

Vaneechoutte, M. and Skoyles, J.R., 1998;

The memetic origin of language: modern humans as musical primates.

Journal of Memetics - Evolutionary Models of Information Transmission, 2 (link no longer functional).

The memetic origin of language: modern humans as musical primates

[Mario Vaneechoutte](#)

*Department of Clinical Chemistry, Microbiology & Immunology,
Blok A, University Hospital, B 9000 Ghent, Belgium*

[Home](#)

[John R. Skoyles](#)

*6 Denning Road, Hampstead,
London NW3 1SU. UK*

Up from dragons. The evolution of intelligence: <http://www.upfromdragons.com/>

[Abstract](#)

[1. Introduction](#)

[1.1. Memetics and the origin of language](#)

[1.2. The evolution of information and language](#)

[1.3. The major questions about language](#)

[2. Language components](#)

[2.1. Four competencies](#)

[2.1.1. Mental semantics](#)

[2.1.2. Mental syntax](#)

[2.1.3. Vocal flexibility](#)

[2.1.4. Linguistic syntax](#)

[2.2. Evolution of mental semantics and mental syntax \(the first two competencies\)](#)

[2.2.1. Mental representation and semantic ability](#)

[2.2.2. Mental syntax](#)

[2.2.3. Conclusion](#)

[3. Problems with Chomsky, Pinker and Deacon](#)

[3.1. Noam Chomsky: Is there a universal grammar?](#)

[3.2. Steven Pinker: Is it plausible to assume that natural selection for better speech capacities could have happened?](#)

[3.2.1 Introduction](#)

[3.2.2. Language genes](#)

[3.2.3. Individual vs group fitness](#)

[3.2.4. Can language be explained by assuming better reproductive success of better speakers due to an increase of social status?](#)

3.2.5. Does better speech production really brings along a higher social status?
3.2.6. Does higher social status ensure reproductive success?
3.2.7. The development of language by natural selection is even more difficult to explain than other cultural characteristics like tool making
3.2.8. Summary
3.3. Terrence W. Deacon: Did language and brain co-evolve, leading to both a larger brain and better speech capacities?
3.4. Conclusion
4. Phylogenetic origins of spoken language
4.1. Protolanguages
4.2. The musical primate
4.2.1. Humans have unique adaptations for singing
4.2.2. Musical primates and song birds: examples of convergent evolution?
4.2.2.1. Introduction
4.2.2.2. Music and mating
4.2.2.3. Music and bonding
4.3. Present and past social music
4.3.1. Various observations
4.3.2. Music and group identity
4.4. Conclusion
5. Melody and language learning by children
5.1. Introduction: the semantics of spoken syntax
5.2. Music and language development
6. Final conclusions
6.1. Summary
6.2. How does this approach compare to the hypotheses put forward by Chomsky [16, 17], Pinker [67] and Deacon [21]?
6.3. Why has the musical origin of language hypothesis been overlooked?
6.4. Could large vocabularies alone be sufficient for the development of syntax?
6.5. Musicality may also explain other typically human characteristics
Acknowledgements
Notes
References

Abstract

Song (musicality, singing capacity), we argue, underlies both the evolutionary origin of human language and its development during early childhood. Specifically, we propose that language acquisition depends upon a Music Acquiring Device (MAD) which has been doubled into a Language Acquiring Device (LAD) through memetic evolution. Thus, in opposition to the currently most prominent language origin hypotheses (Pinker, S. 1994. *The Language Instinct*, W. Morrow, N.Y.; Deacon, T.W. 1997. *The Symbolic Species*, W.W. Norton, N.Y.), we contend that language itself was not the underlying selective force which lead to better speaking individuals through natural selection. Instead we suggest that language emerged from the combination of (i) natural selection for increasingly better mental representation abilities during animal evolution (thinking, mental syntax) and (ii) natural selection during recent human evolution for the human ability to sing, and finally (iii) memetic selection that only recently (within the last 100,000 years) reused these priorly evolved abilities to create language. Thus, speech - the use of symbolic sounds linked grammatically - is suggested to be largely a cultural phenomenon, linked to the Upper Palaeolithic revolution. The ability to

sing provided the physical apparatus and neural respiratory control that is now used by speech. The ability to acquire song became the means by which children are able to link animal mental syntax with syntax of spoken language. Several studies strongly indicate that this is achieved by children through a melody-based recognition of intonation, pitch, and melody sequencing and phrasing. Language, we thus conjecture, owes its existence not to innate language learning competencies, but to innate music-associated ones, which - unlike the competencies hypothesized for language - can be straightforwardly explained to have evolved by natural selection.

The question on the origin of language then becomes the question on the origin of song in modern humans or early *Homo sapiens*. At present our ability to sing is unexplained. We hypothesize that song capacity evolved as a means to establish and maintain pair- and group-bonding. Indeed, several convergent examples exist (tropical song birds, whales and porpoises, wolves, gibbons) where song was naturally selected with regard to its capacities for reinforcing social bonds. Anthropologists find song has this function also amongst all human societies.

In conclusion, the ability to sing not only may explain how we came to speak, but may also be a partial answer to some of the very specific sexual and social characteristics so typical for our species and so essential in understanding our recent evolution.

