



SEL3329: Dissertation in Linguistics and English Language

An evolutionary perspective on the implementation of memory in language processing and production

Aislinn Keogh

Supervisor: Maggie Tallerman

Word count: 9,928

I confirm that this piece of work contains no plagiarised material and that I have read and understood the section on Plagiarism in the School Style Guide.

Contents

Introduction.....	3
Research questions.....	3
1. Overview of the cognitive neuroscience of memory.....	3
2. Language origins and evolution.....	7
3. The Working Memory Model.....	9
4. Coolidge and Wynn: the evolution of working memory.....	11
5. Experimental evidence.....	12
6. Clinical evidence.....	18
7. Developmental evidence.....	22
8. Long-term memory and language.....	24
Conclusion.....	26
References.....	28
Appendix 1: glossary of acronyms.....	31
Appendix 2: index of examples.....	32

Introduction

Psycholinguistic research aims to understand the cognitive and neurobiological functions that allow humans to acquire, produce and process language. One area of interest for researchers is memory, both in terms of which elements of language are stored in long-term memory and how short-term or working memory is active in language production and perception. A related issue is how these memory capacities may have evolved alongside and selected for various aspects of language. This paper will discuss memory and language, providing an evolutionary perspective on the existing literature and a discussion of how enhanced memory capabilities may have impacted linguistic complexity in the evolutionary ancestors of *Homo sapiens*. Evidence will be taken from cognitive science, psycholinguistics, clinical research, child language acquisition, evolutionary anthropology and generative linguistics.

My three main research questions are provided below. Section one will outline some cognitive and neurological background in the field of memory research. Section two will examine some of the literature on language origins and evolution, detailing some of the phenomena that may be of interest to the study of memory in linguistics. Section three will attend to the Working Memory model (Baddeley and Hitch 1974) and section four will discuss current research into the evolution of working memory. Sections five to seven will outline some of the relevant experimental, clinical and developmental psycholinguistic studies in working memory which may shed light on the possible initial stages of language. Finally, section eight will discuss the involvement of long-term memory in language.

Research Questions

1. What can be learned about the evolution of grammar and the structure and lexical inventory of protolanguage from experimental, clinical and developmental psycholinguistic memory research?
2. In what ways would enhanced memory capabilities have been adaptive for language?
3. Could memory have constituted a selective pressure for any aspects of language?

1. Overview of the cognitive neuroscience of memory

Human memory can be divided into three basic levels: sensory (immediate), short-term and long-term memory. Sensory memory is the most volatile, retaining items for only a few seconds before they are either deleted or processed further by a more stable component of memory. Short-term memory (STM) refers both to the initial formation of a memory that will later become permanent and to the working memory (WM) component, that which

retains items for up to a few minutes after a stimulus while they are processed in attention. By the process of consolidation, some short-term memory items are made more durable over a period of days, eventually becoming permanent in the long-term memory (LTM). In a computer's memory, there are superficially similar levels: keystrokes are stored in a type-ahead buffer (immediate memory), then transferred to RAM (short-term memory), and can eventually be moved to the hard disk (long-term memory), the only location in which information survives after the power is turned off. All three levels assign one piece of information to one physical slot with an address, used to determine what exactly has been stored in each location. However, unlike the memory of a computer, human LTM does not consist of a finite number of available slots at which individual pieces of information can be stored, but seems to be infinite and distributed: each new piece of information that is learned appears to be stored in a number of places, overlapping with previous memories instead of being assigned to an empty slot (Calvin and Bickerton 2000).

Eichenbaum (2002: 2) suggests four crucial elements involved in the study of memory. *Connection* refers to “how memory is encoded within the dynamics [...] of connections between nerve cells”. *Cognition* refers to the nature of memories at a psychological level of analysis. *Compartmentalisation* refers to the localisation within the brain of different kinds of memory, and *Consolidation* refers to when and how memories become permanent. These components will now be discussed in turn, providing a multi-level description of memory in the brain.

1.1 Connection

Memory must have a cellular basis in the physical brain. The basic mechanism of memory encoding is the modification of functional connections between neurons, the cells of the brain and nervous system (Eichenbaum 2002: 29).

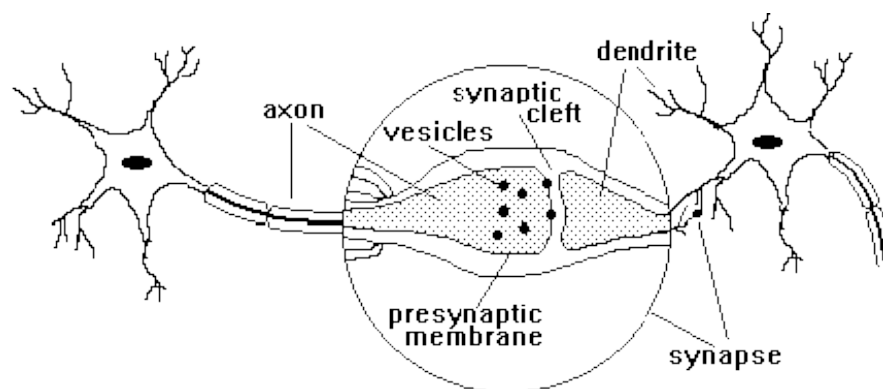


Figure 1: Neuron (Source: http://euclid.ii.metu.edu.tr/~ion526/demo/chapter1/section1.1/figure1_3e.html)

Figure 1 shows the structure of a neuronal connection, known as a synapse. The axon carries electrical signals away from the cell body to the presynaptic membrane, an enlargement of the axonal ending designed for neuronal transmission. The dendrites, or postsynaptic membrane, contain specialised receptors which receive neurotransmitters from the presynaptic cell. The connection shown in figure 1 is an axodendritic synapse, between the axon of one neuron and the dendrites of another. Neurons can also make axosomatic connections, between the cell body of one neuron and the axon of another, and axoaxonic connections, between the axons of two neurons.

STM involves local mechanisms within synapses, such as the phenomena of sensitisation – increased reaction to a certain stimulus – in which the strength of a synapse's signal is increased, causing an increase in reflex magnitude. Long-term memory is thought to require additional mechanisms: either the growth of existing synapses, or the addition of new ones by means of protein synthesis. For example, for sensitisation to a stimulus to become long-term, the number of synaptic terminals in sensory neurons must be greatly increased, resulting also in an increase in the number of dendrites to accommodate the extra electrical signals being passed between neurons (Eichenbaum 2002: 50). However, in both levels of memory, the basic cellular mechanism is an alteration in synaptic efficacy. It is important to note that, while certain brain structures are strongly associated with memory, the changes that mediate memory do not occur in specific 'memory cells'; neurons are highly adaptable and it takes relatively few activations to induce an adaptation for memory, and not many more activations for this to be fixed permanently (2002: 52).

1.2 Cognition

As well as the previously defined cellular basis, memory must presumably have a cognitive basis. Whilst learned reflexes do constitute some form of memory insofar as they do impose some changes on synapses, a cognitive level of description is all that allows us to distinguish between mere habit and true memory, the difference between the ability to repeat specific actions under the same conditions and independent representations or recollections of the past (Eichenbaum 2002: 17). *Declarative memory* and *procedural memory* are frequently discussed in the literature; declarative memory refers to the capacity to consciously remember everyday facts and events (cognitive memory), whereas procedural memory is the ability to perform certain cognitive or motor actions without explicit attention, such as tying one's shoes (non-cognitive memory). The instinctive tendency to consider this a 'real' distinction requires us to make use of a psychological level of description, despite the fact that, in a neurobiological description, it has no physical reality.

William James (1901: 648) argued that true memory is characterised by conscious retrieval of memories from a network of associated information: “it is the knowledge of an event, or fact, of which in the meantime we have not been thinking, with the additional consciousness that we have thought or experienced it before”. Bartlett (1932) also studied recollection, demonstrating that remembering is not a simple process of random recovery and loss, but that remembered items are predictable on the basis of their content and relation to the subjects’ other experiences. He therefore proposed that memory is reconstructive rather than simply reproductive, and that consciousness plays a central role in the reconstruction process.

