THE HISTORY OF THE HOLISTIC PROTOLANGUAGE IDEA

9834332

MSc EVOLUTION OF LANGUAGE AND COGNITION

THE UNIVERSITY OF EDINBURGH

DECEMBER 2009

TABLE OF CONTENTS

ABSTRACT

Research in language evolution has increased dramatically over the last two decades. The subject is of necessity multi-disciplinary and research from diverse fields of inquiry has helped inform the various hypotheses offered for language evolution scenarios. Computational and mathematical modelling, evolutionary biology, archaeology as well as linguistics are just some of the areas that help shape our conception of the origins of language. However, a hitherto untapped area in language evolution research is the history of the subject itself. A longitudinal study of certain theories related to the question is worthwhile in that it can both highlight the staying power of specific ideas, and perhaps stimulate discussion as to why they have persisted. One such idea is that of the holistic protolanguage. This dissertation will focus on the persistence of the claim that there was a stage in language evolution that was entirely holistic. As such the time-frame I will be looking at will cover the mid-to-late 18th century to the present day.

Word count (excluding references, footnotes and abstract): 16,300

PART 1: CURRENT PROTOLANGUAGE THEORY

1.1 INTRODUCTION

A recent series of articles and exchanges on the linguistics web site *Language Log* clearly shows the continuing commitment to certain ideas surrounding the emergence of language. These ideas revolve around the conceptualisation of a stage in language evolution that has come to be known as *Protolanguage*. In his guest post on *Language Log* W. Tecumseh Fitch (2009) revives the idea of a Musical Protolanguage stage in language evolution by citing the work of Charles Darwin (1871). This theory in itself is not without support from other authors and it is therefore not surprising that it should reappear again at this time (cf. Mithen, 2005; Brown, 2000). What is surprising is the immediate response and critique it provoked from a prominent supporter of a competing theory of protolanguage, Derek Bickerton (2009). The subsequent (sometimes heated) debate underlies issues I believe to be at the heart of thinking about the evolution of language; namely, the question of language's contiguity with animal communication; whether protolanguages that bridge the evolutionary gap were holistic or compositional; and how much of language's structural complexity is dependent on an innate, syntactical component versus more general constraints acting on cultural transmission.

Another interesting outcome of Fitch's article is the reference to an early appearance of the idea of a holistic protolanguage by the linguist Otto Jespersen (1922). However, an even earlier view that expresses the same idea of holistic proto-utterances can be found in Bentham (1843)[1]. This idea is still salient today (Wray, 1998; Arbib, 2005b; Bowie, 2008). Indeed, an entire issue of the journal *Interaction Studies* (Vol. 9:1, 2008), has been devoted to the idea of "Holophrasis vs Compositionality".

This review will focus on the history of ideas that relate to the issue of holistic protolanguage, and its persistence through the years, be they musical, manual-gestural, or speech-like in nature. Along the way we will have to consider language as related to animal communication in general; the general merits of individual theories of protolanguage extant today; and their place in the wider study of language

[1] Much of Bentham's work was published posthumously by his executor in 1843. It is therefore difficult to ascertain the date for Bentham's *Essay on Language*. It was probably written in the late 18th century – much earlier than the publishing date of 1843.

evolution. Although the main counterpoint to this theory is taken to be the synthetic/compositional view, the focus in this discussion will be holistic theories; the opposing view will be limited to the role of exemplifying the differences between both whilst devoting the lion share of discussion to the holistic.

1.2 NOTES ON TERMS

In the following review, I will be using a number of terms that are central to some of the competing theories of human protolanguage. The central dichotomy to have emerged in a number of key books and articles (Bickerton, 1990; Wray, 1998; Tallerman, 2007; Arbib, 2005b) is between a *holistic* (or holophrastic) versus a *compositional* view of protolanguage. An utterance is holistic when it is "a complete communicative act" (Arbib, 2008: p.154) exclusively associated with one proposition where that meaning is taken from the whole not the sum of its parts, and "there are no component parts that could be recombined to create a new message" (Wray, 1998, p.51). Wray's example is characteristic of a holistic meaning-signal mapping: *tebima* would be the utterance and 'give that to her' the meaning, with no part of tebima 'standing for' any of the elements in the meaning.

An important distinction to be noted with reference to the use of the term holistic in this essay is with the study of *holism* in the philosophy of language. In particular semantic holism stresses that the meaning of a word or sentence in any language can only be understood with reference to the whole of that language (Fodor & Lepore, 1992). Here we will only be concerned with the narrower definition of holistic as described above.

In contrast, the *compositional* depiction of protolanguage stresses that the first words were simple one-to-one mappings between signals and their referents, but without syntax. More accurately, the route from simple words without syntax to words with syntax is *compositional*, while the protolanguage itself can be referred to as *synthetic*. Modern language characterized in this way conveys meaning through expressions that are a function of their sub-parts. In other words the sentence 'the cat sat on the mat' is meaningful because we know the meaning of the elements it is composed of: 'cat', 'mat', 'on' etc.

With reference to the holistic route from protolanguage to modern language the process has been described variously as *segmentation*, *fractionation* and *analysis*. To avoid confusion, especially concerning *analysis* which can also describe an opposite meaning to *holistic*; I will use Wray's (1998) term *segmentation* when discussing holistic protolanguage development. For the sake of clarity, the term *protolanguage* as used in this essay will refer exclusively to the hypothesised, intermediate communicative stage(s) between a non-linguistic forebear and a language-using (in the modern sense) human.

1.3 THE HOLISTIC AND SYNTHETIC VIEWS

Alison Wray (1998, 2000, & 2002) and Michael Arbib (2002, 2005a & b, & 2008) are the two most prominent researchers proposing a holistic stage in language evolution. In this view communication at some point was effected by means of individual utterances that expressed entire propositions. Only later in time did compositionality emerge when these utterances were broken down – or segmented – into more familiar word-like units. Although both of these authors share certain views – a stage of holistic utterances and a process of cultural development that acted on these to yield modern compositional language – they also differ on a number of points. Wray (1998 & 2000) seeks to link directly back to earlier non-hominid communication, citing functional similarities between modern human and non-human primates, claiming that "holistic utterances in chimp communication are a subset of those in human language" (Wray, 2000, p.289).

Arbib (2002 & 2005b), whilst also discussing a primate link, breaks his analysis down into many intermediate stages involving the neural restructuring of a 'mirror system' of neurons. Arbib also hypothesises a stage of pantomime and 'protosign' before hominids were then "language ready" (Arbib, 2005b). Wray (2000, p.292) acknowledges the need for biological evolution in developing control in the "transition to arbitrary phonetic representation" but focuses more on the similarities between the social function of communication systems of primates and modern human language and the segmentation process (see Tallerman, 2007, for a critique).

Bickerton (1990, 1995, 2000, & 2002), contra Wray and Arbib, proposes that the protolanguage spoken by our hominid ancestors should be regarded as the simple,

non-syntactic concatenation of basic word-like meaning-signal pairs. For a non-linguist this seems the most intuitively parsimonious given the way we look for definitions of single words in dictionaries.

The core hypothesis proposed by Bickerton (1995, p.51) is that language evolution comprises a "two-stage process: first a stage in which there was a lexicon without syntax, then a stage in which infinitely productive mechanisms emerged to create syntax as we know it". Although the initial words in the proto-lexicon are evolved out of the prior cognitive abilities of pre-hominids (Bickerton, 1990, pp.74-75), these words in contrast to Wray's model do not have antecedents in the call systems of primates (Bickerton, 1995, pp.56-57). This, Bickerton reasons, is due to the proposal that "the antecedent hominid call system and the burgeoning protolanguage had to be kept separate" because the reactions to call systems are potentially more crucial to an individual's survival than the uses of protolanguage (ibid). The uses Bickerton discusses are ecology-based, so that, again, contra Wray (1998), the "selective pressure that started us on the road to modern human language cannot have come from social intelligence" (Bickerton, 2002, p.209). Instead the main function was to exchange information about foraging and environmental conditions.

1.4 MUSICAL THEORIES

Musical precursors to language have long been hypothesised (Thomas, 1995; for a review see Fitch, in press). The related idea that language and music share certain properties has fuelled language evolution theories in the last decade (Brown, 2000; Mithen, 2005; and Fitch, in press). Earlier writers and thinkers had conceived of a link between music and language. Taken together, along with Darwin, these writers link language origins with, not only animal cries, but also an explicitly musical stage of human communication. The philosopher Condillac (1714-1780) conceives of "A primeval song-language" that "is the transition that leads the first societies from instinctive cries to language and reflection" (Thomas, 1995, p.72). However, Charles Darwin (1871) provides the template for explanations of language evolution using evidence from the world of comparative biology (Merker & Okanoya, 2007; Fitch, 2006). In current models that posit a musical precursor for language, evidence from species who use song as a means of display figure largely, due to structural

similarities between language and song. There is also a connection based on the neural and vocal requirements of birds, for example, to produce streams of organized sound. Interestingly, many of the musical hypotheses extant converge on a stage of holistic protolanguage. In part 3 of this paper I shall discuss these models in more depth.

1.5 PLAN

The weight of this review will be on the various theories centred on the idea of a holistic protolanguage. This idea goes back to at least the late 18th century in the form of Jeremy Bentham's *Essay on Language* (1843). Therefore Part 2 will deal with early theories of language evolution and thus provide a backdrop against which Bentham's theory emerged. At the same time Part 2 will also deal with how various models, past and present, have dealt with the problem of how language, specifically meaning, could have arisen in the first place.

Part 3 will move on from the discussion of meaning to look at the nature of holistic utterances themselves, providing a breakdown of the views of two contemporary researchers. Wray (1998) and Arbib (2005b) have presented full accounts of what they envisage a holistic protolanguage to have looked like. In this section I will also look at the relationship with music that other authors have put forward.

Part 4 takes us out of protolanguage and into language, looking at how the various theories we have hitherto encountered could possibly have developed into modern language. Prominent in this section will be Wray's (1998) theory of segmentation which, to varying degrees, other advocates of the holistic view have supported (Arbib, 2005b, Mithen 2005). As an adjunct to this, research in the area of computational modelling, which has shed light on the feasibility of a possible trajectory out of a holistic protolanguage and into modern, compositional human language, will be discussed.

PART 2: THE PROBLEM WITH GETTING MEANING-SIGNAL PAIRS UP AND RUNNING

2.1 ENLIGHTENMENT IDEAS OF LANGUAGE ORIGINS

2.1.1 INTRODUCTION

The period of history commonly referred to as the Enlightenment can roughly be considered to span the cultural and intellectual developments in Europe from the 1670s to the early nineteenth century (Kors, 2003, p.xvii). The Enlightenment period, moreover, represents a marked shift in *where* learning and the dissemination of ideas took place. For centuries the church was at the centre of European education and literacy, and learning was to be found in the monasteries and the courts. Thus philosophy was aligned with catholic dogma and the natural sciences were made to conform to a homo-centric view of the universe, as exemplified by Galileo's forced recant of his astronomical discovery that the earth revolved around the sun.

During the Enlightenment period, however, there arose the academy or society: organized associations committed to discourse in all areas if inquiry – science, philosophy and history. Groups could devote themselves to questions

> mostly taken from the natural sciences, but academies also concerned themselves with language, literature and history – subjects which were otherwise pushed into the background by the all-powerful faculties of theology and law (Im Hof, 1994, p.108).

The growth of academies over the entire Enlightenment period charts another trend that correlates with the move away from the church. The earliest academies set themselves the philosophical task of "the defence by rational argument of the divine personality and the moral responsibility of man" (ibid, p.111). However, the rise of smaller groups of thinkers and the spread of learning amongst the lesser nobility and the emergent middle classes led to an interest in devoting "energies more and more to utilitarian ends" (ibid). That is, the focus of inquiry remained the same – philosophy, the natural sciences etc. – but the application was increasingly towards "the realms of society, economy, law, and human relationships" (Kors, 2003, p.xvii). In short there

was an increasing secularisation in European thinking and inquiry. It was in this changing climate that a growing number of investigations into language origins arose.