Keywords: origin, human, language, natural selection, cultural evolution, music, intonation, rhythm, song, children

1. Introduction

1.1. Memetics and the origin of language

A major topic of memetics is the transmission of information by words. Thus better knowledge about the origins of language could throw light on many of the issues that are presently debated in memetics. Understanding what language is about is also important because it can be put that language is the only essential difference between human and animal existence, a difference which enables to explain most or all of the features characteristic for human psychology and human behaviour (e.g. [92]), which in turn explains why some memes - which we define broadly as bits of behaviourally transmissible information [Note 1] - are spread more successfully than others [Note 2].

Finally, memetic selection tends to get ignored by theories seeking to explain the origins of language. Most prominent theories instead argue for a gene- rather than meme-based origin. Some, for example, conjecture that language arose from Darwinian, adaptationist selection processes by which better speakers had greater reproductive success (Pinker [67], Smith and Szathmari [83]). Also Deacon [21] relies on genes (genetic assimilation of phenotypic characteristics (Baldwinian evolution)) and long term evolution to explain how language could arise, although memes (i.e. the use of symbols) already play an important role in this Baldwinian evolution. To the opposite, the memetic selection hypothesis defended here assumes that all preadaptations for language production and language understanding were naturally selected for other reasons than language, wherefrom language emerged and evolved rapidly and only recently by a process of cultural evolution. Thus, we do not reject natural selection - indeed, our approach largely depends upon it - but we try to understand how and when cultural/memetic selection comes into play and eventually takes over. The emergence of language in a human community by interaction between humans and symbols is not specifically addressed here, but will be the issue of a forthcoming paper [82].

1.2. The evolution of information and language

Informationally, 'life' can be considered as a giant chemical process which took off some 4 billion years ago - with the origin of the first self-replicating cell (also see [\[Note 2\]](#)). For instance, the product of one enzyme can be used as the substrate for other enzymes, cells interact by means of hormones, humors and neurotransmitters and multicellular colonies do so by means of pheromones and scent molecules. And across these organisational levels, the interactions between enzymes enable cells to interact by means of hormones, humours and neurotransmitters, while the cellular interactions enable multicellular colonies to interact by means of pheromones and scent molecules.

When brain (and eye and ear) possessing animals arose, a new manner of information transmission was introduced into biology. Now, living creatures could exchange information in a nonchemical manner through sound and sight, which increased the speed and the flexibility of informing each other and of influencing each others behaviours or of adapting one's behaviour to the cues provided by others, i.e. behavioural instead of chemical interaction. One of the characteristics of such information is that it is 'inherited' in a nongenetic manner [\[35, 93\]](#). One easily imagines how it is impossible to inherit rules for complex social behaviour genetically while on the other hand such morals, habits are readily 'learned' (rather: assimilated, absorbed) by young animals, e.g. through (emotional) punishment and reward (also see [\[Note 1\]](#)).

Dawkins [\[20\]](#) has called these 'culturally inherited' bits of information, 'memes'. Their development went hand-in-hand with parallel evolution of semantic abilities: these signals have to be turned by the brain into information that can be processed by cells that interact chemically (that is, by neurons and their information transmission by neurotransmitters). For example, an alarm cry (auditory cue) has to get linked/recoded by semantic abilities into the neurological (biochemical) concept or category or mental representation of 'dangerous situation'; this then triggers the various neurotransmitter- and/or hormone-transmitted cellular interactions that occur in fright and fear.

Importantly, the alarm cry is already here functioning as a kind of symbol, since similar sounds or signs may have a completely different meaning depending on the situation or the species. The full progress to symbolism in communication - spoken language as used by humans - requires that further encoding takes place: not only are symbols (words or signs) linked to mental images (linguistic semantics), but also word order, affixes, and other morphological modifications/operations/processes enable the communication of relationships between representations (linguistic syntax). Of course, the possibilities for individual thinking and for information transmission between individuals as a result of symbolic language are increased exponentially again.

1.3. The major questions about language

Two questions exist about language development and its origins.

- First, 'How is it that the complex phenomenon of spoken language arose in the animal world?'
- Second: 'How is it possible that children acquire language with such ease irrespective of the language they learn?'

We argue that none of the presently available hypotheses [\[21, 67\]](#) provide final answers to the questions about the origin of language both during phylogeny of the human kind and during ontogeny of the individual human being. First, we briefly review the present dominant approaches.

The present prominent hypothesis on the phylogenetic origin of language is the natural selection approach. This claims that speech - and its associated characteristics like voice, and specialized brain regions and the ability to comprehend syntax of spoken language - was selected by gradual natural selection of genetic

changes, made possible by the selective advantage of speech itself [66, 67, 83] or was made possible by genetic assimilation [21].

The present prominent hypothesis on the origin of language during individual ontogeny [16, 17] - a hypothesis taken for granted by e.g. Pinker [67] and Smith and Szathmari [83] - is that children have a language acquiring device (LAD), that uses an innate Universal Grammar. Syntax, according to this view, is acquired by a child by setting a few parameters of the 'innate grammar' according to those in their parents language.