1.3 Compartmentalisation

Research has shown that memory is not localised independently, and is instead tied to processing regions of the brain that mediate certain perceptual, motor and cognitive functions. Indeed, Eichenbaum (2002: 192) claims that memory and information processing are indistinguishable in the cortex. This paper will not be particularly concerned with the localisation of memories, suffice to say that, whilst the medial temporal lobe, specifically the hippocampus, is traditionally associated with declarative memory, it is important to be wary of viewing it as the single centre thereof. Rather, multiple systems interact and different memory systems (cognitive vs. non-cognitive) make use of different pathways and different functional brain regions, as shown in figure 2.

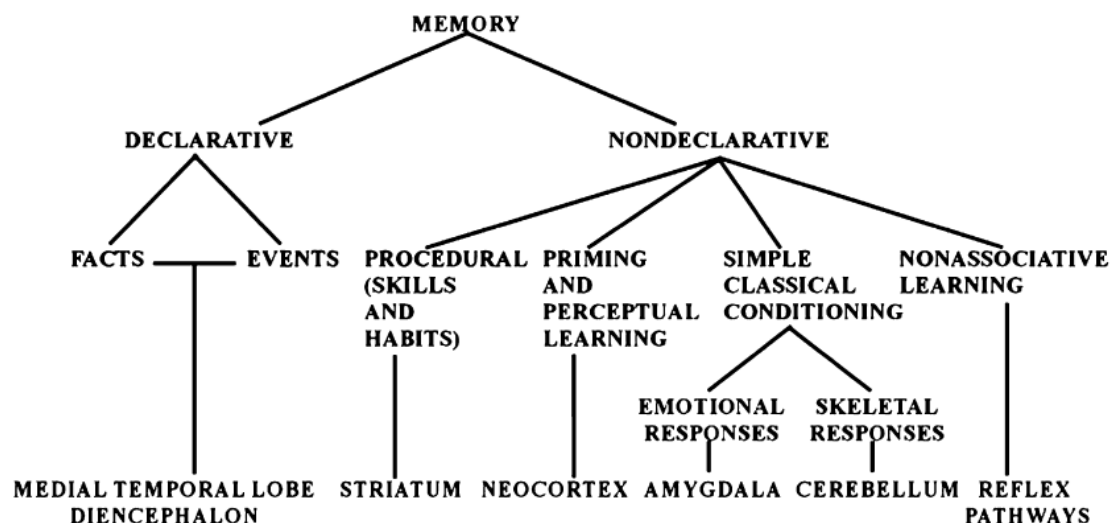


Figure 2: Different memory systems. (Source: Ardila 2012: 203)

1.4 Consolidation

The study of consolidation seeks to understand how new memories make the transition from their initially volatile state to being permanently fixed. There are two different

aspects of consolidation: *fixation* within synapses occurs over minutes to hours, whereas *reorganisation* in brain systems takes weeks to years, and is arguably never complete. Fixation is a cellular process by which short-term synaptic modifications result in permanent changes in neuronal connections (Eichenbaum 2002: 285). The necessity for a process of reorganisation is shown by Ribot's Law of retrograde amnesia (Ribot 1882), whereby memories formed more recently in time before the brain injury are more susceptible to be lost, demonstrating that the fixation and organisation of memories does require a certain amount of time to be completed. This paper will not place much emphasis on *how* memories become fixed, but will simply assume a functional distinction between short and long-term memory without much reference to the involved brain processes.

2. Language origins and evolution

Language is a unique and defining human characteristic, and questions over its origins are numerous. There are also many serious problems with providing an account of the evolution of language. Firstly, our knowledge of how language works in the brain is limited, and even if more was known about its neurobiological basis, brains do not fossilise, so the only evidence we have for their development is endocasts of fossilised skulls. These are useful evidence of absolute brain volume, but complex abilities like language are more likely linked to neural changes which are not easily observed from brain anatomy. We also cannot study naturally-occurring proxies for language in other species, because no other species naturally exhibits any behaviour resembling language. The closest proxies we have are animal communication systems (ACS) like birdsong, or language experiments wherein certain species can be taught a rudimentary language. Another issue is that writing systems are our only direct evidence for language history but they do not appear until as recently as 5,000 years ago, by which time it is clear that humans already had a full language faculty. Language is proposed to have first arisen as early as *Homo habilis*, found in the fossil record between 2.3 and 1.4 million years ago (mya), or as recently as *Homo erectus*, found in the fossil record between 1.8 million and 500,000 years ago.

Despite these issues, there is much that can be hypothesised about language evolution based on indirect evidence. Such evidence has been used to postulate a stage in hominin evolution at which communication was more advanced than the primate calls of our last common ancestor (LCA) with chimpanzees but was not yet as complex as modern language: protolanguage.

2.1 Protolanguage

In order to be considered a qualitative advance from the communication systems of primates, protolanguage must have contained discrete, meaningful symbols – as opposed to

the analogue or graded nature of primate calls – known as protowords. However, considering it as a less complex precursor to language, there are assumed to be some significant structural differences. Jackendoff and Pinker (2005: 223) consider protolanguage to consist of randomly strung-together words, whereas in language “semantic relations among the words are conveyed by syntactic and morphological structure”. A lack of true syntactic structure has serious implications for the nature of protolanguage.

Firstly, language makes use of hierarchical structure whereby words are systematically combined into phrases and phrases into sentences. Calvin and Bickerton (2000: 34) do not suggest that protolanguage was non-combinatorial, but claim that there was no structure in the way words were combined: “in protolanguage, there are no units of intermediate size between the single word and the entire utterance”. Furthermore, there is good evidence (see section 5.7) that phrases in language are assembled prior to being pronounced, whereas protowords are thought to be “sent separately to the part of the brain that controls the motor organs of speech” (Calvin and Bickerton 2000: 42). The systematic combination of words in language depends on distinguished word categories and what positions can be occupied by different types of word; without these category-based rules, combination would simply be linear.

Language also relies on the principle of argument structure by which events and their participants are systematically related. Predication is centred around the verb which, cross-linguistically, can select a minimum of one and a maximum of three obligatory arguments to complete the propositional content of the sentence. Humans recognise that sentences are incomplete and ungrammatical if arguments are randomly missing, but in a pre-syntactic protolanguage, speakers would have had no reason to expect all the necessary action words and participants. Without argument structure, any item may be randomly missing from any position.

A lack of empirical evidence for the protolanguage stage requires us to rely on indirect sources. Bickerton (1981) argues that protolanguage is not a mere artefact from our ancestors but that it is still present in the human brain, surfacing when our language is somehow disrupted. Cases of this include pidgin languages, where complex grammar is lacking and speech is significantly slower, non-native speakers attempting to make themselves understood in a foreign country and even adults in their first language when under the influence of alcohol. Protolanguage can also refer to the type of language acquired by members of the *Pan* genus in ape language research (ALR) experiments, such as Kanzi (Savage-Rumbaugh et. al 1998), where some simple language production in the form of gesture can be observed. Whilst controversial to some extent, these examples do

provide some evidence that language has built on what went before the *Homo* genus and does not need to be regarded as 'all or nothing' (e.g. Berwick 1998).

2.2 The scope of this paper

Based on the evidence above, I will assume a protolanguage stage and propose that memory research can contribute to our understanding of this stage. The research questions provided in the introduction will be addressed using evidence from adult, clinical and child psycholinguistic studies in memory, leading to the provision of a number of hypotheses regarding the nature of protolanguage in greater specificity.

3. The Working Memory Model

The Working Memory model was proposed by Baddeley and Hitch (1974) to describe the type of short-term memory used for temporary storage and the processing of information. This is distinct from the type of short-term memory discussed in section 1.1, which refers instead to the initial stage of forming a new memory that will later become permanent; working memory specifically refers to the type of temporary memory used for tasks such as remembering a phone number for long enough to write it down. The crucial element of WM as opposed to general STM is that it is attentional: it does not just consist of storage, but of information manipulation in active attention.

Neurological research indicates that WM "is largely a frontal-lobe neural network" (Coolidge and Wynn 2008: 80). It is this area of the brain that is most enlarged in the *Homo* genus compared with our most closely related genus *Pan* and has been found to be responsible for the most highly complex cognitive functions such as problem solving, decision making and planning. This indicates that WM is a characteristic human cognitive function.