2.1.2 EARLY LANGUAGE THEORIES

Of course, secularisation does not mean philosophers suddenly stopped believing in God, nor does it imply that the increase in scientific knowledge and discovery meant that man suddenly made God obsolete. Rather, the disciplines themselves – philosophy, history, science – existed, and developed, in an increasingly independent world from church authority. Thus, as James H. Stam (1976, p.14) notes, early language origins theories such as those of Giambattista Vico (1668-1744) framed language origins with reference to the Bible. In this view man is created in the image of God and Adam, being the first man, is the inventor of language.

Later theories, notably Jean-Jacques Rousseau's (1712-1778), moved away from the idea of divine origins to claim that "speech being the first social institution, owes its form to natural causes alone" (Rousseau & Herder, 1781/1966, p.5). He also proposed that language was, of necessity, invented by man who, recognising his fellow man as sentient and thoughtful, aroused "the desire or need to communicate his feelings and thoughts" making "him seek the means to do so" (Rousseau & Herder, 1781/1966, p.5). This begs the question: how can sentience and thought be perceived in another? What were humans doing, behaviourally speaking, that differentiated them from the rest of the world that impelled the need to communicate in a way not already instantiated? Stam (1976, p.82) also identifies a similar "cul-de-sac" in Rousseau's conception of language origins; namely that language was hard to conceive without prior social organization but that paradoxically social organization could not arise without language in place already.

That this kind of issue was recognized by Rousseau is no small matter, research into the topic still stimulates intense discussion. Bickerton (1990, pp.8-9) framing the problem in evolutionary terms, highlights what he calls the Continuity Paradox. By definition the evolution of language involves development out of a prior system and according to Bickerton there is no such system (that is, no *communication* system). Language in his view is primarily a representational system and as such doesn't fit into the continuity models proposed some researchers (Wray, 1998; Arbib, 2005;

Mithen, 2005; Fitch, in press). Bickerton (2002) also stresses the need for an "ecology-based" account of language origins that takes into consideration the difference between early hominids and modern primates environments and argues that we can't take for granted our knowledge concerning hominid environmental conditions (Bickerton, 2002, p.213-214). Thus, we have to consider carefully the context in which a protolanguage would have arisen, and its use in that environment to avoid getting stuck in similar paradoxes. Rousseau similarly rallies against the "shortcomings of Europeans" who "philosophize on the origins of things exclusively in terms of what happens within their own milieu", failing to realize that the "human race originated in warm climes" (Rousseau & Herder, 1781/1966, p.30).

There is a striking sense of modernity in some of these theories. Writing in 1772 Johann Gottfried Herder (1744-1803), as well as denying a divine origin for language, recognised that animals are in some sense constrained by their environments; that their form and behaviour is a result of the habitat they have adapted to. In relation to communication: "The narrower the sphere of an animal, the less its need for language" (Rousseau & Herder, 1772/1966, p.105). In contrast, humans, able to exist in a wide variety of contexts, must have a more general means of communication. According to Herder they have "senses for all things and hence naturally weaker and duller senses for each one" (ibid).

At this point it should be noted that the previous paragraph contains a similar chicken and egg scenario to that of Rousseau. Where does the human ability to survive successfully in many different ecological niches come from? Is it facilitated by having language-like communication that can be applied to many different situations? If so then we have to explain how a general purpose language arose in a formerly narrow, constraining environment.

This caveat aside we can illustrate the differences between human and animal communication that Herder envisaged. For example, where the Vervet monkey has a call system that specifies danger from different types of predator: snake, eagle, or leopard; this will elicit appropriate flee responses: look to the ground, seek ground cover, or run up a tree respectively. This is all they can 'say' and do about the predators they are likely to encounter. Human language on the other hand can convey information about the size or number of predators; it can imagine the likelihood of a predator in the future, or recall a particularly nasty encounter in the past. Language users can even choose to ignore the information as spurious. The communication

system of Vervet monkeys is thus narrowly confined to its environment as opposed to the seemingly endless generativity of language.

With the advent of evolutionary research, the growing study of ethology and linguistics there is a need to clearly delineate between animal communication and human language. Herder clearly views animal communication in general as a type of language: "the natural language of that animal species" (Rousseau & Herder, 1772/1966, p.89). But nonetheless it is interesting to note the appeal to an animal's environment as shaping force in its communicative behaviour.

Although Rousseau ended up in a hopeless paradox concerning language genesis, he believed it to have arisen through invention. Firstly language, or rather protolanguage, arose distinctly through the modularity of speech rather than gesture/pantomime. In distinguishing the senses of sight and sound for communicating Rousseau holds the view that pantomime is the most immediate method, depending "less upon conventions" and rendering "more exact imitation" (Rousseau & Herder, 1772/1966, pp.6-9). However, pantomime as related to action seems to be more iconic in Rousseau's view and tied to the needs of individuals. The vital distinction then, which led to the auditory modality being chosen over the visual one, is summed up as follows:

> "It seems then that need dictated the first gestures, while the passions stimulated the first words…It is neither hunger nor thirst but love, hatred, pity, anger, which drew from them the first words" (ibid, pp.11-12)

Thus gesture/pantomime is related to the activity of survival and speech to the social sphere. Language arose by the "invention of the most ancient words" to help select a mate and manage inter-group relationships (ibid). The nature of the words was musical and contained much onomatopoeia and imitation; in addition they were in the form of aphorisms – short 'sentences' that contained truths or maxims. Thus, Stam (1976, p.89) interprets meaning to be at the level of the sentence rather than the word in Rousseau's theory. However, we are no closer to a satisfactory answer as to how a shared set of arbitrary signal-meaning pairs could have arisen. Herder (Rousseau & Herder, 1772/1966, pp.117-118) similarly credits the origin of the first words to man's invention, as a result of reflection on the perceivable world. Through perception man is able to distinguish objects in the world and, upon reflection,

organize and categorize them based on their characteristic features. Herder's example is the naming of a sheep due to its distinguishing characteristic – its bleat. Unlike animals who act according only to instinct, man, "his soul in reflective exercise seeks a distinguishing mark – the sheep bleats" (ibid). Repeated encounters evoke previous ones and the "distinguishing mark of the sheep became…the name of the sheep" (ibid). Thus, humans could be said to have created the first words for objects that display behaviours that can be imitated. Herder is appealing implicitly to a pre-linguistically rich cognitive architecture, albeit one which he sees as being well in advance of anything found in other animals (Rousseau & Herder, 1772/1966, pp.105-106).

The problem is, again, how the individuals in a group agree on which part of the environment corresponds to the newly minted symbol. An issue related to reference and meaning is addressed in the work of Quine (1960) who identified what he called the "Gavagai" problem. This is the potential difficulty arising between two speakers of different languages trying to convey what they mean to each other, but can be extended to individuals in the present scenario. To take Herder's example, if someone imitated the aforementioned sheep with the idea of using that imitation as a word referring to the sheep then the hearer can't know exactly what is meant by use of the new word. If it merely means 'there is a sheep' then the word is of little use if uttered to a hearer who can already see the sheep. Even if the sheep is absent, uttering the word is hardly going to accomplish much unless your goal is just to put the idea of a sheep in the hearer's mind, which is not particularly useful. Further still, the utterance might express a variety of propositions: get the sheep; kill the sheep; avoid the sheep; move the sheep etc. The point is that protolanguage as a communication system is useful only if it can achieve things that benefit the users, and explaining a protolanguage without bridging these fundamental gaps will leave a theory wanting.

2.1.3 ANIMAL COGNITION

Despite taking language origins out of the hands (or mouth) of Adam, Herder and Rousseau seem to deny animals the cognitive faculties out of which language might have arisen. Herder's earlier reference (p.12) to the soul still suggests humanity's divine origins, only instead of being 'given' language; we were given 'mind' instead.

However, we can hardly criticise the lack of a theory describing cognitive continuity; Darwin was over half a century away and the mental lives of animals is still a divisive issue to some thinkers (Hurford, 2007b, p.9).

The idea that a richer cognitive architecture than previously envisaged preceded language is the subject of much of Hurford's work in the field of language origins. Hurford (2003 & 2007a) seeks to provide a "neural basis of predicate-argument structure" theoretically depicting a basic representational system in, not only, pre-linguistic humans but perhaps a larger class of mammals. This system is based on a simplified version of mathematician Gottlob Frege's logical scheme, which allowed the mapping from one domain – sentences of a language – to another – their logical representations. Hurford (2003) concerns himself mainly with 1-place predicates of the form PREDICATE (x), where (x) is a variable taken to be some object in the world and PREDICATE is some characteristic that can be attributed to (x) (p.264). For example 'the ball is red' would be represented as ∃x [BALL (x) & RED (x)]. ∃ is the existential quantifier and says (x) exists, the predicates BALL and RED assign attributes to (x). The argument of the paper is that these simple predicates correspond to properties apprehended by the senses in a wide variety of animals (Hurford, 2003: p.264). Furthermore, they existed before human communication and are fundamental to concepts in animals (ibid, p.263). Additionally, the variable (x) is correlated with instantiations of "whole objects attended to" (ibid, p.275). In interacting with its environment an animal will plan its behaviour according to information gleaned from its senses. The crucial information will come in the form of the predicate-argument structure (Hurford, 2007a, p.535). The rather abstract notion of predicate and argument is rooted by Hurford (2003) to the underlying neural correlates in the dorsal and ventral visual streams of the brain: "Both cortical streams process information about the intrinsic properties of objects and their spatial locations" (pp.267-268). Elsewhere, Hurford (2007a) has expanded on this idea by proposing that this hierarchical organization was 'co-opted' by syntax and that the ubiquitous "S/NP distinction [in modern language] now serves a pragmatic communication function" (p.528). That is, the Topic/Comment distinction that underpins the handling of new and given information in discourse.

Hurford (2003) stresses that possession of the PREDICATE (x) system is not enough to propel a species towards language – it is merely one of the many proposed

pre-adaptations that language was able to utilise in its development (p.264). This vein of research can, if given further empirical support, provide a fuller picture of what is needed to get to the stage of protolanguage. Although Hurford aligns himself with Bickerton's (1990) proposal of a "primary representation system" serving "in the first instance merely to label protoconcepts derived from prelinguistic experience" (Bickerton, p.91 quoted in Hurford, 2003); the subsequent 'shape' of the protolanguage is by no means a given. Whether or not the predicate-argument paradigm would, from a protolanguages inception, structure utterances in a bipartite fashion; or facilitate the subsequent segmentation of initially holistic utterances is a question for further research. For the present it is important to note the grounding of any protolanguage in a prior cognitive system.

2.2 THE ROOTS OF THE HOLISTIC VIEW

2.2.1 JAMES BURNET

As we have seen from section 2.1.2., some thinkers held to the idea that humanity's ultimate language ability arose out of the expression of emotions and passions. Rousseau seems to have had some conception of words as being holistic (he uses the term aphorism), associating what we would call words with propositional content more equivalent to a sentence. He didn't develop this idea further so we can only speculate about the nature of his aphorisms.