In this paper, we argue, in opposition to these approaches, first that song production and song interpretation capacities were the essential, naturally selected, preadaptations that enabled language, which readily evolved in a cultural (memetic) manner. In other words, speech preadaptations were naturally selected but only in regard to singing and not in regard to the later use they came to have in language. Second, linked to this 'song being the preadaptation for speech' approach, we argue that children learn spoken language by means of innate melody recognition capacity (Music Acquisition Device or MAD). If genetic evolution contributed to our abilities to learn language, it was in an indirect manner by providing us with abilities to sing. Thus, language learning devices can in fact be considered as memetically adapted song learning ones.

2. Language components

2.1. Four competencies

The capacity to produce and comprehend spoken information presupposes several cognitive abilities.

2.1.1. Mental semantics

The mind must be able to make mental images (virtual representations) that represent externally perceivable objects, agents and situations.

Furthermore, the mind must be able to semantically link sounds (or visual signals in animal behaviour, writing or sign-languages) to these mental representations to enable their communication, i.e., to make sense of them or to convey meaning. These links are symbolic - that is arbitrary or established by convention. This is illustrated by the fact that many words exist in different languages for the same concept or thing. Accordingly, many completely different writing conventions exist.

2.1.2. Mental syntax

The mind must be able to establish the relations and interactions between these representations, whereby some are active or originators (agents, subjects), while others undergo a change in situation or receive actions (patients, objects).

2.1.3. Vocal flexibility

Spoken language depends on the vocal dexterity to produce a wide range of consonants, vowels and intonations.

2.1.4. Linguistic syntax

The mind must be able to process the word order, syntactic 'roles' such as verb, subject and morphological modifications (syntactic affixes and internal phoneme changes) so that it can produce and comprehend them when communicating with others.

2.2. Evolution of mental semantics and mental syntax (the first two competencies)

2.2.1. Mental representation and semantic ability

The ability to form mental representations, mental images of the environment and to categorize these objects does not need explanation in the context of the origin of language. It is clear that categorization and generalisation is a selectively advantageous trait to any heterotrophic, multicellular, mobile, brained organism (i.e., to most animals), since this enables an animal to reduce the reaction time upon perception. If animals, for example, could not generalize the concept of a certain species of tree, they would be forced to investigate each tree as to whether its fruits were edible. It is easily understood how better and better mental representation or categorization capacities were continuously naturally selected for throughout animal evolution.

We define semantic skill as the ability to link visual or auditory stimuli to a mental representation. This skill is nothing new as animals readily assign meaning to auditory or visual signals. For example, a dog observing a bell ringing will quickly learn to link it to the subsequent appearance of food - Pavlovian conditioning. The dog's brain finds no difficulties in linking an arbitrary sound to an internal representation of a real or possible event. Such semantic abilities can be expected to have been selected particularly for auditory and visual signalling to aid communication between members of a species. Thus, social animals seem well equipped to acquire and use culturally inherited semantic meanings.

The semantic ability to link auditory sounds and visual signs to meaning, it should be noted, lies at the heart of culture, and of memes (see also [1.2](#)). Culture, indeed, can be broadly defined as the exchange of information by auditory and visual behavioural cues. Thus, semantic ability enables the existence of behavioural memes (cry x means danger, sound y means anger, melody z means affection, ...) that survive as replicated bits of information with high heritability within learnt culture. This can be seen in the rules for social behaviour that get culturally, behaviourally, memetically inherited among social animals from generation to generation. The composition of memes in an animal culture is often more stable than the genetic composition of the population. These behaviours are the forerunners of the symbolic memes

It should be noticed that linking arbitrary sounds (words) to specific meanings in essence requires no novel skills: also dogs can learn it. From this it follows that we do not need to explain any quantitative differences in semantics. All we need to explain is why humans are so good at this.

2.2.2. Mental syntax

Also we need not explain the existence of mental syntax in the context of language development: the expansion in thinking and intelligence by means of increased mental syntactical abilities is observed throughout vertebrate and invertebrate taxa. Animals recognize different agents and their interactions and the causal links between ongoing processes connecting them. We have now plenty of examples of mental representation possibility and of generalization, categorization and causal reasoning, i.e. thinking in animals [\[36, 37, 40, 53\]](#). There is now strong indication that chimps even succeed in forming mental representations of the knowledge present in the mind of another subject [\[Note 3\]](#).

Mental syntax offers a strong selective advantage because it enables animals to predict possible outcomes of

current situations (aided by memory of past events and their outcomes) and it helps them to make the best choices between different possible actions.

2.2.3. Conclusion

We conclude that mental representation and semantics (linking of observable behaviour to a mental representation or conveying meaning to an observation) on the one hand and mental syntax (recognition of causal links) on the other hand are two abilities that were naturally selected for long before humans appeared and long before the rise of spoken language. Before we explain how vocal flexibility and linguistic syntax arose, we will briefly summarize why we need other hypotheses than those proposed by Chomsky [16, 17], Pinker [67] and - to a large extent - Deacon [21].

3. Problems with Chomsky, Pinker and Deacon

3.1. Noam Chomsky: Is there a universal grammar?

Chomsky [16, 17] argues that syntactical skills are novel to human communication, arising in each variety of human language from parameters set in an innate Universal Grammar. However, current knowledge does not provide evidence of something like innate Universal Grammar.