Figure 3: The Working Memory Model (source: Baddeley 2003: 191)

As shown in figure 3, the original model consisted of three components. The central executive is utilised for the allocation of input to the other two components and retrieval of information from other components of memory (Gathercole and Baddeley 1993: 4). Baddeley (2000: 420) views it as an "attentional control system with no intrinsic storage capacity". Storage functions are allocated to the other two components, known as slave

systems; the phonological loop is specialised for short-term processing of verbally coded information, and the visuo-spatial sketchpad is specialised for visually coded information. The phonological loop further divides into two components: the phonological store contains phonologically coded representations which are subject to the effects of interference and decay, and a rehearsal process recodes non-auditory material into a phonological form and refreshes representations in the store as they decay (Baddeley 1986). The visuo-spatial sketchpad will not be discussed here as there is little evidence that it plays a significant role in language. Crucially, this model makes a claim that the only linguistic aspect of WM is phonological: there is no specific syntactic component.

Baddeley (2000) introduces a fourth component of the WM model: the episodic buffer. This is a mechanism that can integrate information from a variety of other subsystems into a temporary representation that can be retrieved by the central executive by means of conscious awareness (2000: 421). The introduction of this component deals with some issues of the original model, specifically: the ambiguity as to how material in LTM can be reactivated to create new structures which themselves can be manipulated, how information from the two slave systems can be combined if, as the model seems to assume, they are separate components, and how to explain the storage of material in quantities that seemed to exceed the individual capacities of either slave system (Baddeley 2003). The model of WM that will be assumed henceforth, therefore, is shown in figure 4. This diagram shows explicitly how the episodic buffer is capable of integrating the information contained within both original slave systems, as well as information from LTM.

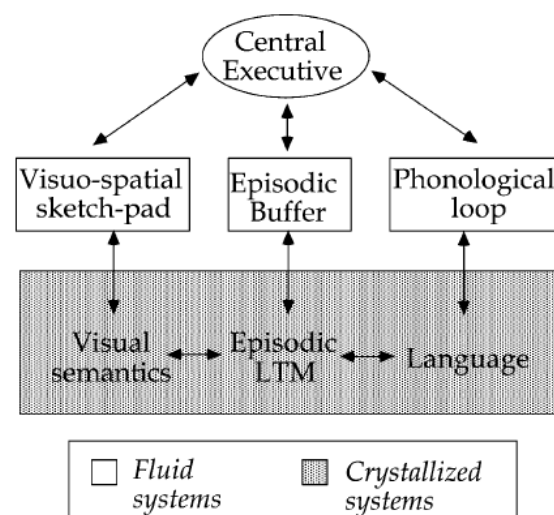


Figure 4: The WM Model with the added Episodic Buffer (Source: Baddeley 2003: 203)

3.1 Jackendoff: a parallel architecture model of WM

Jackendoff (2007: 13) claims that Baddeley's model of WM, while useful for explaining memorisation of nonsense syllables, is inadequate for characterising the perception of actual spoken language. Whilst Jackendoff is rather too quick to dismiss the focus of Baddeley's research and the usefulness of his model, his alternative model of WM is worth considering and may help to explain some of the data in sections five to seven.

The problem Jackendoff presents is the tension in the literature between two conceptions of WM. On one account, it is functionally separate from LTM, in which case "lexical retrieval involves copying or binding the long-term code of a word into WM" (2007: 12). On the other account, WM is simply the currently activated part of LTM, in which case "lexical retrieval consists simply of activating the word's long-term encoding" (2007: 12). There is a serious issue with the second view, known in Jackendoff (2002: section 3.5) as the 'problem of 2': activation does not explain how words are joined up in phrases, and there is no way to account for sentences in which the same word occurs twice or more. On the first account, it is easy to imagine that there can simply be two copies of the same material in WM, each of which has its own relations to the other material (2007: 12).

Jackendoff (2007: 13) wants to think of WM as a 'blackboard' "on which structures are constructed online". His model has three subdivisions for each component of grammar he discusses – syntax, semantics and phonology – and the ability to establish links between them. Because he envisages a parallel and not serial processing, it is unproblematic to think that subdivisions are capable of interaction. This model is more highly specified than Baddeley's conception of the phonological loop, and Jackendoff claims that it avoids Baddeley's tendency to relegate language processing to the central executive without explanation (2007: 14). It also helps explain some phenomena of natural language, such as the distinction between permanent and transient linkages. In idioms, whole phrases are linked and stored in LTM, whereas in the construction of novel sentences, lexical entries are stored separately in LTM, sent to WM and combined there (2007: 13). This distinction is shown in example 1:

[1] [pass][the][shovel] vs. [kick the bucket]

4. Coolidge and Wynn: the evolution of WM

Coolidge and Wynn (e.g. 2005, 2008) consider the evolution both of language and of various other metacognitive executive functions in humans. They claim that, in light of evidence that human WM capacities greatly exceed those of even our nearest primate

relatives, a relatively recent enhancement of WM capacities may be at the heart of modern complex thought. The change is characterised as “small but significant” (2008: 81), meaning that hominin species preceding *Homo sapiens* certainly had advanced cognitive capabilities, but that a relatively small enhancement of these capabilities was the last major factor in the emergence of modern thinking.

This hypothesis remains controversial, in part due to the lack of direct physical evidence: WM is linked to the organisation of neural networks and not known to be connected in any way to absolute brain volume, and neural changes leave few marks on the anatomy of the brain, making endocasts and other fossil evidence largely unhelpful (2008: 81). Archaeological evidence for the evolution of WM instead comes in the form of evidence of the executive functions enabled by enhanced working memory (EWM: Coolidge and Wynn 2005), such as group planning and “temporally and spatially remote action” (2008: 81). Most of the archaeological evidence available demonstrates activities that probably did not place a great burden on WM, like gathering food and using fire, but some technological evidence, like certain foraging technologies dating to around 20 thousand years ago (kya), points to a need for greater WM (2008: 82).

Coolidge and Wynn propose a recent genetic change (60 – 130kya) which finally modernised the human mind; the idea of such a genetic mutation is supported by the fact that WM can be shown in modern humans to have a strong genetic basis (2008: 83). The implications of EWM would be particularly dramatic for language, both in comprehension and production, as will be shown in the discussion in sections five to seven inclusive. Specifically, EWM would have had significant consequences for the ability to use complex grammar. Coolidge (2012: 218) claims that grammatical language is selected for and emerges as a means of bypassing the limits of WM. Elements of grammar such as word order and productive morphology are stored in LTM (see section 8) and therefore imposing such structure on language reduces the amount of processing that has to be undertaken in WM.

Martin-Loaches suggests that EWM was based on “an increase in the number of neurons that can be activated in parallel” (2006: 228), arguing that the phonological loop is too highly specialised for item span to be the only factor in increasing language complexity. Indeed, neurological evidence suggests that several brain areas are relevant to syntax comprehension, so a fairly unspecified increase in neuron efficiency could certainly have had significant consequences for language, as well as other functions manipulated in WM.

5. Experimental evidence

In the absence of any direct evidence for the evolution of metacognitive functions like WM in hominin ancestors, psycholinguistic experiments can show the limitations of WM in modern humans, and based on physical evidence of cerebral expansion, we can speculate that these limitations were greater in previous species. However, the crucial caveat to this is that such indirect evidence is always controversial, and there are many other confounding factors that make modern humans significantly different from our ancestors.

Much research has been done to demonstrate the capacities of verbal WM utilising digit, word and sentence recall tests. These are designed to determine the item span of WM by presenting subjects with a list of items to remember and testing how many on average can be repeated back in the correct order before WM is exhausted by the task. All kinds of confounding influences have been investigated, including the effects of word length, phonological similarity, word frequency and semantic similarity. This section will discuss some of the relevant experiments.

5.1 Ebbinghaus: nonsense syllables

Ebbinghaus (1885) was interested in memory span and designed a method utilising a random selection of 'nonsense syllables' taken from the set of all possible syllables constructed from a vowel or diphthong between two consonants. He believed that his chosen material was advantageous to test memory precisely because of its lack of meaning, as there would be no influence from interest, acoustic 'beauty' or word length, so the results would reflect span in and of itself with minimum interference. His subjects learned the material by reading and were subsequently tested on it.