Although the term 'holistic' was not in existence until the early twentieth century, one of the earliest theories of language origins to which it might apply, in the modern sense of Wray (1998) and Arbib (2005b), is that of Lord Monboddo (James Burnet). Monboddo (1714-1799) devoted six volumes to his work *Of the Origin and Progress of Language* (1773-1792). Like Rousseau, Monboddo characterized the initial vocalizations made by man as "nothing but an improvement or refinement upon the natural cries of the animal" (Burnet, 1773, p.318). Monboddo envisages a steady progress or development in the ability to speak. Language arises incrementally from cries that have a limited tonal range through to greater articulatory control. Articulation gradually improves and more complex vowel and consonant sequences are able to be realised. The trajectory from animal-like cries to increasingly fine-

grained articulation in the vocal tract is clearly, for Monboddo, a discovery: "it was natural that so sagacious an animal as man should go on further, and come at last to…articulation" (ibid, p.321). Invention, or at least recognition of a good idea, is invoked in this theory. Wells (1987, p.20) relates that the earliest communication between individuals was proposed by Monboddo to involve facial expressions, painting and imitative sounds alongside the cries we have noted. However, man's sagacity led him to persevere with vocal sounds as facial expressions and painting had drawbacks when it came to communicating in the dark or at a distance. Thus, through continued use the first phonologically structured 'protowords' would have developed "distinguished only by a few vowels and consonants" (Burnet, 1773, p.322). Monboddo is quite explicit about the nature of this stage in his protolanguage, that "the first articulate sounds that were formed denoted whole sentences" expressing an individual's desires in the "common business" of the community (ibid, p.395). This is a view also to be found in the writings of Jeremy Bentham to whom we now turn.

2.2.2 JEREMY BENTHAM

Jeremy Bentham (1748-1832) gives us, by far, the earliest and fullest description of what we now call the holistic utterance. However, before dealing with his theory of language origins it is necessary to look at his motivation for viewing the first utterances in this way. Bentham recognized the problems in meaning all too well. His study of English law and later of legislation led him to become "conscious of the scope for ambiguity and confusion which arose from the wealth of synonyms which existed and from the variety of meanings which a single word might have" (Dinwiddy, 1989, p.39). Thus, if reform was to be sought in law and legislation it must be achieved via the clarification of its language:

> Fiction, tautology, technicality, circuity, irregularity, inconsistency remain. But above all, the pestilential breath of Fiction poisons the sense of every instrument it comes near.
>
> (quoted in Ogden (1932, p.xvii) from Works Vol. I, p.235)

Indeed, as Ogden (pp.xx-xxi) notes in his introduction to Bentham's *Theory of Fictions*, the "Fictions of Law" were to occupy Bentham to the extent that he postponed publication of *An Introduction to the Principles of Morals and Legislation* (1789) until he had dealt sufficiently with this problem. Furthermore, "many of the most illuminating footnotes are concerned with linguistic difficulties and particularly with ramifications of fictional analysis" (ibid, p.xxi).

The Fictions of Law can be characterised by way of example. Definition for Bentham involves the categorisation of objects into genera, species and sub-species, and definition is the logical result of apprehending "some peculiar character or quality by which it stands distinguished from all other objects included in that same collection – from all other *species* of that same *genus*" (Bentham, 1843, Works Vol. VIII, p.593, italics in original). So for concrete objects such as 'table' we could define it with reference to its membership of the genus 'furniture'. Additionally, we can refine the definition by distinguishing our 'table' from other species of furniture by listing its characteristic features in contra-distinction to other types of furniture. In contrast, the words 'right', 'power' and 'obligation' (in their legal sense) to use Bentham's examples, although in the same word class (i.e. nouns) cannot be defined as above; none of these words is a "species of anything" (Ogden, p.lxxvii). They are thus characterised as "fictitious entities". They are not real in the way that tables, chairs or animals are and cannot be defined in terms of a hierarchy of sense relations. The problem with reference to law and legislation is that words such as 'obligation' lack, in Bentham's view, a clear definition and undergo a kind of reification that only belongs to 'real' entities (Ogden, p.7).

Bentham's initial motivation for inquiries of a linguistic nature was thus bound to the problem of meaning, specifically of meaning in relation to law. For Bentham, language was closely tied to logic (Dinwiddy, 1989, p.39), and it is in the philosophy of logic where recognition of his theories on language is to be found (Quine, 1966). The answer to ambiguity in meaning, however, came in the form of Bentham's *Theory of Fictions*.

2.2.3 THE THEORY OF FICTIONS

Bentham's theory held that the subject matter of thought, language and perception were entities. That is, anything representable by a noun-substantive (i.e. a noun denoting an object, material or immaterial). In addition entities are divided into *perceptible* or *inferential*: perceptible entities are apprehended by people empirically through the "immediate testimony of their senses"; inferential entities are conversely only known through "reflection…inferred from a chain of reasoning" (Ogden, pp.7-8). A further bipartite distinction is made for both perceptible and inferential entities and that is that they can be either real or fictitious.

As we have seen above the distinction between the real and the fictitious rests upon whether we can define it successfully or not. Real, perceptible entities having an objective physical presence can be defined in a categorical way with reference to real world properties. On subsequent occasions can be brought to mind in their absence by use of that word. Real inferential entities are not available to the human senses but are postulated through a reasoning process. Thus Bentham can class the soul as such an entity

Fictional entities, however, require a different explanatory process from that of definition. If the use of words that describe fictitious entities are defined without caution they can cause "a propensity and disposition to suppose the existence, the real existence, of a correspondent object" (Ogden, 1932, p.xxxv). Overcoming this confusion "involved defining a term not by presenting a direct equivalent of it, but by what Bentham called *paraphrasis*: by providing equivalents of all desired sentences containing the term" (Quine, 1966, p.659). From this logical analysis of the semantics of words and their definitions came the proposal that the earliest words were holistic in nature: "their language is all in propositions" (Bentham, 1843, Works Vol VIII, p.322). The nature of holistic utterances will be dealt with in section 3. For the present we should note that although Bentham based his description of an early protolanguage on his logical analysis of modern language, his thoughts about the ultimate origin of this communication system were similar to his contemporaries. Like them and every researcher since, we can only hypothesize how a putative protolanguage emerged.

His treatment of meaning, however, is certainly relevant today. It is apparent that the holistic view has persisted within theories of language evolution; but it was also invoked in the mid-nineteenth century in the form of Bentham's concept of *paraphrasis*.

Willard Quine's (1966) recognition of Bentham as a precursor of ideas to be found in Gottlo Frege's (1848-1925) and Bertrand Russell (1872-1970) highlights a central problem in semantics: the relationship between meaning, truth and object (Quine, 1960, ch.1). Frege is quoted as saying: "We ought always to keep before our eyes a complete proposition. Only in a proposition have the words really a meaning…It is enough if the proposition taken as a whole has a sense" (quoted in Hallett, 1967, p.10).

There is an idea that the communicative-pragmatic aspect of language deals with propositions in there entirety. Sentences that we use day-to-day in order to socially navigate our interactions, whilst capable of yielding a full syntactic analysis, are often processed 'whole'. For example "how do you do?" or "can I help you?" are cited as examples of formulaic utterances (Wray, 2002).

This, by extension becomes an issue in language evolution – how does a system of arbitrary meaning/signal pairs become instantiated in a group of individuals. As we have seen, many early thinkers proposed that language was an invention, by early humans as a way of expressing their thoughts; Bentham included (cf. Bentham, 1843, Vol VIII, p.323). However, the perspective on this scenario has changed given the developments in the many fields of science that inform the question of language evolution. Indeed, a theory of evolution that was to encompass all living species and their behaviour, man included, was not to appear for over fifty years. As we will come to see a separate treatment of man, including language was over sixty years away.

2.3 OTTO JESPERSEN

Outlining Darwin's theory of language origins for a modern audience, Tecumseh Fitch (2009) recognises at least one problem in his Darwin "redux": "complex phrasal semantics remains unexplained by Darwin's model". Central to a theory of language evolution is how and when do arbitrary strings of phonemes come to have the shared meanings that they do. How can a symbols meaning be conveyed (by arbitrary representations) to someone who does not already know the meaning without recourse to yet more arbitrary, meaningless symbols? Darwin's hypothesis was that "language owes its origins to the imitation and modification of various natural sounds" and that through mimicry "may not some unusually wise ape-like animal have imitated the

growl of a beast of prey, and thus told his fellow-monkeys the nature of the expected danger?" (Darwin, 1871, pp.109-110).

In section 3 we will treat more fully the theory of Charles Darwin (1871) concerning language origins. For now it will suffice to note that Darwin's theory helped give shape to the ideas of the early 20th century linguist Otto Jespersen. He takes up Darwin's idea of a singing ancestor and, according to Fitch (2009) remedies the semantic "oversight" that Darwin's theory failed to account for the combining of meaning to sung phrases.

Building on the work of the philologists of the previous century Otto Jespersen utilised the most "fruitful source of information" available for any inquiry into the origins of language – that is, the history of language (p.416). Two other sources are considered: children's acquisition of their native language and the "language of primitive races" (ibid). Regarding the former type of data, Jespersen considers the earliest babbling of babies to be a possible area of research in finding "some clue to the infancy of the language of the race" (Jespersen, 1922, p.417). This view is related (although supporting an opposed theory) to Bickerton's hypothesis (1990, p.110; 1996, p.50) that sees the language of children under the age of two as 'fossils' of a synthetic protolanguage.

The latter data source we can dismiss, along with Jespersen, linguists having rejected notions that so-called civilized societies' language structure would be qualitatively different from those languages used by people occupying a different cultural/ecological niche. Any child from any background can acquire any language given a 'normal' upbringing in a language speaking environment. Jespersen, however, does not reject the "languages of contemporary savages" on these grounds for he still holds to the idea that these languages may illustrate "a linguistic stage that is nearer to that in which speech originated." (p.27). Rather, he implicitly sees language change as, in some sense, directed towards an optimal state:

> "it should never be forgotten [that] even the most backward race has many centuries of linguistic evolution behind it, and that the conditions therefore may, or must, be very different from those of primeval man." (p.27)

This kind of view has fallen out of favour in subsequent linguistic research. The prevailing approach is that of uniformitarianism: in some sense all human languages

are the same. It is true that any infant in the appropriate environment can learn any language, but does this say more about the innate learning mechanisms involved in language acquisition than the structure of the languages themselves? There is both a biological and a cultural element to acquiring a language and cultural changes are more rapid, shaped by the performance and cognitive constraints of the language community. Whether a languages structure is affected by the cultural environment in which it is used is a controversial subject (see Everett (2005) and related comments for a study exploring this view). If the syntactic structure of language can be related to the uses of a linguistic community then it might have implications for the earliest structure of language tied to the needs of its speakers.

The researches carried out by the likes of Rask, Grimm and Bopp revealed tendencies in the way that languages change over time. Studying the comparative differences between cognate words in different languages revealed a systematic set of corresponding phonemic changes. From these "sound shifts" languages were grouped together into language families where current languages descended from earlier 'parents' such as Proto-Germanic and Indo-European (O'Grady et al, p.358). Thus, Jespersen reasoning backwards tries to reverse engineer the likely phonological 'shape' of the earliest primitive language. Noting that in speech there is a general "tendency to make pronunciation more easy", Jespersen reasons that earlier forms would be characterized by a higher frequency of "all kinds of difficult sounds" (Jespersen, 1922, pp.418-419). Using the same kind of inductive reasoning, the disappearance of tonal and pitch aspects in some languages, giving way to stress patterns that have obscured the tone element, is argued to support a more musical language in the past. As an adjunct to this it is observed that "the modulation of sentences ["sentence melody"] is strongly influenced by the effect of intense emotions in causing stronger and more rapid raisings and sinkings of the tone" (p.420). In contrast, calm, conversational speech oscillates relatively little between high and low, the pitch being more monotonous. Jespersen sees the loss of musicality as a result of an ongoing 'civilizing' process whereby passion is tempered in more complex societies (ibid).

From this reasoning, based on the study of historical linguistics, the conclusion is reached "that we must imagine primitive language as consisting …of very long words, full of difficult sounds, and sung rather than spoken" (p.421). It is not difficult

to imagine that from this the earliest languages would have been extremely phonologically complex (Wells, 1987, p.52).