A peculiar feature of human language is the high degree of diversity in each of its characteristics. For example, in configurative languages - like Indo-European languages, subject (S), object (O) and verb (V) can mathematically be ordered in 6 manners. Although SOV (45%) and SVO (42%) are predominant [1], five of the six possibilities are used (there are no examples known of OSV), by itself a strong indication that almost any conceivable order can be used. Some languages (like Dutch) also use mixtures of SOV and SVO, dependent on the hierarchical position of the sentence. Even more curiously, there are nonconfigurative languages, like Guugu Yimidhirr from N.E. Australia (see also below).

Aitchison has reviewed the difficulties in finding underlying universal rules in grammar of spoken languages [1].

However, to Chomskyans all this diversity is an illusion and under the surface all of these languages are dialects of 'earthpeak' - as Pinker [67] puts it. The syntactical differences are due to different parameter settings given to an innate Universal Grammar. There is a problem with this answer: the notion of Universal Grammar, frankly, is philosophical speculation. We refer to Botha [11], Harris [39], Tomasello [89], Allot [3], Bates & Goodman [6] for useful sources underpinning our skepticism. This nonfactual status of the Universal Grammar hypothesis has been ignored since Chomskyans have been remarkably successful in promoting the idea that Universal Grammar is to language what molecules are to chemistry, gravitation to astronomy or DNA to biology - an established fact.

Not only is 'universal grammar' not universal, it does not even concern many aspects of grammar. Really 'odd' languages exist like Nootka and Mohawk, (two native American languages), Lisu from Burma and Mam from West-Guatemala. The latter for example has a rich vocabulary for the action of laying, depending on the position of laying (on belly, on back, on side), depending upon whether a human or an animal is laying, upon telling whether one lays sick or drunk, etc.). Some languages have no gender classes, some two, other three and Sotho even has six. Furthermore, languages use a limited subset from 757 phones (observed in a total of 317 languages) very differently. They do so with a varying number of consonants and vowels from as low as 11 in the case of the Polynesian Mura to 148 in the African !xu or !Kung. Most average between 20

and 35 [52]. This is not explained by Chomskyan theory.

It appears that 'Universal Grammar' is not a scientific fact but a program which for theoretical reasons assumes that language has a universal core. It is an assumption which Chomskyans have constantly failed to establish. Chomskyan linguistics, we further note, is ignored by many linguists. Rather than being the only approach to grammar - as has been suggested by Pinker [67], it is only one of several. Further, the people who we might expect to make use of it - that is those seeking to computerise speech, completely ignore Chomskyan theory. Also many cognitive psychologists reject it as it fails conspicuously to fit the process by which syntax is acquired by children [6].

The noncredibility of 'Universal Grammar' leaves us with a hard problem: how could natural selection have created the diversity we find amongst human languages? Diversity does not offer the user of any language any advantage. (The only people, we might note, that gain from it are linguists who can make careers based upon studying obscure languages).

In the case of phonetics this is particularly problematic since it is known that infants before nine months are prepared to hear the phone contrasts present in all languages [27, 86, 90], an ability which is lost once they are familiar with their own language [4]. What advantage could exist for such an ability?

This is an unacknowledged but puzzling anomaly in the evolution of phonology and one can wonder why there should exist such a variety when only a small subset of phones (roughly one in twenty to one in forty) are used in any particular language.

In our opinion, this diversity creates work for phoneticians, but it makes no sense evolutionary, except in one circumstance: that linguistic evolution was not responsible for selecting the processes responsible for phone differences but instead coopted existing diversity to the - phonetically more limited - needs of speech. In this view, speech evolution limited itself to developing means to use preexisting information processing sensitivities in the temporal/parietal and motor cortices.

In fact we know this is the situation in phonetics: animals as different to us from chinchillas [46] to quails [45] can hear phones. The auditory cortex of monkeys is as able as that of humans to hear the auditory features which characterise phones [85]. Neurons in the homologous areas to Wernicke's area process phonetic parameters such as fundamental frequencies, voice onset times and place of articulation (for instance, [85]). Phonetics appears to be a case where an important component of speech was not a direct product of natural selection but one that came about from a reuse of processes that had been already evolved much earlier for other reasons.

In conclusion, although we argue that something like universal mental syntax exists (see 2.1.2), we fail to see how spoken syntax could rely on a universal linguistic grammar.

3.2. Steven Pinker: Is it plausible to assume that natural selection for better speech capacities could have happened?

3.2.1 Introduction

Can language be understood as 'an organ', a 'language instinct' [67], that was developed by gradual natural selection in which better speakers had more reproductive success, resulting in the selective survival of genes encoding for such better language abilities?

Pinker [67] and Pinker & Bloom [66] have suggested that the Chomskyan innate Universal Grammar arose

by natural selection. There are many problems with this proposal. Bickerton [9], for example, in spite of being committed to the idea that an innate Universal Grammar arose by natural selection, felt the problems of this happening were so great that it could only be explained by a single and extraordinary macromutation, which is clearly unacceptable to any evolutionary biologist.