Ebbinghaus' results are unsurprising: a series of syllables was learned more quickly with fewer members in the series (1885: 48) and with increased number of readings before testing (1885: 52), and retention was best the closer in time the testing and the learning occurred (1885: 62). Interestingly, Ebbinghaus also found that if a subject relearned a sequence they had previously forgotten, forgetting occurred more slowly after the second learning (1885: 81); this suggests that the neuronal connections associated with the learning have remained at least somewhat changed even if the information cannot be directly remembered.

Ebbinghaus alludes to an effect on memory from acoustic similarity, but for him this refers to internal similarity between the two consonants in one syllable, as opposed to

similarity between the syllables of a series, and he does not conduct any specific experimentation in this area.

5.2 Baddeley: word recall

Baddeley has undertaken various studies in verbal WM, as a function of word length and of acoustic, formal (letter structure) and/or semantic similarity. His experiments deal exclusively with presentation, either oral or visual, and memorisation of word lists, not with WM as involved in the comprehension and production of natural language sentences. This lack of attention to real spoken language is sometimes criticised, but the word list experiments are a good place to start to discover the very basic structure and span of WM in and of itself before attending to how it deals with more complex functions like sentence production, which probably involve additional input from various other systems.

Baddeley (1966) presented a set of three experiments to test STM for word sequences as a function of acoustic, formal and semantic similarity, looking to determine whether short-term and long-term memory differ significantly on what cues they rely on to retain information (1966: 362). His hypothesis is that STM will use acoustic cues and LTM will make more use of meaning. His experiments showed, in support of Conrad (1964: discussed in 5.3), that “sequences of items which are hard to discriminate in noise are also hard to remember” (1966: 362); in other words, phonological similarity had a significant negative effect on the subjects’ memory for the word lists. Formal and semantic similarity had very slight impacts (<10% difference from control sample) but were statistically relatively insignificant, supporting Baddeley’s original hypothesis. However, a strong criticism can be made against the research as too many variables changed between each experiment, for example the subjects used, the size of the pool of words from which the sequences were chosen and the frequency of the words used. All of these factors may have had additional influences on the results.

Baddeley, Thomson and Buchanan (1975) examine the ‘item-based hypothesis’, the proposition of Miller (1956) that STM capacity is a constant number of items, versus the ‘time-based hypothesis’ that memory span is not constant and instead varies with the temporal duration of words to be recalled. The experiments presented again test memory as a function of word length, but this time attempt to tease apart the confounding influences that comprise the word length effect: the spoken duration of a word and the number of syllables it contains (1975: 578). Decay theory (Broadbent 1958) assumes that forgetting occurs as a function of time (1975: 578) and therefore predicts a word length effect based on the temporal duration of words; on the other hand, an interference model would predict a word

length effect based on the number of items, with the item being a fundamental speech unit such as the phoneme or syllable. This set of experiments ensures that other variables, such as word frequency and language source, are kept constant in order to examine only the word length effect. The results from all eight experiments were consistent with decay theory, showing that “the temporal duration of items is a powerful determinant of memory span” (1975: 580). In particular, the word length effect remained when the number of phonemes or syllables was constant, showing that the phonological component of WM is more sensitive to time than ‘chunks’ or items.

5.3 Conrad: acoustic confusability

Conrad (1964) and Conrad and Hull (1964) tested memory for letter sequences as a function of acoustic similarity or confusability. The material used in the experiments consists of the spoken names of letters (e.g. b = /bi:/). Letters are separated into two groups depending on the degree of within-group confusability: for example, B, C and D are considered to be highly confusable, whereas A, M and X exhibit a low degree of confusability. The experiments test both the ability of subjects to identify letters when they were spoken against a white noise background, and their ability to recall the letters. All the experiments show a significant correlation between the letters which were confused in the listening test and those confused in recall. From these results, Conrad suggests that memory span is more a function of the number of acoustically similar items in a set than the total number of items.

5.4 Blumenthal: prompted sentence recall

Blumenthal's (1967) experiment sought to demonstrate subjects' success at recalling sentences to which they had previously been exposed, prompted by the presentation of one 'sentence word'. The material used consisted of a control set of 15 active sentences and two sets of 10 passive sentences with the same phrase structure but differing underlying grammatical relations e.g. *Gloves were made by tailors*, with an agent *by*-phrase (standard passive) vs. *Gloves were made by hand*, with a non-agent adverbial *by*-phrase (1967: 204). Subjects listened to test sentences and repeated them back immediately, and then listened to prompt-words and attempted to repeat back the corresponding sentence for each word. Responses were scored as correct if they were identical to or synonymous with the original. The results show that “recall differences corresponded to differences in linguistically defined sentence structures” (1967: 205), meaning that differing grammatical relations of the prompt-words affected the success of recall; the final nouns from the standard passive sentences resulted in successful recall, whereas the final nouns involved in the non-agent adverbial *by*-

phrase were not effective prompts. Blumenthal's experiment demonstrates that there is an observable effect of grammatical relations on STM. However, this effect is probably not relevant to the early stages of language evolution if protolanguage is assumed to lack word categories and therefore grammatical relations (see section 2.1).

5.5 Studies in speech production

Evidence from various studies in speech production shows that there is a component of WM corresponding to the stage of speech motor programming prior to output: the speech output buffer. This is a temporary store where information retrieved from LTM is kept whilst the output utterance is formed. Various studies show that this buffer is dissociable from the phonological loop and that "normal phonological loop function is not necessary for speech planning" (Gathercole and Baddeley 1993: 89). The crucial implication of the speech output buffer is that entire utterances are stored prior to output: this requires a significant degree of WM capacity which is presumed to be lacking in protolanguage, meaning that each protolanguage word is retrieved from LTM and sent to the speech motor organs separately.

Speech errors such as spoonerisms constitute good evidence that the speech output buffer stores utterances prior to output and that it does so in a phonetic domain: the phoneme is the unit of exchange in such errors, and clearly this interaction across word boundaries could not occur if the words of an utterance were represented entirely separately prior to output. Word and list-length effects on speech latency have also been found by Klapp (1974) and Sternberg et al (1978) respectively. Klapp found that word length did not influence speech latency when the target word was known in advance but was strongly correlated with speech latency when visual presentation of the word was the signal to speak. In Sternberg et al.'s experiment, subjects had to speak a list of words known in advance on a 'Go' signal: the latency in responding was strongly and linearly correlated with the number of words in the utterance. The effect demonstrated in these experiments seems to "reflect the construction of a representation of the entire intended utterance prior to response" (Gathercole and Baddeley 1993: 83).

5.7 General conclusions

The studies examined above exemplify some of the factors affecting the operation of WM. The task of speech perception is demonstrably easier when the material presented is phonologically dissimilar, high frequency in the target language and temporally short. In speech production, it is clear that a component of WM is responsible for constructing and representing utterances in their entirety before they are spoken. Finally, WM is a language-external constraint on the length and complexity of utterances, both in perception and

production: presumably both the phonological loop and the speech output buffer are limited in how complex an utterance they can represent.

There are implications of these results for our understanding of language evolution. Building on Coolidge and Wynn's (e.g. 2005, 2008) assumption that WM is a capacity that has increased significantly recently in the *Homo* genus, it follows that the limits on WM would have been greater for the first linguistic hominins. In practice, it therefore seems reasonable to hypothesise that protolanguage was lexically limited to words of short articulatory duration and maximally distinct in phonology. Such a lexicon would make easier the task of speech perception, and successful perception is required for words to be remembered and stored in LTM; words that could not be easily understood in speech would not be easily remembered and would therefore not be functional for speakers. It is, however, important to remember that our LCA with the *Pan* genus did not have a vocal tract capable of speech, so the speech capabilities of early hominins were almost certainly not great enough to build a lexicon of radically phonologically distinct words; an increase in utterance length likely constituted a selective pressure for improvements in the vocal tract. The crucial thing to achieve is the maximum degree of dissimilarity possible given the phonetic inventory available; despite also being minimal pairs, /pæt/ and /kæt/ are easier to distinguish than /pæt/ and /p^hæt/, for example, due to the smaller number of features shared by the initial segments.