Let us now clarify and chart the proposed development of language in early humans set out by Jespersen:

i. "Our remotest ancestors" sang in an "exclamative", non-communicative way, expressing "an inner craving of the individual without any thought of any fellow creatures" (p.436). This is likened to the roar of animals or the singing of birds, by which I interpret Jespersen as meaning a lack of intentionality.
ii. From this "frivolous use" the vocal apparatus was exapted to serve an increasingly communicative function "so that it became more and more fitted to express everything that touched human souls" (p.436). Jespersen characterizes this system as "neither deep nor wise" (p.435).
iii. The next stage deals with the question of how sound and meaning came to be associated. For a small number of cases (onomatopoeia and interjections) the process was easy and "such words were at once employed and understood as signs for the corresponding idea" (p.438). However, the bulk of the first words would have been proper names referring to distinct individuals. This is imagined to have happened through the association of a song or "leitmotiv" with an individual, arising through the pairing together of potential mates. Others, noting the co-occurrence of song and individual, "would occasionally banter him by imitating and repeating" (ibid).

In Part 3 we will see that this last idea has persisted and been promoted in a situation that sees a musical protolanguage functioning as it does in songbirds (Merker & Okanoya, 2007). This too invokes the holistic paradigm by associating songs with whole situations or events. However, for the moment we will look at how meaning is handled by the theories of Alison Wray and Michael Arbib.

2.4 THE MODERN VIEW

Wray (1998), in supporting a holistic protolanguage, appears to sidestep the issues of meaning origins by invoking continuity with primate communication. Wray argues that the vocal and gestural signals of primates function in relation to "self-preservation, the preservation of others and the servicing of and adjustment of relationships" (1998, p.51). In this way, the argument goes, the signals are intentional, and being employed to achieve a definite effect in an interaction, and are thus propositional in content: "The signals are holistic: there are no component parts that could be recombined to create a new message" (ibid). But Wray makes no mention of the difference between human and primate communication with regard to the fact that signals are able to be learned across generations, whereas primate signals are learned and negotiated between individuals over a lifetime but not passed on to the next generation (cf. Tomasello et al, 1997). This is definitely discontinuous and should be accounted for as it is the infants of the next generation that must acquire associations between signals and meanings, holistic or otherwise.

In the "smooth, gradual, and beneficial" development to hominid protolanguage, the resultant communication system is imagined to be similar to that of primates "but capable of a more complex inventory" (1998, pp.50-51). Continuity is all very well but there needs to be more detail if the issue of referential meaning developing is to be addressed without having to explain why it isn't more ubiquitous in other primates.

The work of Arbib (2002, 2005b & 2008) whilst appealing to the idea of continuity, has also stressed the importance of a pantomimic stage in human language evolution. The crucial bridge that needs to be crossed in getting to a "language-ready brain" is grounding symbols between speakers and hearers, so that what counts for the speaker correspondingly counts for the hearer (Arbib, 2005, p.105). Arbib refers to this as the "parity requirement" (ibid, p.106). This requirement is met "because Broca's area evolved atop the mirror system for grasping, with its capacity to generate and recognize a set of actions" (ibid).

The crux of Arbib's argument, then, rests on the substantial body of research outlining the so-called Mirror Neurons and their implication in language evolution (see Rizzolatti & Arbib, 1998 for an overview). According to this hypothesis there exist a class of neurons – in an area of the monkey brain known as F5 – called mirror neurons. These are active when certain actions are performed by a monkey and also when those same actions are observed. Area F5 of the monkey brain has a

homologous region in the human brain which is part of Broca's Area – one area associated with speech. Not only is Broca's Area associated with grasping gestures it is implicated in comprehending sentences and has been linked to Wernicke's Area – a region of the brain related to sentence production (Lieberman, 2006, pp.196-197; Arbib, 2005, p.106).

The mirror system, which is thought to have been present in the common ancestor of monkey and human, is hypothesized to have further developed in the common ancestor of humans and chimpanzees – around 5-6 m.y.a. The resulting system is said to facilitate simple imitation: "imitation of short, novel sequences of object-directed actions through repeated exposure" (Arbib, 2005, p.115). The Rubicon, if you will, that then sets the stage for hominid development is the evolved brain mechanisms that allow for complex imitation: "acquiring (longer) novel sequences of more abstract actions in a single trial" (ibid). In Part 3.4 it will be seen that the mirror system provides the foundation for a complex, evolutionary scenario where manual, as well as vocal, gestures evolve in tandem.

2.5 THE PAST RE-IMAGINED

From the foregoing, it can be seen that during the Enlightenment period – particularly, the late 18th-early 19th centuries – there was vigorous interest in the question of language origins. Aarsleff (1982, p.147) identifies the 18th century as one which debated the question of language with "greater zeal, frequency, consistency, and depth of insight" than any other. Had he been writing today Aarsleff might have remarked that with the progress of modern science, the 18th century zeal for inquiry into language origins has found its mirror in the burgeoning research and literature in the field over the last two decades. Furthermore, it might be argued that the parallel extends to the fact that interest in the question was, in part, motivated by the causal factor of scientific progress.

The Enlightenment period saw acceptance of Newtonian physics, the beginnings of the Industrial Revolution and the increasing spread of knowledge in general through the academies, books and encyclopaedias (Heilbron, 2003, p.134). In our own time we have seen the human genome mapped, the brain investigated in ever more

detail and the development of computers such that they provide powerful ways of modelling sophisticated and dynamic systems (for instance language) over time.

However, despite the relative sophistication of technology and a concomitant increase in information, a good idea is a good idea, regardless of time and place. This is perhaps why we can see so much overlap between theories of language evolution from the 18th century and those of the last twenty years. Early theories are necessarily framed by the state knowledge available at the time and the tools available to test these theories. In the subsequent sections the nature of various holistic protolanguages will be fleshed out in more detail. Hopefully what will emerge is a picture that increasingly becomes informed by how much research-based knowledge is brought to bear on the subject.

PART 3: THE PROPOSED NATURE OF THE HOLISTIC MEANING-SIGNAL PAIR RELATIONSHIP

3.1 INTRODUCTION

Whilst the previous section looked at the inherent problems in establishing meaning in a group of protolanguage users, this section will primarily be concerned with what holistic units of meaning might have looked like. This will include a consideration of protolanguage models that link music and communication.

Bentham's model of language origins as we have seen stems from his investigations into modern English. It is, however, also tied into his overarching concept of utility. For the sake of clarity I will outline the basic scenario of language origins proposed by Bentham and then discuss it with reference to the idea of utilitarianism.

3.1.1 BENTHAM REVISITED

In his *Essay on Language* (Works Vol. VIII, 1843) Bentham seems to propose that thought is prior to language and that language has arisen chiefly through the need to communicate (Works Vol. VIII, p.320-323). Further to this, language has two distinct but related uses which are referred to as the "purely self-regarding" and the "extra-regarding" (ibid, p.301). The self-regarding mode of language is related entirely to the individual's thoughts and ideas and is conceived of as helping "improvement of thought" (ibid,). The extra-regarding use of language is geared towards *communication* of thought. Moreover, it "is the use to which language is indebted for its existence, it was, for a long time, not only the only use actually made, but the only one which was even so much as in contemplation" (ibid). That is, language is seen as essentially a social construct, and, moreover, a tool with which "we communicate with others either to convey information, excite emotions, or prompt certain courses of action" (Stam, 1976, p.42).

For Bentham "the primary and only original use is the communication of thought, the conveyance of thought from mind to mind" (Works Vol VIII, p.320). In the philosophy of the principle of utility (also known as the greatest happiness principle), objects are seen to have the property of utility if they result in an increase

of pleasure, benefit or happiness for the individual or community; and a corresponding decrease in pain or unhappiness (Honderich, 1995, p.85). Given, then, that we are governed by pleasure and pain, and that our actions to some extent are occupied with the respective increase and decrease of these states can we assume that Bentham saw this as a motive force in the emergence of communication? I think we can, given that he clearly viewed language as a function of the need to convey ones thoughts, and that these thoughts would largely have been occupied with increasing an individual's pleasure or reducing their pain. In the context of a group, then, it would seem that achieving the greatest happiness principle could only be possible by communicating to ensure those ends.

Having set out the uses of language in early society, but not stating an explicitly causal relationship between utility and language origins he goes on to describe those utterances:

> The first words must in their import, have been equivalent to whole sentences, to sentences expressive, for example, of suffering, of enjoyment, of desire, of aversion. Of this original language, the parts of speech called interjections are examples. (p.322)

From this communication system words that represented propositions – "the original sentences" – were, through a process of analysis, "as it were, broken down into words, these words into syllables" (p.323). The composition of these propositional utterances is not explicitly outlined, whether they were phonetically complex as Jespersen would later imagine them; or if they were composed of relatively simple syllabic units as is suggested by the preceding quote. Bentham does, however, offer the following analogy where letters are to words so were "words but so many fragments" of the earlier holistic utterances.

We will leave a more detailed discussion of the route to modern language until Part 4. Now I will consider theories whose nature is musical as well as holistic.

3.2 A MUSICAL INTERLUDE

3.2.1 DARWIN AND JESPERSEN

In *The Descent of Man* (1871) Darwin proposed three stages to the evolution of language. The first is a development of pre-human cognition "before even the most imperfect form of speech could have come into use" (p.110). Darwin's second stage (which Fitch (in press) calls prosodic protolanguage) involved the evolution of singing "producing true musical cadences" which served the functions of courtship and marking territory, which can be seen in analogous birdsong. Additionally, these songs expressed emotions such as love and jealousy (p.109). The final transitional stage on the way to language rests on "the imitation and modification of various natural sounds, the voices of other animals, and man's own instinctive cries, aided by signs and gestures" (ibid).

Darwin's theory of a musical protolanguage seems to inaugurate a much more rigorous, scientific approach to the question of language origins. Earlier theories such as those outlined in Part 2 were addressed in a more philosophical and speculative manner without much recourse to empirical evidence, which, to be fair, did not exist in most cases. In contrast Darwin, as he had done in *Origin of Species* (1859), brought a wealth of comparative biological insight to the question. This methodological approach can also be seen in Jespersen's work. He clarified, linguistically, aspects of Darwin's model to put forward his own ideas with regard to linking music and language. Darwin's theory is developed from the wide range of comparative data he gathered, particularly the insight of the analogous song learning of birds with human language learning (1871, pp.108-109).

In first discussing Jespersen (section 2.3) we noted the musical characteristics of his protolanguage. The first three stages depicted there chart the development of meaning out of initially meaningless, sung phrases. The third stage might, in current parlance, be depicted as maintaining dyadic, interpersonal relationships. Now we come to the stage that can most accurately be described as holistic in nature. 'Stage 4', as I have called it, might have originated to augment group cohesion:

> If a certain number of people have together witnessed some incident and have accompanied it with some sort of impromptu song or refrain, the two ideas are

> associated, and later on the same song will tend to call forth in the memory of those who were present the idea of the whole situation. (p.440)

Meaning is a group creation based on melodic, holistic utterances. Through processes of language change there is "a progressive tendency from inseparable irregular conglomerations to freely and regularly combinable short elements" (p.429). We will come to this in Part 4 after looking at more recent theories that appeal to musical origins. In particular the quote above, outlining the association of songs with situations, seems to foreshadow a similar hypothesis we will encounter in the next section.

3.2.2 FITCH AND PHONOLOGY

Championing Darwin and Jespersen, Fitch (in press) advocates a vocal-learning based account of language emergence. The musical route is based on a similar appreciation for the comparative evidence of vocal learning in other species as well as being a parsimonious account of phonology and vocal tract development. A major criticism of Wray's model (Tallerman, 2007) has been the lack of a coherent explanation of phonological development (see section 3.4.1 below).