The following quote summarizes how Pinker and Bloom [66] propose that natural selection could have played a role in the development of language:

"Furthermore, in a group of communicators competing for attention and sympathies there is a premium on the ability to engage, interest, and persuade listeners. This in turn encourages the development of discourse and rhetorical skills and the pragmatically-relevant grammatical devices that support them. Symons' [87] observation that tribal chiefs are often both gifted orators and highly polygynous is a splendid prod to any imagination that cannot conceive of how linguistic skills could make a Darwinian difference."

Below are some of our objections.

3.2.2. Language genes

Natural selection for language only works if it can be genetically inherited. However, thus far no language genes have been reported in spite of intensive searches that have been made for inherited disorders of language. The case of the family known by the initials KE with an inherited disorder demonstrates the failure of this search, paradoxically by the enthusiasm with which this example has been misreported as an inheritable language disorder. From early on in life members of this family suffer a devastating neurological dysfunction that requires many of them to communicate by sign-language. The disorder affects the coordination of orofacial musculature, both for nonlinguistic uses and for speaking. The syntax problems of these people are reported as their primary problem by those seeking a gene specific for language [32, 33, 67]. Still, recent clinical reports upon this family stress that this claim is false since virtually all aspects of their expressive language - from syntax to articulation - is found impaired [30].

However, let us suppose - for the sake of argument - that specific language abilities are genetically encoded. Would such genes increase the reproductive success of better speaking individuals? There are several problems, some of which are addressed below.

3.2.3. Individual vs group fitness

First, the advantages offered by speech, like more successful hunting, would have benefited all individuals in a hunter-gathering band of early humans, even those with less well developed language capacities.

Indeed, better communication possibilities favour the group (or the species) as a whole and as such it seems implausible that natural selection, which works on differential reproductive success of specific genes, could have worked at all. Group selection means here that reproduction of all genes present in a group is influenced in a similar manner by newly developed behaviours.

Accordingly, Allott [3] notes:

"However, in the case of humans there can also be cultural selection, behavioural selection at the group level, where the patterns of behaviour adopted are not tied to individual genetic differences."

Although Ridley [70] has convincingly argued that in nature group selection usually is a much weaker selective force compared to natural selection, in cases where memetic information transmission plays a role (like in imitating/learning behaviour) we might understand easily how group selection could play its role in evolution. Changes in individual behaviour - regardless whether these changes are genetically encoded or not - can be taken over by other members of a group. In case this behaviour happens to confer some selective advantage, every member of the group can quickly profit of this, regardless the genetic make-up of the individual. Certain groups as a whole then may be favoured since they acquired some behaviour. Therefore an initial coincidental link will exist between certain behaviours (memes) and the collection of genes which happen to be present among the group members with this behaviour. So, all individuals and genes will be favoured, and this will obscure natural selection for the gene which possibly led to the successful behaviour. Moreover, many complex social behaviours do not have a genetic basis but can originate as coincidental inventions of some individual (see the example of Japanese macaques below and our remarks on language genes (3.2.2)).

Taken to its extremes, Pinker's claim that the complexity of language arose as a gradually selected feature compares to stating that our tool making refinements were a consequence of genetic selection. As such, people who had a mutant gene which gave them the possibility to keep a fire burning were reproductively more successful than those without the gene. A later mutant enabled some to make fire by firestones and his/her offspring was reproductively more successful than people who did not possess this genetically encoded capacity, because it is obviously more advantageous to be able to make your fire yourself whenever you want to. Thus the gene for making fire with firestones spread in the population and outcompeted the 'keep the fire burning' gene. Of course, the genetic mutants which could make matches were better off and outcompeted the firestone firemakers through better reproductive success. But alas, present day mutants which use lighters are competing out the match using genetic mutants.

The silliness of this argument is obvious. Still, this is largely what is being claimed by the gradual natural selection approach of Pinker about the origin of language. Moreover, it is even far more difficult to explain selection of language this way than it is to explain tool use (see 3.2.7).

3.2.4. Can language be explained by assuming better reproductive success of better speakers due to an increase of social status?

Even in the case of genetic encoding of cultural phenomena like language, and even in case where individuals gain a higher social status which results from their socially highly valued and (for the sake of the argument) genetically encoded cultural behaviour (but see 3.2.5), this new skill which is first owned by an individual with a mutant gene or a novel recombination of genes, must remain hidden to other members of the species. Because of mimicking capacity, other members will readily copy the art such that the eventual higher social rank brought by the new trick and which might lead to more successful reproduction, is readily lost. The mutant parent even must hide this skill for its own offspring, otherwise both mutant and wild type offspring will take advantage of it, and again no natural selection will be possible.

The example of the Japanese macaques who readily adopted washing sand from sweet potatoes as it was first done by one member of the group is well known (see [Note 1] for remarks on imitation). Whether or not this single group member had some gene for this behaviour (which we heavily doubt), the gene could not lead to higher social status as a result of the behavioural change it introduced, since several group members readily behaved the same way.

3.2.5. Does better speech production really bring along a higher social status?

Pinker [67] claims that people with better speech capacities have more chances to acquire a higher social

rank, becoming a tribe leader or politician, and from this it is inferred that they will have higher reproductive success leading to spread of genes for better speech.