Evidence from Calvin and Bickerton's (2000) proxies for protolanguage demonstrates that words are despatched separately to the speech organs, resulting in speech being much slower and more laboured in protolanguage scenarios than in full language. With each word being pronounced separately, the task of speech perception is again made easier and words more likely to be remembered, in line with Berko Gleason and Bernstein Ratner's (1993: 140) observation that words are more intelligible in citation form than in fluent speech: words could never have become conventionalised if they were not intelligible. Precision of pronunciation was perhaps not a factor when there was no fluent speech, but would have had to increase along with an increase in utterance length to ensure that perception was not compromised.

5.7.1 Coolidge and Wynn: recursion

Recursion is a grammatical mechanism allowing the insertion of "a structural unit of a particular type [...] within another unit of the same type" (Tallerman et al. 2009: 140). Recursion allows a phrase to be the object of a higher phrase, giving rise to theoretically infinite utterances. However, in practice, recursion is restrained by language-external factors, crucially, the analytical capacity WM: "the number of recursions must be held and processed

in attention if they are to be understood” (Coolidge and Wynn 2005: 230). Coolidge and Wynn stress that sentences can be grammatical but exhaust WM to the extent that a native speaker could not parse them.

EWM enlarges the recursive capacity of language, yielding an increase in the length and complexity of sentences. This is logically adaptive for two main reasons. Firstly, Jackendoff and Pinker (2005: 230) suggest that recursion in language reflects a recursive thought system, so the ability to communicate recursively gives rise to the ability to make explicit a greater complexity of thoughts. Secondly, recursion is used in speech to achieve specific aims which are beneficial to the speaker, such as persuading, manipulating and deceiving (Coolidge et. al 2010). If we presume that the first hominins to have language had lesser WM capabilities than modern humans, it is clear that recursion could not have been a feature of protolanguage.

6. Clinical evidence

Studies on patients with specific acquired memory deficits provide a valuable source of evidence as to the contribution of normally functioning memory to language. When the disorders result from a specific brain injury, we can also gain a better understanding of what kinds of processing certain brain areas are involved in, which may be useful for interpreting fossil evidence of brain anatomy. This section will look at some of the deficits that have been studied.

6.1 Deficits in phonological WM

Specific deficits of phonological WM included material being encoded in a noisy or unreliable form, phoneme or word order being weakly represented, rapidly decaying representations and an inefficient subvocal rehearsal process (Gathercole and Baddeley 1993: 71). The results of such deficits are poorly specified phonological representations, low discriminability and lack of durability. In this way, poor WM can result in poor LTM for new material as the cognitive processes involved in stabilising representations in a more permanent form cannot take place immediately as the new item is encountered (1993: 72). Furthermore, individuals with low WM capacities have less opportunity in parsing to exploit contextual clues, which impairs the ability of normal subjects to derive the meaning of a new word when it is presented in a sentence (Daneman and Green 1986).

Vallar and Baddeley (1984) studied the memory deficits of PV, a patient who had suffered a stroke, suggesting that “her primary impairment was a damaged phonological store” (Gathercole and Baddeley 1993: 45). Additionally, she lacked the ability to make use

of subvocal rehearsal. She was shown in experiments to retain auditorily presented material very poorly, and her inability to use the rehearsal process was explained by the fact that use thereof, either to maintain decaying phonological representations or to recode visually presented material, would feed information into her severely impaired phonological store, so the process was of no use to her. In later experiments (Baddeley et al. 1988), her ability to learn non-words was shown to be extremely poor. In one recall test, she was incapable of immediately repeating any non-words of more than two syllables. In another test, she was presented with pairs of words – the first a high-frequency word from her first language, the second a non-word – and experienced 100% failure to recall any of the non-words when presented with the high-frequency word as a stimulus, in comparison with a success rate of up to 80% in the control subjects.

6.2 Speech production disorders

The principal speech production disorder referred to in the literature is Broca's aphasia (Broca 1885): this results primarily from damage to Broca's area of the brain (figure 5), with the severity of the aphasia being determined by the extent of the damage beyond Broca's area alone (Ardila 2012: 205).

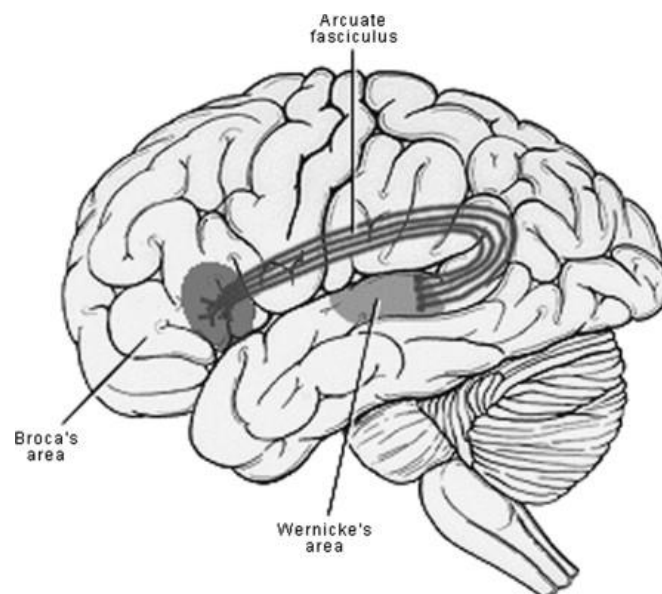


Figure 5: Broca's area of the brain (source: Ardila 2012: 199)

The typical features of Broca's aphasia are non-fluent or laboured speech, agrammatic speech (lack of function words and grammatical inflections), impaired phonological WM skills in repetition tasks and reliance on content words (nouns and some verbs). Gathercole and Baddeley (1993: 87) claim that speech production deficits in Broca's patients are often mirrored in comprehension and memory deficits, and that there are

patients with severe disturbances of phonological WM who have normal output, so Broca's aphasia is likely to be due to an impairment in a specific syntactic mechanism that contributes both to comprehension and output. However, Ardila (2012: 205) claims that more recent evidence points to Broca's patients retaining their *knowledge* of grammar but selecting simpler syntax and shorter phrases because they place less of a burden on their impaired WM. Ardila (2012: 206) admits that there is, as yet, no general consensus as to the precise contribution of Broca's area to language, and what other factors interact in producing the characteristic features of Broca's aphasia.

6.3 Deficits of language comprehension

Spontaneous speech production is usually normal in patients with phonological WM deficits, but they demonstrate difficulties in understanding long and syntactically complex sentences (Gathercole and Baddeley 1993: 212). Saffran and Martin (1975) showed that, in sentence repetition tasks, patients can often paraphrase where they are incapable of repeating verbatim, but that the paraphrases were consistently semantically incorrect for passive constructions (e.g. *the girl was given a present by the boy*) and centre-embedded sentences (e.g. *the boy the man saw carried the box*). This evidence seems to show that the phonological loop is central in retaining order information, which becomes crucial for parsing sentences of the type above where subject-verb-object (SVO) order is not strictly preserved. Baddeley et al. (1987) suggest that the phonological loop can provide a representation that can be used to support off-line analysis; word order is represented and retained while word meanings are accessed and interpreted. When the size and duration of the representation needed to achieve accurate interpretation exceeds the capacity of an individual's phonological loop, comprehension deficits ensue (Gathercole and Baddeley 1993: 215). This explains why patients can successfully parse simple sentences but encounter difficulties with longer and more complex constructions: a more sizeable representation is required to parse a centre-embedded sentence, and it needs to be more durable while SVO order is reconstructed.

Wernicke's aphasia (Wernicke 1874) is a specific kind of deficit in language comprehension resulting from damage to Wernicke's area (figure 5). Syntax is well preserved in production, but patients have problems parsing input, discriminating acoustic information and associating word forms with meaning (Ardila 2012: 201). These deficits result in a decrease in the lexical repertoire and deficiencies in word selection, so despite the preservation of sequencing, speech output lacks meaningful content.

