language, gestures (as will be argued below), segments, syllables, words, and phrases.

Preparation of this chapter was supported in part by Haskins Laboratories and by grants NICHD HD-01994, NIDCD DC-00403 to Haskins Laboratories.