However, there are several pitfalls in this line of reasoning. First, one should not confuse current macrosocial politics - where indeed leadership often has to do more with ones' public image, which indeed partially depends on ones' linguistic capacities - with the original small tribe policies. Under these original conditions, being the leader is often the mere consequence of having a father who was the previous leader, regardless one's (linguistic) skills. Second, one can question whether it were especially better speech capacities which led to high social rank and thus reproductive success. Being a successful hunter, a good parent, an efficient food gatherer, a socially enjoyable person (which depends not necessarily on speech capacities), a very aggressive and physically strong male or a good singer or a sexually attractive partner, are all other and probably more important reasons of why an individual could be reproductively successful. Physical attractiveness may even be a more important reason for reproductive success than social rank in humans (see [3.2.6](#)) and later on we will argue in favour of the attractiveness of male singers - above male orators - on females (see [3.2.5](#)). Whatever, it appears that natural selection will be too weak a force to explain language by the social status it might provide, since speech happens to be only one of many possible other factors which determine social status.

Anthropological studies moreover show that in the small hunter-gatherer bands, the 'big man' is not distinguishable from the other members of the band [[25](#), [26](#)]. To quote Richard Lee upon the hunter-gatherer !Kung: 'None is arrogant, overbearing, boastful, or aloof. In !Kung terms these traits absolutely disqualify a person as a leader and may even engender forms of ostracism... Another trait emphatically not found among traditional camp leaders is a desire for wealth or acquisitiveness... Whatever their personal influence over group decisions, they never translate this into more wealth or more leisure time than other group members have' [[47](#)]. The kind of social organisation (tribes and kingdoms) which Pinker has in mind and where speech eventually might have increased social status, eventually (but doubtfully, see [3.2.6](#)) leading to a minor reproductive advantage, does only exist since about 10 000 years, well after the origin of language [[25](#), [26](#)].

3.2.6. Does higher social status ensure reproductive success?

There is strong evidence - for example from analogy with social animals - that reproductive success is indeed closely linked to social rank.

However, just in case of humans - where this link between social status and reproductive success is needed most to supply natural selection as an explanation for language - it may not strictly be applied.

Humans, living in fission-fusion societies with strong pair bonding and with prolonged periods of absence of the males, appear to be a special case. It has been suggested that females indeed do prefer partners for life with high social rank, thus ensuring material advantages for raising offspring, but that they try to choose physically - genetically attractive partners for sexual reproduction. Strong evidence for adulterous behaviour of females - at least in original human tribes - comes from the many highly complicated adaptations of both female and male reproductive behaviour at the level of oocytes and spermatozoides. For instance, it has been shown that males produce killer spermatozoides - able to kill spermatozoides from other males - and that these are produced especially when there may be suspicion of adulterous behaviour of females (for instance after long absence of the male). On the other hand, it appears that the female body can regulate which sperm of different partners is preferentially taken up, for instance by - subconscious - regulation of orgasmic experience [[5](#)].

In conclusion, the high social rank of a human male not necessarily confers absolute ensurance for better reproductive success.

3.2.7. The development of language by natural selection is even more difficult to explain than other cultural characteristics like tool making

For language the problem for a natural selection explanation is even more difficult to overcome than it is for other cultural traits like tool making. Not only producers must be selected, but at the same time very different mutations - those mutations which enable understanding of what producers say - have to be selected gradually and naturally. We have previously pointed to the same bottleneck in the explanation of behavioural mate recognition systems [93] and this problem for language has been formulated also by Geschwind [31]. The arguments by Pinker & Bloom [66] to resolve this paradox, are far from convincing, and finally they have to rely on the Baldwin effect (see also Deacon [21]). It appears an odd supposition to state that better story-tellers will gain high social status, when one has to explain how 'better story understanding' genes have to be selected independently at the same time.

3.2.8. Summary

In summary, there is as yet no convincing evidence that language genes exist (see 3.2.2). Second, in cases where behaviours can be inherited by learning and mimicking, natural selection for individual genes can be a weaker selective force than group selection, whereby group selection favours all genes in a social group indifferently. Natural selection can explain the increase of general abilities: better vision, higher intelligence, better singing capacities. When it comes to explain how specific, directly observable and mimickable abilities - like speech and like making tools - can be selected, group selection becomes important enough to counteract or overwhelm natural selection, because new findings of individuals will be taken over by others, whatever their genes (see 3.2.3). For the same reason of mimickability, it is unlikely that mimickable behaviours will lead to higher social rank (3.2.4).

Third, trying to explain how language could be naturally selected by assuming that better speech entails a higher social status which in turn leads to reproductive success, can be criticized by showing that speech was (and is) only one of several factors in determining social status (see 3.2.5), and that it is uncertain that social status of humans is a guarantee for reproductive success (see 3.2.6). Fourth, it should be noticed that there is not really a link between being the 'leader' of a hunter-gatherer band and social status or reproductive success (see 3.2.5). The example of Symons [87], adopted by Pinker [67] on the reproductive success of tribal chiefs (often both gifted orators and highly polygynous) then is not really applicable to the humans which first developed language. Finally, it is difficult to see how natural selection for better speech could work, when realizing the difficulty of selection for better speech understanding to occur simultaneously (see 3.2.7).