6.4 Nonfluent aphasia

Code (2011) discusses the relevance of aphasic lexical speech automatisms (LSAs) to the evolution of protolanguage. LSAs are the typically occurring patterns of real words seen in severe nonfluent aphasia; in many cases, they ultimately evolve into total agrammatism. Code explains that LSAs are “made up of real words, their features best characterised by formulaic language models” (2011: 138), meaning that utterances are produced automatically and not newly or individually generated. Lending support to experimental findings that high frequency words are more easily remembered, it is these words that are preserved in severe aphasia and produced in LSAs. Speakers are also often unaware of the inappropriateness of their utterances to the present context; the majority of LSAs “have no referential, contextual or intentional connection with the speaker’s world” (2011: 138). A common LSA type is [pronoun + modal/auxiliary verb] e.g. *I want* or *I can*, and many aphasics produce identical LSAs, and speakers frequently produce utterances of a type they were never previously in the habit of using, most notably in the case of expletives, emphasising the significantly reduced semantic range available and level of automatism in severe cases.

There are also non-lexical speech automatisms (NLSAs) which are non-propositional, consisting simply of the most primitive phonetic utterance produceable: the CV syllable. These are the only speech outputs preserved in the most severe cases of aphasia. Code suggests that such non-propositional language “may be seen as evolutionarily pre-linguistic” (1994: 140) as it is this kind of vocalisation that also exists in primates, serving only to communicate some level of emotion. Both types of automated utterance, LSAs and NLSAs, along with other types of formulaic language like expletives and stock phrases, have been shown to be processed largely in the right hemisphere, unlike left hemisphere processing of combinatorial language. Code therefore suggests that they represent a stage pre-dating the enlargement of Broca’s area, and that only the creative and propositional aspects of language have become left-lateralised (2011: 139).

Code emphasises that LSAs in pathology like aphasia have no intention, and therefore carry no true meaning, whereas we must presume that the equivalent structures in protolanguage were intentional and meaningful, as in other examples of formulaic language in normal modern adults. However, their neurobiological roots in ancient neural systems that we share with the *Pan* genus (2011: 139) and lack of syntactic structure make LSAs seem like good evidence for possible characteristics of protolanguage utterances.

6.5 General conclusions

According to Ardila (2012), there are two basic linguistic operations: selecting and sequencing. Selecting, the lexical process, is impaired in deficits of language comprehension as patients lose the ability to associate word forms with their semantic content. Sequencing, the grammatical process, is impaired in deficits of speech production as patients retain their knowledge of meaning but lose the ability to express it in grammatical form. Clinical evidence of the two main types of aphasia, Broca's and Wernicke's, show that these two processes are independent of each other in the brain. As such, there is no reason to think that they had to evolve simultaneously, which fits with our picture of protolanguage as being reliant on content and lacking in complex grammar. The difference in development of the corresponding areas in chimpanzee brains also explains why, in ALR, the animals are able to acquire some lexical/semantic knowledge but word order is often random in production (e.g. Savage-Rumbaugh et al. 1998).

Whilst not assuming that Broca's and Wernicke's areas are the only contributors to language, their development is certainly vital to an evolutionary account of language. The presence of some lexical/semantic capabilities in the *Pan* genus indicates that enlargement or rewiring of Wernicke's area must have begun before the split with our LCA. Fossil endocasts support this conclusion, showing that the leftward asymmetry of Wernicke's area originated before the divergence (Spocte et al. 2010). However, language as a grammatical system appears to be exclusive to *Homo sapiens*, suggesting Broca's area has undergone relatively recent and extensive development. This conclusion is supported by brain imaging: the size of the entire human temporal lobe (wherein is contained Broca's area) is "significantly larger than predicted for a primate brain of human size" and the complexity of some of its component structures like the pathway between Broca's area and the left posterior temporal cortex (Ardila 2012: 204) are significantly greater than is found in the *Pan* genus. This enlargement of the temporal lobe has been calculated to have occurred around 2-3kya, well into the *Homo sapiens* era (Kochetkova 1973). This again supports the theory of protolanguage that pre-*sapiens* hominins may have had lexical/semantic communication systems, but lacked complex grammar.

7. Developmental evidence

It is thought that ontogeny recapitulates phylogeny to some extent, meaning that evidence of the development of language in childhood may be useful in the study of language evolution in hominin ancestors. As with evidence from modern adults, developmental evidence must be considered with a certain amount of scepticism. Not only

do *Homo sapiens* children, unlike the first hominins to have protolanguage, have a full language faculty, but they also grow up in the company of fully-lingual adults who often make a conscious effort to assist their language learning. Because of these clear advantages to the child, it is difficult to know how relevant the data really is.

Research has found that children show an increase over years in their ability to temporarily retain verbal information; the average four year old has an auditory digit span of 2-3 items which, by the age of 14, has increased to 7-8 items (Gathercole and Baddeley 1993: 25). There are three logically possible explanations for this: the memory capacities could be absent in neonates and emerge at a particular developmental stage, the capacities may be initially present and expand with maturation, or the capacities may be present with the efficiency of each component increasing with age. Evidence supports the final possibility, with all major developmental changes identified in each component of memory being in “the operating efficiency of the working memory subsystems and in the increasingly effective use of strategies to maximise working memory functioning” (1993: 26). It seems that the increase in span is not due to a genuine enlargement, but due to an increase with age in the rate of subvocal rehearsal and articulatory speed, both of which are processes which render the memory system more efficient. Close ties have been found between rate of articulation, rate of rehearsal and success of recall and are claimed to show that rehearsal is a process of “subvocal articulation in real time” (1993: 27), meaning that limitations on articulation speed directly influence the ability to rehearse and therefore how many items are lost to decay.

Case et al.’s (1982) ‘constant total space hypothesis’ states these findings as reflecting a memory processing space that remains constant in its entirety but changes in the size of its subcomponents. Total processing space consists of storage space and operating space, and the claim is that older children need less space to perform intellectual operations with increasingly efficient encoding and retrieval operations (Gathercole and Baddeley 1993: 32).

Brown (1973) also observed functional semantic relationships in the two-word stage of children’s language development, such as *action* + *agent* and *entity* + *attributive*, shown respectively in examples 2 and 3. These examples show that there are ways of relating words in the absence of complex grammar, which would have been crucial in multi-word protolanguage utterances.

[2] daddy kiss

[3] water hot

7.1 General conclusions

If ontogeny does indeed recapitulate phylogeny, findings from developmental research may suggest that EWM (Coolidge and Wynn 2005) in hominin ancestors was specifically caused by a reorganisation of the processing space which allowed more efficient use of WM, rather than by an increase in brain volume which allowed a larger processing space. This clearly fits with evidence that neural organisation is far more important to WM than brain volume, especially since notable cerebral expansions are found in the fossil record much earlier than Coolidge and Wynn hypothesise EWM to have appeared, and it is *Homo neanderthalensis*, not *Homo sapiens*, with the largest brain in the hominin line. Developmental research also suggests that, before a significant enhancement in hominin WM, articulation was probably slow and utterances of short length; these are exactly the features Calvin and Bickerton (2000) attribute to protolanguage.

8. Long-term memory and language

Language is processed for speech and comprehension in WM, but it must be the case that elements are stored in LTM and then transferred to WM for processing as WM is not a storage system in and of itself. Whilst it may have increased in size in the *Homo* genus, archaeological evidence of tool use in archaic hominins and ALR results seem to suggest that some capacity for LTM was present long before any linguistic selective pressures.

8.1 Lexicon

The lexicon is a clear example of a component of language that is held in long-term storage, both the phonological shape of words and their associated meanings. Carstairs-McCarthy stresses that “absolute brain volume does not correlate directly with vocabulary size” (1999: 11), meaning that fossil endocasts alone are not sufficient evidence for when hominins first possessed a lexicon. However, the size of the human lexicon is certainly a defining characteristic. Chimpanzees have shown themselves to be capable of acquiring referential symbols in ALR experiments, but the most extensive vocabularies have only reached around 300 words, as opposed to the 30-60,000 words mastered by human adults (Coolidge and Wynn 2008).