Focusing on the congruencies between music and language Fitch (in press) notes the close analogy between aspects of phonology and music. Music is universal to all human cultures as are phonological systems; both involve a generative system of building phrases from small meaningless units (notes and phonemes); and both involve hierarchical combination (Fitch, in press, Ch.14.0, para.7). This parallel is fundamental to Fitch's prosodic protolanguage theory in that it appeals to the convergent evolutionary evidence of species of songbird and is parsimonious in explaining some of the later structural elements of language. Fitch thus describes animal song as "bare phonology: generative, arbitrary vocalization lacking discrete meaning" (ibid, para.6). Because phonology has the generative feature of producing sequences composed of smaller units it has the ability to generate novelty, albeit meaningless novelty. Recent research suggests that in birds at least there is a patterning to song-strings that can be described grammatically with a finite-state syntax (Okanoya, 2002); and, more controversially, with a context-free recursive grammar (Gentner et al, 2006).

The capacity for vocal learning in evolutionarily disparate species has informed much recent comparative literature dealing with music and language evolution (Merker & Okanoya, 2007; Mithen, 2009; Fitch, in press). This vein of research highlights the importance of vocal learning and the nature of the 'strings' that are learnt in animal songs (Merker & Okanoya, 2007, p.406). Animal songs are distinct from animal calls in that the latter are largely innate and fixed in their species specific signalling system. Animal song, on the other hand, supplies "the raw materials for a learned (initially arbitrary) system of meaning assignments" (ibid, p.405). The impetus for finding signals that display arbitrary and combinatorial properties is the desire to overcome the problem of initializing a system of meaning-signal pairs – a problem we have encountered previously.

Merker & Okanoya (2007, p.407) highlight the vocal aspect of song and in particular the capacity for vocal learning – "the learned duplication of an auditory model by voice" – a behaviour that is rare among mammals. The authors hint at a possible link between hominid brain expansion and vocal learning, citing evidence for this link in birds: evidence suggesting a positive correlation between the telencephalic-to-brain ratio and the presence of singing and vocal learning (p.408). Across species, animal song, achieved through a neurally supported system of vocal learning has been a successful evolutionary solution for communication and display. Found in birds, humans, whales and seals (amongst others) the comparative analysis reveals members of disparate taxa with analogous traits. This makes a case for a musical stage in language evolution more plausible. If evolution converged on this solution before, there is no reason to assume it couldn't again. This is especially relevant when the ubiquity of music in human culture is considered.

In the context of each species' unique ecological and social niche vocal learning, while functioning as per its adaptation, will undergo different pressures in disparate contexts. This informs Merker & Okanoya's (2007, p.410) thought experiment. They imagine a scenario where an "elaborate learned song" is in place in a mammal without restrictions on seasonal breeding. Thus, song, in the service of display for the purpose of mate attraction and territoriality, and also not tied to any fixed pattern of breeding "would be spread out over the full range of the groups life circumstances, initially in haphazard fashion" (p.411). What they propose next is that learners in this population would be exposed to initially meaningless (in the sense of linguistic meaning) song-strings in different situations and would, in the course of learning, associate these

song-strings with a particular context. Strings would then be stored in memory together with the context. Iterations of this scenario over generations through the "learner bottleneck[2]…would ensure a gradual assortative segregation of song-strings by context" resulting in a repertoire of songs that is statistically shaped into a set of correlated song contexts (2007, p.411). This picture – not unlike Jespersen's one above – is tentatively offered as a possible explanation for the emergence of meaning-signal pairs which "is simply the context in which it is habitually sung" (ibid). A potential problem with this is how individuals come to have the same meaning associated with the same song. Although the authors don't discuss the context in which learning occurs, one would assume that to tie a song to a context many individuals would have to experience both song-string and context together in order for each one to both produce a song given the stimulus of a context and to evoke a context in the mind of the hearer.

This aside, the scenario is of a musical holistic protolanguage that progresses to modern language. How that happens will be addressed in Part 4. For now we will continue to look at how current research has interpreted the musical protolanguage hypothesis, first scientifically postulated by Darwin almost 140 years ago.

3.2.3 MITHEN, WRAY AND HOLISTIC PROTOLANGUAGE

The musical protolanguage theory, like the holistic counterpart, seems to have been resurrected at a time when greater understanding of human cognition can ground previous theoretical questions in convergent data. The idea of a holistic protolanguage has persisted and recently been supported by a number of researchers. The same cannot be said, as Mithen (2009) points out, of musical theories of language evolution. This has to do with a hitherto singular focus on the core objects of linguistics – words and grammar. This has, for Mithen, marginalised the potential discussion of the emotional lives of our pre-linguistic ancestors, which he sees as relevant based on evidence from infant directed speech and primate social interaction (2009, pp.59-60). The key is to view music, in evolutionary terms, as an emotional and affective strategy. With reference to great apes Mithen feels that "their social

[2] The learning bottleneck refers to the fact that in language transmission only a subset of the language can ever be heard by an individual. Therefore general structural principles will tend to survive. In the above context only a subset of the song-strings possible will be heard.

behavior cannot be understood without reference to their emotional states and how these are manipulated by vocalizations and physical actions" (2009: p.60). This is an appeal to behavioural continuity with an ancestral, pre-hominid primate species, as is his view that (in line with Wray, 1998) primate vocalizations are holistic "like human musical phrases" (2009, p.63).

On an ontogenetic scale the prevalence of infant directed speech in all cultures suggests that as well as helping an infant acquire their native language through prosodic cues which aid word segmentation and perception of phonetic categories (Thiessen et al, 2005; & Werker et al, 2007); there may be an explanation that highlights IDS as functioning "in terms of its emotional impact on the infant" (Mithen, 2009, p.59). That is, the musicality of IDS may be a factor in strengthening the infant-caregiver bond. Given that Mithen supports the holistic protolanguage hypothesis (2005, ch.9 & 15) why does he further assert its musical nature? The following are the six reasons he gives for "a degree of musicality" in his holistic protolanguage:

i. Anatomical evidence from the archaeological record shows the hominid line to have been able to make a wide variety of sounds (vocal tract anatomy); and a variety of gestures, including dance (bipedalism).
ii. The ubiquity of singing and dancing in human culture serving social bonding purposes.
iii. A similar prevalence of emotional expression and induction in others "by using vocalizations with particular pitch sequences and rhythms, and their equivalence in gesture and body movement".
iv. The increased period of neotony in humans, as a result of bipedalism narrowing the birth canal, led to increased infant dependency. IDS is a 'fossil' of a musical strategy used in looking after infants.
v. In line with Darwin's theory of sexual selection singing was initially a display for mate attraction; the increased vocal skill was then exapted for protolanguage.
vi. Evidence from contemporary traditional societies of mimicry of the natural world. (adapted from Mithen, 2009)

Mithen's work to date (2005 & 2009) has been to synthesize a number of ideas from different fields of inquiry into a unified theory of language evolution. This is his "Hmmmmm" model of protolanguage: Holistic, manipulative, multi-modal, musical, and mimetic.

Steven Brown's (2000) "Musilanguage" model is similar in many respects to Mithen's and also shares with Fitch an appreciation of "phonological syntax" (Brown, 2000, p.294). Along with Mithen, Brown (2000, p.278) sees in music an emotional function, what he calls "sound emotion", where "particular sound patterns…are used to convey emotional meaning". Both authors also share the view that language and music developed out of a musical communication system that was a precursor to modern language and music (Brown, 2000, p.277; Mithen, 2005, ch.9). However, whilst exploiting the combinatorial properties of phonology to support later syntax, it is not clear where meaning might have come from in Brown's model. In the growing literature that compares language and music (see Fitch, 2005 & 2007; see also Besson & Schön, 2003, for a cognitive perspective), meaning is seen as a fundamental difference between the two domains. Nevertheless, Brown seems to view language and music "as fitting along a spectrum instead of occupying two discrete…universes" (2000, p.278). At one end of this spectrum is what Brown calls "sound reference" which is the semantics of language: reference and symbolic meaning. At the other end we find "sound emotion" where emotional meaning is conveyed by "particular sound patterns" (ibid). Thus, meaning is assumed in the domain of music along a spectrum that represents degrees of difference where others have seen musical and language meaning as fundamentally distinct (Fitch, 2005, p.31).

Botha (2009) offers a sustained critique of both Mithen and Brown's models. The underlying objection is a lack of clarity concerning linguistic theory. In particular Botha (2009, pp.66-67) cites Brown's indeterminate definition of language as characteristic of the underlying faults with his model. In short Brown's model is "based on various untenable assumptions about what syntax is, about what semantics is, and about how syntax and semantics are related to each other and to phonology" (p.71). Similarly, Mithen's theory fails to account and fully explain the segmentation process and the emergence of syntax (although see section 4), and is thus rejected as a possible explanation for a linked evolutionary history between language and music.

Having briefly looked at musical theories of protolanguage we now return to the work of Alison Wray.

3.3 ALISON WRAY

As we have already observed, Wray (1998 & 2000) advocates a continuous development in the hominid line from an earlier, essentially primate system of communication. Based on observations of modern primates, the earlier, pre-hominid, system is taken to be characterized by "noise and gesture". The "noise" is phonologically unstructured and lacks the physiological and articulatory control required to produce the vocal gestures of modern human speech. The gestures are functional, most often involved in dyadic exchanges to facilitate relationships, sex, play, grooming etc. Out of this, presumably via ecological pressures, hominid communication developed, which was similar to the prior system "but capable of a more complex inventory of functional exchanges and thus needing a larger set of utterances" (Wray, 1998, p.51). The noises and gestures of hominid ancestors would, of necessity, gradually change the vocal tract to "make more sounds available, so an increasing number of discrete utterances could be created" (ibid). Thus, in this way, a basic phonemic inventory is established as vocal articulation becomes more refined in the service of producing auditorially distinct utterances.

However, one of Maggie Tallerman's main contentions is that Wray (1998) assumes the "prior existence of discrete segments" that holistic utterances are composed of (2007, p.585); a complaint she also levels at computational models that explore the emergence of compositionality (Kirby, 2000). These segments are at the syllabic and phonetic levels of speech production: syllabic segments are recombinable at the utterance level to produce further distinct utterances, and phonemes can be recombined to produce the syllabic building blocks of utterances. The problem is, for Tallerman, that in the segmentation process that Wray hypothesizes there is a presupposition of distinct phonetic elements that go hand-in-hand with words (in the modern sense):

> you can't have morphemes without phonemes (since morphemes are composed of phonemes) and you can't have phonemes without words, since you have to have semantic contrasts and minimal/near minimal pairs in order to know what the phonemes are. (Tallerman, 2007, p.587).

Wray's (1998 & 2000) theory does seem to gloss over an explanation as to how vocal control might have emerged in the hominid line. Especially as the fine, intentional

articulatory control needed to produce holistic utterances Wray envisages are generally not available to primates. However, Tallerman's, suggestion that this makes a holistic theory of protolanguage untenable can be countered by the fact that Wray's failure to adequately account for a certain aspect of the evolutionary picture merely requires a deeper consideration of these issues. The absence of an explanation concerning vocal control highlights a gap in the model rather invalidating it altogether.

A related point can be made with reference to computational models that build in certain elements at the start. Language evolution is a complex interaction of processes involving, amongst other things, genetic preadaptations, natural selection, ecological pressures, and cultural transmission. A computer model seeks to isolate explananda by simplifying the problem – too many variables will render the problem too complex to model. In the case of Kiby (2000), to investigate the emergence of linguistic structure as well as the development of a modern human vocal tract, and the concurrent development of cognitive structures guided by natural selection would be far too abstruse an undertaking. Tallerman's issues are certainly valid but not insurmountable.

So now we come to the actual utterances that speakers of Wray's protolanguage would use, a small sample of which is represented in table 1.1.

1	/mabu/	→	'keep away'
2	/madu/	→	'take the stick'
3	/mɛbita/	→	'give her the food'
4	/ikatubɛ/	→	'give me the food'
5	/kamɛti/	→	'give her the stone'

Table 1. Characteristic protolanguage utterances with their glosses on the right. Adapted from Wray (1998).