3.3. Terrence W. Deacon: Did language and brain co-evolve, leading to both a larger brain and better speech capacities?

The hypothesis of Deacon [21] is well summarized by the following quote:

"Considering the incredible extent of vocal abilities in modern humans as compared to any other mammal, and the intimate relationship between syntax and speech, it should not surprise us that vocal speech was in continual development for a significant fraction of human prehistory. The pace of evolutionary change would hardly suggest that such an unprecedented, well-integrated, and highly efficient medium could have arisen without a long exposure to the influence of natural selection. But if the use of speech is as much as 2 million years old, then it

would have been evolving through most of its prehistory in the context of a somewhat limited vocal capacity. It is during this period that most predispositions for language processing would have arisen via Baldwinian evolution. This has very significant implications for the sorts of speech adaptation that are present in modern humans." (page 358-359).

Here Deacon [21], like Pinker [67], relies on long term (2 million years) gradual evolution through selective advantage offered by the use of symbols, and he relies on Baldwinian evolution. It should be noted however that Deacon [21] clearly dismisses the notion of Universal Grammar [16, 17].

For the moment, it will suffice to say that we claim that our large brains did not expand to enable language and that language did not cause brain expansion. For example, microcephalics [77] and individuals with only half the brain of normal humans and so with brain masses within the upper limit of nonhuman primates - can learn normal speech [81]. It might help, but you do not need a large human brain to be able to speak. Furthermore, the archaeological evidence indicates a late origin for language [58]. For further comments on Deacon, see 6.2.

3.4. Conclusion

We have summarized some possible criticisms on the most reknown hypotheses on the origin of language and we have indicated why we have difficulties in accepting the existence of some kind of Chomskyan 'universal grammar' (except for some basic mental syntax which we share with animals (see 2.2.2)) and why a genetic explanation, adaptationist [67] or assimilative [21] seems implausible to us. There are several other criticisms possible [3, 82, 89].

Still, this denial of a direct role of natural selection in the origin of speech and the arguments in favour of a cultural evolution process to understand the origin of language do not supply us with the concrete genetic preadaptations we need to understand how both production of symbolic sounds (vocal flexibility) and giving structural value to words in sentences (linguistic syntax) have been achieved.

We will try to answer the first question on vocal flexibility largely by evolutionary considerations about the phylogenetic origin of language (section 4), while the second question will be approached in an attempt to understand how infants acquire language (section 5).

It is now finally time to readdress one of the most ancient explanations for the origin of language: our musicality or singing capacity, which is essential in explaining both the phylogenetic and developmental origin of language. Both the origin of language and its development in children, we argue, can be best understood by recognising that we are musical or singing primates in the first place.

4. Phylogenetic origins of spoken language

4.1. Protolanguages

Are there cues to protolanguage in close relatives of ours? Burling [13] states: "Since our surviving primate communication system remains sharply distinct from language, it is implausible that it could have served as the base from which language evolved. We are more likely to find hints about language origins by studying how primates use their minds than by studying how they communicate." The same conclusion was reached by Jonker [42]. This is not in contradiction with our claim for some universal mental syntax among higher animals (see 2.1.2).

However, we should mention that opinions differ:

"The analysis of the so-called long calls in chimpanzees and bonobos make it likely that the group-living great apes preserved the ability to create syntactically different calls, which would be developed by requirements of social life. A call repertoire emerged in these species which contained a large number of call variants at group level available for each group member via social learning. This type of animal call is different from ordinary animal communication; it shows some features of human language." [91].

There is also some controversy with regard to the speech capacities of our ancestors. Is speech already present in *H. erectus*? Since when can *H. sapiens* (which originated about half a million years ago) speak. Could *H. sapiens neanderthalensis* speak?

The archaeological evidence indicates that planning and other complex activities date back at the earliest perhaps 60,000 years ago. Noble and Davidson [63] argue that increasing tool use capacities, the occurrence of cultural artifacts (paintings, statues), and burial practices follow from the mental activity enabled by language. Such behavioural evidence of language starts with the Upper Palaeolithic around 40,000 years ago. Maybe significantly, it was at this time that anatomically modern humans started to replace Neanderthals which only became extinct between 40,000 to 32,000 years ago. Others also argue in favour of a late origin of vocal language [58].

Here, we adopt the point of view that spoken, symbolic language is quite different from primate languages and that it originated only recently.

4.2. The musical primate

4.2.1. Humans have unique adaptations for singing.

The idea that the origin of speech lies in our ability to sing can be traced back to at least Jean Jacques Rousseau, in the seventeenth century [73]. It was suggested by the famous linguist Wilhelm von Humboldt in the nineteenth century [94] and by Otto Jespersen early in this one [41]. However, this approach to language has been ignored in more modern times. Indicative is that the word 'music' lacks in the index of the recent books of Pinker [67] and Deacon [21]. In recent times, music has received serious attention by some linguists [48], but this was done within the Chomskyan paradigm and did not address the origin of language.