8.2 Morphology

Productive morphology, like regular verb endings, is also an element stored in LTM. Jackendoff (2002: section 6) claims that productive morphology reduces the number of entries needed in the lexicon. Markers such as those for plural, gender, person, number and

tense are stored in LTM and can be affixed to word stems in WM before production. Therefore, only word stems have to be stored in the lexicon along with one copy of each morphological marker in the given language; although there may be many productive affixes in a language with rich morphology, this method of storage still adds up to far fewer entries than storing each stem along with each of its inflected forms.

8.3 Syntax

Word order is the crucial factor pointed to by Coolidge in modernising language and reducing the load on WM: “one possible way to bypass the limits of working memory capacity is to have an imposed order in the word lexicon such that individuals could expect, for example, that actions followed agents” (2012: 218). However, there are several reasons that this evaluation is not necessarily linguistically sound. First, there are languages that have almost no restrictions on word order and rely on other structural clues like rich morphological marking (most notably Australian languages e.g. Walpiri). Secondly, even in languages with fixed word order, there are constructions that disobey the normal rules and can still be parsed by normal adults. English conforms to an SVO order, with the subject usually being an agent of some kind, but the passive construction makes sentences in which the agent is relegated to final position (examples 4 and 5):

[4] Chris gave Lee the book

[5] Lee was given the book by Chris

Furthermore, as discussed in Bickerton (2012), syntax is absolutely not a simple matter of linear word order, with true relationships between words being established hierarchically; linear order is simply a by-product of the fact that speech must occur in real time (2012: 456). Word order certainly could play a part in two or three-word protolanguage utterances, as it does in child language (note the potential for semantic difference between *daddy tickle* and *tickle daddy*), but alone it does not constitute the basis for full language.

Despite these issues, there is certainly some weight to Coolidge’s hypothesis if the focus is adapted from ‘word order’ to ‘syntax’. It is true that WM is exhausted by the task of parsing linguistic input if there is no structure to give clues to meaning, and ambiguity could well constitute a selective pressure for syntactic structure, along with the need to process in attention longer utterances as the vocal tract improved. However, other than proposing that they *do*, Coolidge gives no explanation as to *how* the limits of WM may have selected for grammatical structure.

I suggest that the answer lies in Jackendoff's (2002: section 8) evolutionary perspective on his Parallel Architecture. Jackendoff presents a sequence of distinct stages, an incremental evolution of syntax from protolanguage, which provides one plausible explanation as to precisely how fixed word order and, ultimately, syntax, may have emerged. He refers to a level of linguistic competence achieved and sometimes never exceeded by adults in a second language called 'The Basic Variety' (BV: Klein and Perdue 1997). BV lacks complex syntax, but the relevant feature is the use of semantically grounded principles of word order, most prominently Agent First and Focus Last. Agent First seems to be the default principle of modern language with agents typically expressed in subject position, and it disambiguates a large proportion of utterances. For example, a speaker employing Agent First would only use *chase elephant John* to mean that the elephant chased John, not that John chased the elephant. The principle of Grouping, whereby modifiers appear adjacent to what they modify (e.g. *dog brown eat mouse*), is also a purely semantic and linear principle that does a lot of work to disambiguate utterances (Jackendoff 2002: 248). His model then shows phrases and heads emerging from single words; for example, *the little boy* is just an elaborated version of *boy* and could, as a unit, occupy the same position as *boy* itself. He suggests that "syntactic categories emerged as a result of distinguishing verbs from everything else" (2002: 259) and that a simple noun-verb distinction sets up the possibility for various aspects of syntax and morphology, like specialised vocabulary and inflections, to continue increasing the complexity with which semantic relations are shown. In this way, an initial stage that conveys only simple semantic relations lays the foundation for complex grammatical principles.

Jackendoff's model of how syntax emerges is perfectly complimentary to Coolidge's theory that syntax is selected for to bypass the limitations of WM. Parsing is a crucial task of WM and is achieved far more easily when structure is available to convey meaning, so a principle like Agent First would precisely be stored in LTM and invoked in parsing to reduce the workload for WM. As phrases begin to emerge, such a principle, along with Grouping, would enable speakers to parse a unit like *the little boy* with little more difficulty than *boy* alone. I suggest that an appeal to the emergence of syntax from semantically-based orderings explains how it is possible that grammar could have evolved as a product of WM limitations on language.

Conclusion

My research questions were:

1. What can we learn about the evolution of grammar and the lexical inventory and structure of protolanguage from psycholinguistic, clinical and developmental memory experiments?
2. In what ways were enhanced memory capabilities adaptive for language?
3. Could memory have constituted a selective pressure for any aspects of language?

The evidence from psycholinguistic, clinical and developmental memory experiments lead me to hypothesise that, assuming the first hominins to have language were not in possession of WM capacities as great as those of modern *Homo sapiens*, the initial stage of protolanguage may have been based on an inventory of temporally short and phonologically dissimilar words, pronounced slowly and separately and combined in only short utterances with no embeddedness or recursion. I also support the theory that a relatively recent enlargement in the temporal lobe, especially Broca's area, has been a crucial neurobiological factor in the evolution of complex grammar, and that right-lateralised examples of language reflect the protolanguage stage that pre-dates this enlargement.

I have shown that enhanced WM capacity is adaptive for language, both in production and comprehension. It takes significant WM resources to parse and produce long and complex utterances, but with the ability to do so, far more information can be reliably and quickly communicated and understood. Furthermore, grammar is a significant aid in reducing the size of the task with which language provides WM, as principles of linear order and hierarchical structure are stored in the stable LTM component and can be invoked during processing. These are factors that may have been crucial to our hominin ancestors in a savannah environment rife with predators and scarce with resources (Bickerton 2002: 214); an ability to communicate in greater depth whilst maintaining speed and reliability could have been the difference between life and death. Indeed, it is often suggested that one reason *Homo sapiens* survived *Homo neanderthalensis* was some kind of linguistic deficiency in the latter (e.g. Trinkaus and Shipman 1993).

Finally, with the modifications I have suggested to the focus of his hypothesis, I am in agreement with Coolidge (2012) that the limitations of WM were one plausible factor in selecting for certain aspects of grammar. Considering the psycholinguistic evidence for WM and the anthropological evidence for the living environment of early hominins, and invoking Jackendoff's (2002) Parallel Architecture, I suggest that ambiguity in protolanguage would have constituted a selective pressure for greater efficiency and reliability in the communication of information, and grammatical structure was the way this was ultimately achieved.

References

- Ardila, A. 2012. 'Interaction between lexical and grammatical systems in the brain'. *Physics of Life Reviews* 9: 198 – 214.
- Baddeley, A. 1966. 'Short-term memory for word sequences as a function of acoustic, semantic and formal similarity'. *Quarterly Journal of Experimental Psychology* 18: 362 – 365.
- Baddeley, A. 1986. *Working Memory*. Oxford: Clarendon Press.
- Baddeley, A. 2000. 'The episodic buffer: a new component of working memory?'. *Trends in Cognitive Science* 4: 417 – 423.
- Baddeley, A. 2003. 'Working memory and language: an overview'. *Journal of Communication Disorders* 36: 189 – 208.
- Baddeley, A. and G. Hitch. 1974. 'Working Memory'. In G. Bower (ed.) *The psychology of learning and motivation*. New York: Academic Press. 47 – 90.
- Baddeley, A., N. Thomson and M. Buchanan. 1975. 'Word length and the structure of short-term memory'. *Journal of Verbal Learning and Verbal Behaviour*. 14: 575 – 589.
- Baddeley, A., G. Vallar and B. Wilson. 1987. 'Sentence comprehension and phonological memory: some neuropsychological evidence'. In M. Coltheart (ed.) *Attention and Performance, XII*. Hove: Lawrence Erlbaum Associates. 509 – 529.
- Baddeley, A., C. Papagno and G. Vallar. 1988. 'When long-term learning depends on short-term storage'. *Journal of Memory and Language* 27: 586 – 596.
- Bartlett, F. 1932. *Remembering: A Study in Experimental and Social Psychology*. Cambridge: Cambridge University Press.
- Berko Gleason, J. and N. Bernstein Ratner (eds). 1993. *Psycholinguistics*. 2nd ed. USA : Harcourt Brace College Publishers.
- Berwick, R. C. 1998. 'Language evolution and the minimalist program: the origins of syntax'. In Hurford, J. R., M. Studdert-Kennedy and C. Knight (eds) *Approaches to the Evolution of Language*. Cambridge: Cambridge University Press. 320 – 340.
- Bickerton, D. 1981. *Roots of Language*. Michigan: Karoma Publishers.
- Bickerton, D. 2002. 'Foraging versus social intelligence in the evolution of protolanguage'. In Wray, A. (ed) *The Transition to Language*. Oxford: Oxford University Press. 207 – 225.
- Bickerton, D. 2012. 'The origins of syntactic language'. In Tallerman, M. and K. R. Gibson (eds) *The Oxford Handbook of Language Evolution*. Oxford: Oxford University Press. 456 – 468.
- Blumenthal, A. L. 1967. 'Prompted recall of sentences'. *Journal of Verbal Learning and Verbal Behaviour*. 6: 203 – 206.
- Broadbent, D. E. 1958. *Perception and Communication*. London : Pergamon Press.
- Broca, P. 1885. 'Du siège de la faculté du langage articulé dans l'hémisphère gauche du cerveau'. *Bulletins de la Société d'Anthropologie de Paris*. 6: 377 – 393.
- Brown, R. 1973. *A First Language: the Early Stages*. Cambridge: Harvard University Press.