From the above examples we can note a number of aspects. Firstly, there is no internal structure to the utterances. Examples 3-5 are glossed with the verb 'give' but no element is common to all three utterances that could stand for 'give'. Thus they are holistic. Secondly, we can note a simple yet productive phonemic inventory that

displays wide range of distinct sounds – from bilabials (both nasal and plosive) at the front of the vocal cavity (/m/ and /b/) to the voiceless, velar stop at the back (/k/). The vowel sounds also occupy the, more or less, maximally distinct cardinal points of the vowel space – high back and front (/u/ and /i/); front mid and low (/ɛ/ and /a/). As well as these phonetic features there is an ability to combine meaningless but distinct syllabic elements – for example: /ma/, /ka/ and /mɛ/. Thirdly, the propositional content of the utterances "perform the same functions as in primates, expressing and negotiating relationships, and inviting specific actions from others" (Wray, 1998, p.51).

Tallerman (2007) has outlined the problems she sees with Wray's account and by extension with holistic models of protolanguage in general. Tallerman's own view is in line with Bickerton's: "a word-based lexicon evolved by building on ancient conceptual categories which are likely shared by many primates" (2009, p.181).

Although this model is a hypothetical bridge from a previous, animal communication system to, ultimately, modern human language it is not the whole protolanguage story. Wray (2002, p120) identifies "two inherent limitations on a holistic system, when operating alone" which constrain "the quality and type of message that can be sustained". The first limitation arises from the interaction of two parameters which impact on the number of messages the protolanguage could support: the speaker and hearer's ability to produce and perceive the holistic utterances; and the utterances demands upon long-term and working memory. Using a simplified phonetic inventory at the inception of protolanguage (which I take to be implicit in Wray's model due to the incremental nature of vocal tract development she proposes), will constrain the number of distinctly compact messages in an individuals inventory. If the 'proto-lexicon' increases then the distinctiveness of the utterances will increasingly be along the length parameter of utterances, which will affect working memory.

The second limitation relates to how frequently an utterance can be used without it disappearing during the vertical, bottleneck stage of language transmission. When learning a language, an infant is only ever exposed to a subset of the languages total utterances (Kirby, 2000), for modern human language with its infinite generative power it is logical that no child will ever hear all possible utterances of his or her language. For a holistic protolanguage, however, the issue is not so clear cut. The

relationship between an utterance and its meaning is not a function of its subparts and is therefore unique in its place in the lexicon of similar utterances. If this lexicon is small enough then it may well be that the infant does hear all the utterances of the protolanguage and transmission is a matter of learning a static list of items. However, the historical process is not perfect and over time replication of sounds changes, thus changing the utterances. Additionally, it is taken as read that the dynamic environment of hominids would yield new situations that were relevant enough to warrant an increase in the proto-lexicon and so the situation would eventually arise where the infant no longer hears the totality of the protolanguage.

This has repercussions on the holistic protolanguage as a whole if the "balance between specificity [of a message] and resultant frequency" determines whether or not the message is retained in the protolanguage inventory (Wray, 2002, p.121). To balance this fine line it is suggested that "generic denotation" might guarantee an utterances fixture in the protolanguage. However, Wray goes on to reject this on the basis that "since the messages would determine the behaviour of the group, this enforced reductionism would restrict the range of pre-planned actions, encouraging habit over innovation" (ibid). Generic denotation creates ambiguity in the system due to a lack of reference between message and object (122). The holistic utterance (see table 1) bundles candidate referents into one non-decomposable unit making specificity problematic, especially with regard to displaced referents.

Thus, Wray is in a paradoxical position: without individual denotation a holistic protolanguage is denied potentially useful utterances of the type "Fetch X" where "the message is pragmatically associated with an absent referent" (ibid). However, if there were individual denotation this would place a huge burden on the memory load of the users as the list of potentially absent people or things that could be fetched increased, rendering the system untenable.

Wray's solution (2002, 122-124) is to propose an initial holistic system as outlined above which is subsequently augmented by the appearance of individual denotation. This offers some rapprochement with Bickerton's model as it introduces alongside the holistic system of communication a simple means of referring to specifics. Similarly, Wray does not rule out the possibility of a prior stage of manual communication so that her model can also encompass parts of Arbib's.

3.4 MICHAEL ARBIB

Arbib (2005b) outlines a detailed series of evolutionary stages, which show the development of the underlying neural structures he believes to be involved in language evolution. Initially these mirror neurons were proposed to be present in the common ancestor of monkey and human. Later the common ancestor of humans and chimpanzee is thought to have a more developed mirror system that allowed simple imitation of grasping. After the divergence of the hominid line of great apes this mirror system, further modified, allowed for greater complexity of imitation which eventually led to the appearance of what, for the purposes of this essay, I will call protolanguage.

In Arbib's scheme the protolanguage stage is actually a number of important sub-stages that, over their emergent trajectory, switch from the visual, manual-gestural modality to the auditory, vocal modality. Development can be broken down as follows:

Complex imitation → Protosign → Protospeech

These are all stages occurring in the hominid line constituting Arbib's (2005b) stages 4, 5, and 6 respectively. The transition from complex imitation to protosign is further broken down in terms of the abilities needed to support protosign. These are the ability "to engage in pantomime and…make conventional gestures to disambiguate pantomime" (2005b, p.115). Pantomime as the initial communicative act towards language has an advantage of being iconic in that the pantomime shows a similarity with the object or action depicted. The advantage lies in the perceptual similarities between the object and the key features acted out. According to Arbib (ibid) this "provides open-ended communication that works without prior instruction or convention". However, it is just such conventionalization that is needed to 'bridge over' into protosign. This is achieved through gestures that disambiguate the bundle of different meanings that are potentially associated with pantomime. For example, if one were to pantomime 'bird-flight' they might use their arms in a flapping motion or just their hands; either way there is ambiguity in what is being highlighted: the bird or flight in general? What Arbib has proposed (2005b, p.116, 2009b) is that a class of disambiguating gestures were invented to make clear the intention of the pantomime. Thus, a new class of hand movements arose that were less iconic and approached the

quality of arbitrary, symbolic communication. The trade-off between easily recognised pantomimes and non-iconic, conventionalized gestures is "more rapid communication with less neural effort" (2005b, p.116). This brings us to the first stage proper of Arbib's protolanguage.

In getting to protolanguage Arbib has gone to great lengths to flesh out the neural structures underlying the core features that are crucial for the emergence of the "language ready brain". His hypothesis is that development from the last common ancestor of chimpanzee and Homo sapiens was along a path that altered a system of neurons essential to establishing shared meanings based on initially iconic pantomime through a conventionalizing process to a bipartite protolanguage; one which was firstly gestural (manual and oro-facial), and then subsequently vocal. We have seen that the path to protosign was part-neural, part-cultural, evolving out of the already extant gestural system of previous primate communication. Why then, if we know from contemporary studies in sign language that it is clearly language in all its manifest complexity, should speech (and not gesture) come to be the dominant modality – indeed the default – for modern human language?

The answer, for Arbib, lies in the fact that, of the three stages above, protosign and protospeech evolved in tandem. The relationship between protosign and protospeech in his view is one of "scaffolding" (Arbib, 2005a, p.148). Protosign involves manual and facial gestures; protospeech primarily involves vocal gestures which emerged via the scaffolding effect of the evolutionarily quondam protosign. The interaction proposed by Arbib occurred during the period of hominid evolution covering *Homo habilis* to Homo sapiens (roughly 1.75-0.5 m.y.a.). During this time

> biological and cultural evolution along the hominid line saw advances in both protosign and protospeech feeding of each other in an expanding spiral so that…protosign did not attain the status of a full language prior to the emergence of early forms of protospeech (Arbib, 2005a, p.148)

it is at this protospeech stage that Arbib proposes early Homo sapiens to be using holistic utterances. In Part 4 we will address the hypothesis, outlined in Wray (1998) that the structural complexity of modern compositional language emerged through cultural processes.

PART 4: PROPOSALS FOR THE ROUTE OUT OF HOLISTIC PROTOLANGUAGE AND INTO MODERN LANGUAGE

4.1 INTRODUCTION

In section 2.3 we briefly encountered the principle of uniformity (uniformitarianism). Newmeyer (2002), in his critical survey defines it thus: "there is no overall directionality to language change…human languages have always been pretty much the same in terms of the typological distribution of the elements that compose them" (p.360). However, there is mounting opinion that this statement may be more problematic than it appears. As an antidote to the idea that a group of language speakers can be seen as more or less primitive based on the structure of their language it is commendable. And it also takes into consideration our underlying biological ability to acquire any language as our first. Uniformitarianism in this guise deals with language as used by contemporary *Homo sapiens*. Can the uniformity principle be carried over into the study of language evolution? Growing research is pointing towards a situation where it seems to be unfeasible to hold to certain views on language evolution whilst maintaining a broad uniformitarianist stance (Newmeyer, 2002).

In the following section we will sum up the earlier views of how protolanguage might have developed into modern language. Then, after discussing current thinking on this question (Wray and Arbib) we will look at research in the field of computational modelling that provides support for the emergence of compositional structure from initially holistic utterances. These hypotheses are implicitly non-uniformitarian.

4.1.1 BENTHAM, DARWIN & JESPERSEN

As has become apparent over the course of this review the idea of a holistic protolanguage has persisted and recently flourished, as the weight of evidence that has been brought to bear on the question has increased over time. In general, contemporary language evolution research is multi-disciplinary, technologically state of the art, and in constant dialogue with the interested parties (there are now numerous annual conferences devoted to the question). This difference means that in

some areas of the literature I am addressing, current work has far more to say than earlier thinkers could.

Jeremy Bentham took great pains to describe the underlying logic of language and applied this to his theory of language evolution. However, his conception of how protolanguage might have developed into the system of words and grammar used in modern language was less fleshed out. Of the 'route to language' Bentham has this to say:

> To form the words of which language is at present composed has been the work of analysis. The original sentences were, as it were, broken down into words, these words into syllables, and these syllables, by the help of written and visible signs, into letters. (Bentham, 1843, p.323)

Burnet (1773, pp. 395-397) also held the view of a holistic protolanguage but thought that words of the atomistic, modern kind arose alongside the previous system through the purposeful invention by man. Bentham also saw the inventive hand of man at work, but it was applied to protolanguage in use; words arose through "the work of abstraction, the produce of a refined analysis" upon the holistic protolanguage (Bentham, 1843, p.321). This is how modern language came to be, but Bentham still held the view that "in ordinary discourse propositions came entire, it is only on the occasion of some science or art, that…any term is presented by itself" (ibid). I read this as meaning that our day-to-day communication is pragmatically facilitated by propositions and that linguistic compositionality represents a separate tier of analysis that can be applied to language. This is reminiscent of Wray's (1998) conception of formulaic language where "a significant proportion of our day to day utterances are not generated by rule but retrieved whole from the store" (p.58).

Moving the story on, Otto Jespersen (see above, pp.27-28), not unlike Merker and Okanoya (2007) discussed in section 3.2.1; saw shared meanings initializing with the co-occurrence of songs and specific contexts. Jespersen also proposed that this happens through the process of the analysis of existing utterances into constituent elements. His evidence for this claim comes from the study of grammaticalisation as observed by the historical linguists of the previous century. In particular Jespersen highlights the phenomenon of *secretion*: "one portion of an indivisible word comes to acquire a grammatical signification which it had not at first" (Jespersen, 1922, p.384).

Although Jespersen admits that from "the nature of the subject it is impossible to give more than hints" (ibid, p.440); he is clear in seeing the development of language as the action of the historical process working on the earlier melodic utterances. This theoretical position, then, is clearly non-uniformitarian in an evolutionary context. Protolanguage is by definition different from modern language and the transition between the two is a product of recognised historical processes acting on each generation. Also implied, if we invoke the historical process, is a certain amount of directionality to language change. Certain sound changes will only go in one direction, hence, we are able to reconstruct earlier languages (Campbell, 1998, p.115); and grammaticalisation changes, such as main verb > auxiliary, are similarly seen to be largely unidirectional (ibid, p.239). In the following section it will be clear that there is an underlying assumption that protolanguage developed into modern language *after* the appearance of *Homo sapiens*.