- Case, R., D. M. Kurland and J. Goldberg. 1982. 'Operational efficiency and the growth of short-term memory span'. *Journal of Experimental Child Psychology* 33: 386 – 404.
- Calvin, W. and D. Bickerton. 2000. *Lingua Ex Machina: Reconciling Darwin and Chomsky with the Human Brain*. Cambridge, MA and London: MIT Press.
- Carstairs-McCarthy, A. 1999. *The Origins of Complex Language*. New York: Oxford University Press.
- Code, C. 1994. 'Speech automatism production in aphasia'. *Journal of Neurolinguistics* 8: 135 – 148.
- Code, C. 2011. 'Nonfluent aphasia and the evolution of proto-language'. *Journal of Neurolinguistics* 24: 136 – 144.
- Conrad, R. 1964. 'Acoustic confusions in immediate memory'. *British Journal of Psychology*, 55: 75 – 84.
- Conrad, R and A. J. Hull. 1964. 'Information, acoustic confusion and memory span'. *British Journal of Psychology*. 55: 429 – 432.
- Coolidge, F. L. and T. Wynn. 2005. 'The effect of enhanced working memory on language'. *Journal of Human Evolution* 50: 230 – 231.
- Coolidge, F. L. and T. Wynn. 2008. 'Did a small but significant change in working memory capacity empower modern thinking?' In Mellars, P., K. Boyle, O. Bar-Yosef and S. Stringer (eds) *Rethinking the Human Evolution: New Behavioural and Biological Perspectives on the Origin and Dispersal of Modern Humans*. Cambridge: Cambridge University Press 79 – 90.
- Coolidge, F. L., K. A. Overmann and T. Wynn. 2010. 'Recursion: what is it, who has it, and how did it evolve?'. *Wiley Interdisciplinary Reviews: Cognitive Science* 2: 547 – 554.
- Coolidge, F. L. 2012. 'On the emergence of grammatical language as a means of bypassing the limitations of working memory capacity. Comment on "Interaction between lexical and grammatical language systems in the brain" by Alfredo Ardila'. *Physics of Life Reviews* 9: 217 – 218.
- Daneman, M. and I. Green. 1986. 'Individual differences in comprehending and producing words in context'. *Journal of Memory and Language* 25: 1 – 18.
- Ebbinghaus, H. 1885. *Memory: A Contribution to Experimental Psychology*. Translated by Ruger, H. and C. Bussenius (1913). New York: Teachers College, Columbia University.
- Eichenbaum, H. 2002. *The Cognitive Neuroscience of Memory*. Oxford & New York: Oxford University Press.
- Gathercole S and A. Baddeley. 1993. *Working Memory and Language*. Hove: Laurence Erlbaum Associates.
- Halici, U. ND. 'ION 526: Introduction to Neural Networks'. Last accessed 10th February 2013, from: <http://euclid.ii.metu.edu.tr/~ion526/demo/chapter1/section1.1/>
- Jackendoff, R. 2002. *Foundations of Language*. New York: Cambridge University Press.
- Jackendoff, R. and S. Pinker. 2005. 'The nature of the language faculty and its implications for the evolution of language'. *Cognition* 97: 211 – 225.
- Jackendoff, R. 2007. 'A parallel architecture perspective on language processing'. *Brain Research* 1146: 2-22.

- James, W. 1901 c. 1890. *The Principles of Psychology*. London: Macmillan.
- Klapp, S. T. 1974. 'Syllable-dependent pronunciation latencies in number naming: a replication'. *Journal of Experimental Psychology* 102: 1138 – 1140.
- Klein, W. and C. Perdue. 1997. 'The Basic Variety (or: couldn't natural languages be much simpler?)'. *Second Language Research* 13: 301 – 347.
- Kochetkova, V. I. 1973. *Paleoneurology*. Moscow: Moscow State University Press.
- Martín-Loaches, M. 2006. 'On the uniqueness of humankind: is language working memory the final piece that made us human?'. *Journal of Human Evolution* 50: 219 – 222.
- Miller, G. A. 1956. 'The magical number seven, plus or minus two: some limits to our capacity for processing information'. *Psychological Review* 63: 81 – 97.
- Ribot, T. 1882. *Diseases of Memory: An Essay in the Positive Psychology*. New York: D. Appleton and company.
- Saffran, E. M. and N. Martin. 1975. 'Immediate memory for word lists and sentences in a patient with deficient auditory short-term memory'. *Brain and Language* 2: 420 – 433.
- Savage-Rumbaugh, S., S. Shankar and T. Taylor. 1998. *Apes, Language and the Human Mind*. New York and Oxford: Oxford University Press.
- Spocte, M. A., W. D. Hopkins, A. R. Garrison, A. L. Bauernfeind, C. D. Stimson and P. R. Hof. 2010. 'Wernicke's area homologue in chimpanzees (pan troglodytes) and its relation to the appearance of modern human language'. *Proceedings Biological Sciences* 277: 2165 – 2174.
- Sternberg, S., S. Monsell, R. L. Knoll and C. E. Wright. 1978. 'The latency and duration of rapid movement sequences: comparisons of speech and typewriting'. In Stelmach, G. E. (ed.) *Information Processing in Motor Control and Learning*. New York: Academic Press.
- Tallerman, M., F. Newmeyer, D. Bickerton, D. Bouchard, E. Kaan and L. Rizzi. 2009. 'What kinds of syntactic phenomena must biologists, neurobiologists, and computer scientists try to explain and replicate?' In D. Bickerton and E. Szathmáry (eds) *Biological Foundations and Origin of Syntax*. Cambridge, MA and London: MIT Press. 135 – 157.
- Trinkaus, E. and P. Shipman. 1993. *The Neandertals: Changing the Image of Mankind*. London: Jonathan Cape.
- Vallar, G. and A. Baddeley. 1984. 'Fractionation of working memory: neuropsychological evidence for a short-term store.' *Journal of Verbal Learning and Verbal Behaviour* 23: 151 – 161.
- Wernicke, C. 1874. *Der Aphasische Symptomencomplex*. Breslau: Cohn and Weigert.

Appendix 1: Glossary of acronyms

ACS – animal communication systems

ALR – ape-language research

BV – Basic Variety (Klein and Perdue 1997)

CV – a syllable consisting of one consonant followed by one vowel

EWM – enhanced working memory (Coolidge and Wynn 2005)

kya – thousand years ago

LCA – last common ancestor

LSA – lexical speech automatism

LTM – long-term memory

mya – million years ago

NLSA – non-lexical speech automatism

STM – short-term memory

SVO – subject-verb-object

WM – working memory

Appendix 2: index of examples

[1] [pass][the][shovel] vs. [kick the bucket]

[2] daddy kiss

[3] water hot

[4] Chris gave Lee the book

[5] Lee was given the book by Chris