4.2 THE SEGMENTATION ROUTE

The route from a holistic protolanguage to modern human language has been proposed by Wray (1998, p.54) to be a purely cultural process. Moreover, Wray proposes that the first hominids to analyse the holistic protolanguage "used a brain that was fundamentally set up in the same way as ours today" (ibid). This final, gradual development into modern language is supported by Arbib (2005b, p.119), Mithen (2005) and Fitch (in press) – although Fitch (Ch.14.10, para.4) stresses a better fit with a musical model. As such I will focus on Wray's position, supplemented by differing opinions within a similar framework.

In her paper (1998, pp.55-58) she describes this route as segmentation. Bearing in mind Table 1 above (p.34) the process of segmentation is as follows. A protolinguistic community has an inventory of holistic utterances. For each, the meaning of the whole is *not* a function of the meaning of the parts. However, as we have seen the utterances are composed of a finite set of syllabic units (in the examples these have the form CV). In order to generate utterances to cover the range of expressible meanings needed in the community, it follows that the syllabic elements will be used repeatedly in different contexts. This can be seen from the items /mɛbita/ - 'give her the food'; /katubɛ/ - 'give me the food'; /kamɛti/ - 'give her the stone'. It

is a coincidence that /mɛ/ occurs in two different utterances that have a recipient equivalent to 'her', and doesn't occur in the utterance without 'her'. Wray proposes that the analyser (i.e. the hearer) creates a morpheme boundary around /mɛ/ as a result of 'looking' for a phonetic representation of Bickerton's (1998) theta role. This is a prelinguistic hominid neural component involved in "conceptual structure" and social intelligence (1998, p.346). What Bickerton suggests is that at some point a link was established between perception of events involving Agents, Patients and Recipients and an area of the brain that represented "phonetic shapes of words" (ibid).

Returning to Wray, however, notice that in the same three examples 'give' occurs in all three utterances but without the coincidence of a common morpheme. The closest is /ka/ but then /mɛbita/ is a counter-example. Wray suggests three options that might work. The community could just abandon the hypothesis that /ka/ stands for 'give'; hypercorrection might occur so that /mɛbita/ is reanalysed as /mɛbika/; or hypercorrection could occur in the opposite direction: /mɛbita/ is dropped from the 'lexicon', eliminating the counter-example 'give her the food'.

The pressure to analyse utterances into their components and assign meaning presumably comes from the increasing memory load that a growing vocabulary of holistic utterances would create. Whilst recognising this as a possible "catalyst" Wray also points out that other hypotheses are compatible within a holistic framework (1998, p.54). Evidence that this might have been a plausible route comes from first language acquisition studies. Wray notes that modern human infants segment utterances "when the child notices the points at which otherwise identical utterances tolerate paradigmatic variation, and inserts morpheme boundaries." Studies of the primate auditory system suggests that the ability to discriminate the sounds that language is composed of are not specific to humans and further, that a similar ability with reference to perceiving segments further up the phonological hierarchy might be a widespread primate adaptation (Ramus et al, 2000; Hauser et al, 2001). This ability may have further developed in the hominid line and become a preadaptation that was to serve the later protolanguage.

This is merely a starting point towards human language. Development is proposed to be gradual, possibly happening over the period that Homo erectus was extant ca. 1.5 – 0.5 m.y.a. However, it's not clear whether Wray is saying that *Homo*

erectus' brain structure is "fundamentally" the same as modern *Homo sapiens*. Additionally, Wray draws on observations that children aren't hindered by an incomplete analysis of language and in fact use holistic utterances in their productions. She notes that although "the modern child's input is rule-determined…the child does not begin by treating it as such"; the child takes input as complete holistic meanings.

This theory of segmentation is largely supported by the interpretation of current linguistic evidence – language acquisition, phonological analysis, syntactic analysis vs. formulaic language. Recently, however, evidence that supports Wray's view has come from a field of research to which we will now turn.

4.3 SUPPORT FROM COMPUTATIONAL MODELLING

Computational modelling has provided a powerful tool for exploring language evolution. It has helped isolate key parameters in the evolution and emergence of language and language structure. The reason for distinguishing between language and structure results from recent suggestions and support from models that highlight the interaction of different "complex adaptive systems" related to language evolution (Kirby & Hurford, 2001, p.122). Kirby and Hurford (ibid) identify three such systems involved in language: learning (ontogeny); cultural evolution (glossogeny); biological evolution (phylogeny).

Another factor in the importance of computational models has been the way that such systems can be left to iterate over many 'generations' in simulations that provide valuable insights into possible trajectories that language evolution may have taken, given a suitably simplified model of the variables involved (populations, meanings etc.). The simulations have helped to generate guiding principles in approaching the question: iterated learning, learning bottlenecks, learning biases etc. These have arisen purely from a consideration of implementational problems in modelling language evolution. The outcome is that for holistic theories of protolanguage there is hitherto unavailable data that at least suggests a possibility for the emergence of compositionality out of an earlier holistic communication system over a glossogenetic timescale (Kirby, 2000; Kirby, Cornish & Smith, 2008).

Kirby (2000, p.317) describes a computational model the results of which show "the emergence from randomness of simple yet language-like syntax in a population that is not constrained to learn only a compositional language". This experiment explores the cultural development of a simplified language-like behaviour as it is passed down successive generations. However, to get the ball rolling agents are initially linguistic-free but with an internal representation system that, given an input, will model it as a "rewrite" rule that expresses the simplest type of context free, non-compositional grammar:

S/<Agent = Zoltan, Patient = Mary, Predicate = Knows> → 'zoltanknowsmary'

This states that the input signal on the right is represented by the rule on the left, but because the initial agents are 'blank slates', they have no way of internalizing the composite features of the utterance and thus their earlier meanings are holistic chunks. In order for meanings and signals to arise without introducing them from 'outside', the agents in a population at first randomly invent strings to represent meanings. This ensures that the initial language is random.

Key features of this model are the presence of observational learning; a model of population turnover; the absence of selection based on communicative success (i.e. random); and, as we have seen, the population is non-linguistic at the start – emergent biases should be "purely a product of the learners and the population model" (Kirby, 2000, p.305).

In all of the simulation runs in this experiment a series of three stages of behaviour could be observed. In the first stage, after a slow initialization period, random invention plus random noise leads to some stability. Agents have small grammars representing a limited portion of the possible meanings expressible. In this stage there is no structure to strings as related to meaning. In the second stage both grammar size and the number of meanings expressible increase. During this stage, however, the number of meanings outstrips the number of rules in an agent's grammar – that is, the rules became less tied to individual meanings and become more general, consequently decreasing in number. Initially any meaning would have had a separate rewrite rule and thus an agent's representations would consist of a store of rewrite rules paired with corresponding strings.

Another important change associated with this stage is that the grammars become more complex and partially compositional with segments in the strings representing aspects of the meaning. For example Kirby (2000, p.314) notes in one run of the simulation that in the string "dceddd" meaning "S/<Agent = John, Patient = Tünde, Predicate = Knows>" the substring "ce" represents John.

The third stage develops from these partially compositional grammars to stable, almost completely, compositional grammars. This is marked by a further increase (in some cases total) in the number of meanings expressible and a decrease in the size of an agent's grammar.

The conclusion to be taken from this study seems to be that generalisation, due to the learning bottleneck, drives the emergence of compositionality. Although learning a general rule will not be as easy as learning the one-to-one mapping of a meaning-signal pair, given that in the experiment there were 100 meanings, "the probability of acquiring a particular rule given a random sample of meanings increases with the generality of that rule" (Kirby, 2000, p.319).

Fitch (in press), Merker & Okanoya (2007) and Wray (1998) have all cited this work as supporting their individual holistic frameworks. As the model also makes no claims about the possible nature of the signal, other than within the confines of the model itself, it fits into a musical model of protolanguage.

4.4 ARTIFICIAL LANGUAGE LEARNING EXPERIMENTS

Another paradigm exploring the cultural evolution of language is Artificial Language Learning experiments (ALL). Pinker & Bloom (1990) have suggested that features of syntax are not only adaptive – they confer fitness on the individual – but that they are evolvable through natural selection. In contrast, models of the cultural evolution of language using ALL experiments suggest that languages may themselves adapt to the constraints of the human brain (Christiansen & Ellefson, 2002, Kirby, Cornish & Smith, 2008). In this view a language inhabits a niche: the brain. Our brains, like any other brain, have limitations in what they can do. Therefore certain features of language, such as the subjacency principle for example, which seem tailor-made to

their function, might equally be explained as an outcome of more general learning and memory constraints (Christiansen & Ellefson, 2002, pp.349-351).

The results of Kirby, Cornish & Smith (2008) suggest that this scenario may be feasible. Their study is an ALL experiment that involves a diffusion chain system of communication transmission. Participants learn some behaviour from an experimenter; they in turn exhibit this behaviour for the next participant, and so on. The behaviour in question is an 'alien' language which is a set of picture-string pairs. Initially the strings are random sequences of letters assigned to a picture that is a combination of three parameters that can vary in three ways: shape (square, circle or triangle); colour (red, blue or black); and movement (horizontal, spiral or bouncing). Thus a picture could be a red bouncing triangle. Initially the labels attached to the pictures in this language are "generated and assigned randomly, and the first participant in the experiment is trained on this random language" (Kirby et al, 2008, p.10682). Because the strings are random to begin with the initial relationship of string to picture is holistic. For example a black square bouncing is paired with the string 'kihemiwi' (ibid). Like Wray's examples (see table 1) no sub-part of the string refers to any feature of the picture, rather the whole is represented. This means that in order to learn the tokens in this language they have to be rote learned. However, over time there emerged a generalising principle which enabled participants label previously unseen pictures. This is due to systematic underspecification in the language. As items were held back from the training data it was impossible to learn each and every picture-string pair, this resulted in generalization.

The second experiment in this paper involved a modification to the above set-up. In effect homonymy was removed from the language to be learned: "If any strings were assigned to more than 1 meaning, all but 1 of those meanings (chosen at random) was removed from the training data" (ibid, p.10684). The results in subsequent trials showed the emergence of a structured language: "the string associated with a picture consists of substrings expressing color, shape, and motion" (ibid). The absence of ambiguity in the language appears to be important in the emergence of compositional structure.

In part, the motivation for this study was scepticism "as to how well computational models of learning match the abilities and biases of real human learners" (p.10681). Placing the artificial language in the domain of human learning goes some way to redressing this criticism. It also provides a good fit to the general

model of segmentation presented in various holistic protolanguage scenarios in that these are largely hypothesised to have arisen in the minds of anatomically modern humans.

PART 5: CONCLUSION

Protolanguages have been proposed as stages in the evolution of language in order to account for the seemingly insurmountable gap between animal communication and language. Success in how well this gap is bridged depends on our understanding of the problem. A single mutational event in our past that resulted in the linguistic ability of our species is unlikely. But, in contrast, it is hard to imagine the incremental evolution of the complex structure we see in language. Whether protolanguages are viewed as either synthetic or holistic they are a valuable heuristic for investigating how language came to be the way it is. Although the idea of a holistic protolanguage has persisted much longer than people suspect the new paradigm for exploring the question is of an increasing rapprochement between disciplines. Take for instance Hurford's (2003) study that offers a link between the seemingly disparate studies of logic, neuroscience and linguistics. However, we are still a long way from reaching a widely accepted picture of language origins – as the holistic debate can attest. With this in mind then, are we in a better position than Bentham et al. to characterise with any certitude earlier stages of language phylogeny?

I think the answer is an obvious yes; we are in a much more favourable position from which to study language origins. Considering the extensive literature and technology at our disposal in contrast to that of the 18th century it is remarkable that these early thinkers had the fortitude to embrace the question at all.

We have to bear in mind that research is an ongoing process and that as the theoretical landscape changes so to do the theories themselves. But what we must also recognise is that the past still has a lot to offer us. Wray's theory is illustrative of this. Here we can see the convergence of a number of ideas preceded by Bentham. The holistic nature of early language; the idea that this persists today in the form of Wray's formulae and Bentham's interjections; and the idea that holistic utterances were somehow broken down (segmented) and analysed into the words we are familiar with today. Fitch (2009) in his *Language Log* post acknowledges the earlier ideas of Darwin and Jespersen, whilst extending them to fit with what we have learned in the interim – about language, communication and evolution. Good ideas sometimes have to wait for science to catch up in order to test them. So as well as reconciling different scientific disciplines in addressing the question of language evolution, it seems to be worthwhile to engender rapprochement with ideas from the past.

REFERENCES

Aarsleff, H. (1982). *From Locke to Saussure: Essays on the Study of Language and Intellectual History*. London: Athlone.

Arbib, M. A. (2002). The mirror system, imitation, and the evolution of language. In K. Dautenhahn & C. L. Nehaniv (Eds.), *Imitation in Animals and Artifacts*, Cambridge, MA: The MIT Press.

Arbib, M. A. (2005a). Interweaving protosign and protospeech: Further developments beyond the mirror. *Interaction Studies*. 6:2, 145-171.

Arbib, M. A. (2005b). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences*, 28: 105-167.

Arbib, M. A. (2008). Primate vocalization, gesture, and the evolution of human language. *Current Anthropology*, 49: 6, 1053-1076.

Arbib, M. A. & Bickerton, D. (Eds.). (2008). Holophrasis vs compositionality in the emergence of protolanguage. *Interaction Studies*, 9:1. Amsterdam: John Benjamins Publishing Company.

Bentham, J. (1843). *The Works of Jeremy Bentham/ Published Under the Superintendence of his Executor, John Bowring* (Vol. VIII). Edinburgh.

Besson, M. & Schön, D. (2003). Comparison between language and music. In I. Peretz & R. Zatorre (Eds.), *The Cognitive Neuroscience of Music*. Oxford: Oxford University Press.

Bickerton, D. (1990). *Language and Species*. Chicago, IL: Chicago University Press.

Bickerton, D. (1995). *Language and Human Behaviour*. London: UCL Press Limited.

Bickerton, D. (2000). How protolanguage became language. In C. Knight, M. Studdert-Kennedy & J. R. Hurford (Eds.), *The Evolutionary Emergence of Language: Social Functions and the Origins of Linguistic Form*, pp.264-284. Cambridge: Cambridge University Press.

Bickerton, D. (2002). Foraging versus social intelligence in the evolution of protolanguage. In A. Wray (Ed.), *The Transition to Language*, pp. 207-225. Oxford: Oxford University Press.

Bickerton, D. (2009, February 15). Bickerton on Fitch. Message posted to http://languagelog.ldc.upenn.edu/nll/?p=1148

Botha, R. (2009). On musilanguage/''Hmmmmm" as an evolutionary precursor to language. *Language and Communication*. 29, 61-76.

Bowie, J. (2008). Compositional versus holistic theories of language evolution: An interdisciplinary and experimental evaluation. (Doctoral dissertation, The University of Reading, 2008).

Brown, Steven. (2000). The "Musilanguage" model of music evolution. In N. L. Wallin, B. Merker & S. Brown (Eds.), *The Origins of Music*, pp. 271-300. Cambridge, Mass.: The MIT Press.

Burnet, J. (1773-1792/1967) Of the Origin and Progress of Language (Vols. 1-6). England: Menston.

Campbell, L. (1998). *Historical Linguistics: An Introduction*. Edinburgh: Edinburgh University Press.

Christiansen, M. H. & Ellefson, M. R. (2002). Linguistic adaptation without linguistic constraints: The role of sequential learning in language evolution. In A. Wray (Ed.) *The Transition to Language*, pp. 335-358. Oxford: Oxford University Press.

Darwin, C. (1859). *On the origin of species.* London: John Murray.

Darwin, C. (1871). *The Descent of Man and Selection in Relation to Sex*. London: John Murray.

Dinwiddy, J. R. (1989). *Bentham*. Oxford: Oxford University Press.

Everett, D. L. (2005). Cultural constraints on grammar and cognition in Pirahã: Another look at the design features of human language. *Current Anthropology*. Vol. 46, No. 4: 621-646.

Fitch, W. Tecumseh. (2005). The evolution of music in comparative perspective. *Annals of the New York Academy of Sciences*. 1060, 29-46.

Fitch, W. Tecumseh. (2006). The biology and evolution of music: A comparative perspective. *Cognition*. 100: 173-215.

Fitch, W. Tecumseh. (2009, February 12). Musical protolanguage: Darwin's theory of language evolution revisited. Message posted to http://languagelog.ldc.upenn.edu/nll/?p=1136

Fitch, W. Tecumseh. (In press). *The Evolution of Language*. Cambridge: Cambridge University Press.

Fodor, J. & Lepore, E. (1992). *Holism: A Shopper's Guide*. Oxford: Blackwell.

Gentner, T. Q., Fenn, K. M., Margoliash, D. & Nusbaum, H. C. (2006). Recursive syntactic pattern learning by songbirds. *Nature,* 440 (7088), 1204-1207.

Hallett, G. L. (1967). *Wittgenstein's Definition of Meaning as Use*. New York: Fordham University Press.

Hauser, M. D., Newport, E. L., & Aslin, R. N. (2001). Segmentation of the speech stream in a non-human primate: statistical learning in cotton-top tamarins. *Cognition*. 78, B53-B64.

Heilbron, J. L. (Ed.). (2003). *The Oxford Companion to the History of Modern Science*. Oxford: Oxford University Press.

Honderich, T. (Ed.). (1995). *The Oxford Companion to Philosophy*. Oxford: Oxford University Press.

Hurford, J. R. (1998). Introduction: The emergence of syntax. In J. R. Hurford, M. Studdert-Kennedy & C Knight (Eds.), *Approaches to the Evolution of Language: Social and Cognitive Bases*, pp. 299-304. Cambridge: Cambridge University Press.

Hurford, J. R. (2003). The neural basis of predicate-argument structure. *Behavioral and Brain Sciences*. 26, 261-316.

Hurford, J. R. (2007a). The origin of noun phrases: Reference, truth and communication. *Lingua*. 117. 527-542.

Hurford, J. R. (2007b). The evolution of human communication and language. In P. D'Ettore & D. P. Hughes (Eds.), *Sociobiology of Communication: An Interdisciplinary Perspective*, pp. 249-264. Oxford: Oxford University Press.

Hurford, J. R. (2007c). *The Origins of Meaning*. Oxford: Oxford University Press.

Im Hof, U. (1994). *The Enlightenment: An Historical Introduction*. (W. E. Yuill, trans.). Oxford: Blackwell.

Jesperson, O. (1922). *Language: Its Nature, Development and Origin*. New York: W. W. Norton & Co.

Kirby, S. (2000). Syntax without natural selection: How compositionality emerges from vocabulary in a population of learners. In C. Knight, M. Studdert-Kennedy, & J.

R. Hurford (Eds.) *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form*, pp. 303-323. Cambridge: Cambridge University Press.

Kirby, S. & Hurford, J. R. (2001). The emergence of linguistic structure: An overview of the iterated learning model. In A. Cangelosi & D. Parisi (Eds.) *Simulating the Evolution of Language*, pp. 121-148. London: Springer Verlag.

Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: an experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences*. 105, 10681-10686.

Kors, A. C. (Ed.). (2003). *Encyclopaedia of the Enlightenment* (Vols. 1-4). Oxford: Oxford University Press.

McMahon, A. M. S. (1994). *Understanding Language Change*. Cambridge: Cambridge University Press.

Merker, B. & Okanoya, K. (2007). The natural history of human language: Bridging the gaps without magic. In C. Lyon, C. L. Nehaniv, & A. Cangelosi (Eds.), *Emergence of Communication and Language*, pp. 403-420. London: Springer.

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

Mithen, S. (2009). Holistic communication and the co-evolution of language and music: resurrecting an old idea. In R. Botha & C. Knight (Eds.), *The Prehistory of Language*, pp. 58-76. Oxford: Oxford University Press.

Newmeyer, F. J. (2002). Uniformitarian assumptions and language evolution research. In A. Wray (Ed.), *The Transition to Language*, pp.359-375. Oxford: Oxford University Press.

Newmeyer, F. J. (2003). What can the field of linguistics tell us about the origins of language? In M. H. Christiansen & S. Kirby (Eds.), *Language Evolution*, pp. 58-76. Oxford: Oxford University Press.

Ogden, C. K. (1932). *Bentham's Theory of Fictions*. London: Kegan Paul, Trench, Trubner & Co., Ltd.

O'Grady, W., Dobrovolsky, M., & Katamba, F. (Eds.). (1997). *Contemporary Linguistics: An Introduction*. London: Longman.

Pinker, S. & Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain sciences*. Vol. 13, 707-784.

Quine, W. V. (1960). *Word and Object*. Cambridge, Mass.: MIT Press.

Quine, W. V. (1966). Russell's ontological development. *The Journal of Philosophy*. Vol. 63, No. 21.

Rizzolatti, G. & Arbib, M. A. (1998). Language within our grasp. *Trends in Neurosciences*. Vol. 21, No. 5, 188-194.

Ramus, F., Hauser, M. D., Miller, C., Morris, D., & Mehler, J. (2000). Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science*. 288, 349-351.

Rousseau, J-J. & Herder, J. G. (1781/1772/1966). *On the origin of language*. (J. H. Moran & A. Gode, Trans.) Chicago: University of Chicago Press.

Smith, K. (2006). The protolanguage debate: bridging the gap. In A. Cangelosi, A. D. M. Smith, & K. Smith (Eds.), *The Evolution of Language: Proceedings of the 6th International Conference (EVOLANG6)*, pp. 315-322. London/Hackensack: World Scientific Publishing Co. Pte. Ltd.

Stam, J. H. (1976). *Inquiries Into the Origin of Language*. New York: Harper & Row.

Tallerman, M. (2007). Did our ancestors speak a holistic protolanguage? *Lingua*, 117: 579-604.

Tallerman, M. (2009). The origins of the lexicon: how a word-store evolved. In R. Botha & C. Knight (Eds.), *The Prehistory of Language*, pp.181-200. Oxford: Oxford University Press.

Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-Directed Speech Facilitates Word Segmentation. *Infancy*. Vol. 7(1), 53-71.

Thomas, D. A. (1995). *Music and the Origins of Language: Theories From the French Enlightenment*. Cambridge: Cambridge University Press.

Tomasello, M., Call, J., Warren, J., Frost, T., Carpenter, M., & Nagell, K. (1997). The ontogeny of chimpanzee gestural signals: A comparison across groups and generations. *Evolution of Communication. 1,* 223-253.

Wells, G. A. (1987). *The Origin of Language: Aspects of the Discussion from Condillac to Wundt*. La Salle III: Open Court.

Werker, J. F., Pons, F., Dietrich, C., Kajikawa, S., Fais, L. & Amano, S. (2007). Infant-directed speech supports phonetic category learning in English and Japanese. *Cognition*, 103, pp.147-162.

Wray, A. (1998). Protolanguage as a holistic system for social interaction. *Language & Communication,* 18: 47-67.

Wray, A. (2000). Holistic utterances in protolanguage: The link from primates to humans. In C. Knight, M. Studdert-Kennedy & J. R. Hurford (Eds.), *The Evolutionary Emergence of Language: Social function and the origins of linguistic form*, pp. 285-302. Cambridge: Cambridge University Press.

Wray, A. (2002). Dual processing in protolanguage: Performance without competence. In A. Wray (Ed.), *The Transition to Language*. Oxford: Oxford University Press.