Just like song birds possess highly sophisticated syringes, there are very characteristic morphological changes of the human glottis and larynx, unequalled in any mammalian species [75]. Aitchison [1] remarks: "*Our language has more in common with the singing and calling of birds, than with the vocal signals of apes.*"

The resemblance to bird song was noticed already by Charles Darwin [19]:

"(Language) is certainly not a true instinct [Note 4], for every language has to be learnt. It differs, however, widely from all ordinary arts, for man has an instinctive tendency to speak, as we see in the babble of our young children; whilst no child has an instinctive tendency to brew, bake, or write. ... The sounds uttered by birds offer in several respects the nearest analogy to language, for all the members of the same species utter the same instinctive cries expressive of their emotions; and all the kinds which sing, exert their power instinctively; but the actual song, and even the call-notes, are learnt from their parents or foster-parents. These sounds, ..., are no more innate than language is in man."

Provine [68] has shown that a unique overlooked feature of human speech is our ability to integrate respiration and vocalisation. We, as humans, breathe in a way unique among the primates - since only we can neurally modulate sequences of tonal vocalisations upon our expirations. Other primates can vocalise but they are limited to only one vocalization per expiration. For example, both humans and chimpanzees laugh: however, chimpanzees do so by an 'ah', 'ah', 'ah' sequence of repeated inspirations and expirations. In contrast, we do a modulating 'ha, ha, ha, ..' or 'ho, ho, ho ..' upon a single out-breath - this modulation often going on continuous for 16 laughter syllables [68, pp. 40-41]. Moreover, we can subtly tune our series of vocalisations upon a single continuous out-breath. Only amongst birds - not other primates - are there species that possess comparable respiratory-control ability. This underlies the curious fact that while some birds can imitate human speech, the much more closely related chimpanzee or any mammal cannot.

The neural control that allows song was, we suggest, a profound revolution: the 'one breath one-vocalisation' rule stops chimpanzees not only from laughing like humans but also from being able to control the expiration needed to speak. This, as Provine [68] notes, is the reason why attempts to teach spoken language to chimpanzees have failed in spite of them being able to learn sign and token-based languages and even to understand spoken speech [76, p. 40-41].

Neural control of respiration allows many more kinds of vocalizations: over 700 vowel, diphthongs and consonantal phones were found in a sample based upon only one-twentieth of all the world's languages [52]. Moreover, such control allows the concatenation of very complex sequences. Thus, vocalisations upon a single out-breath combine into words, and these in turn combine into clauses, phrases and sentences. Neural control also allows modulations to be superimposed upon these vocalisations, such as intonation (linguistic, pragmatic and emotional), and this can be upon a wide variety of speech types like whisper, song, chant, scream, motherese, 'Donald-Duck speech' and ventriloquism.

The tonal modulation of song is not only enabled by neural control but also by anatomical specialisation of the vocal tract for producing a wide variety of pitches and timbres. The peculiarity of our vocal tract is usually attributed to enabling speech, although it is sometimes also considered as a mere consequence of postural changes between the head and thorax that accompanied the upright stance and human-style bipedal locomotion (see also the [postscript](#)). However, the anatomical characteristics of the vocal tract are more closely linked to our capacity to sing than to our capacity to speak. People cannot sing without fully using all their vocal tract. However, people can speak without using large parts of the vocal tract (for instance in buccal speech, more familiarly known as Donald-Duck speech). Although normal speech contains a range of vowels and consonants that fully exploit the vocal tract, sufficient variety amongst the world's languages exists to suggest that intelligible speech only needs a subset of possibilities, exploiting only part of the vocal tract's pronunciation potential.

Without the neural control that enables song, speech could not exist. But which came first? We argue that we can speak because we can sing, and not that we can sing because we can speak, also for parsimonious reasons: the capacity to speak requires in addition to respirational control also syntax, phonology and the capacity to use and learn a vocabulary of words (see also the remarks in [3.2.7](#)), while singing requires none of these (songs can exist without words). Second, in the development of speech by children, melody - in terms of interest in and production of intonation and rhythm - comes before other aspects such as phonology, syntax and vocabulary (see section [5](#)).

4.2.2. Musical primates and song birds: examples of convergent evolution?

4.2.2.1. Introduction

The exact reason for the origin of singing behaviour is beyond the scope of this paper, but it is clear that the ability to sing has been naturally selected on many separate occasions - e.g. birds, whales and gibbons. Where this has happened, there have often been highly complex adaptations both anatomical and neural. The major idea here is that the complex changes which were necessary to develop an organ which eventually could be used for symbolic language production were selected for singing and not for speech. Convergent evolution to what may have happened to modern humans can be observed in song birds. Also song birds developed highly complex adaptations, anatomical and neural, as a result of natural selection for better song capacities [\[Note 5\]](#).

4.2.2.2. Music and mating

Song production and song preference play an important role in mating in song birds. Possibly music had a similar role originally in human mating - and it still has to some extent. Below are some of the several possible examples of the central role of music in courting behaviour. In several cultures males indeed bring serenades for their beloved. Also, male singers and musicians in general exert strong physical attractiveness on females (some females even have orgasmic experiences during concerts). Much poetry and love texts sound silly when proclaimed, but are quite acceptable and even touching and convincing when sung. Adolescents meet through singing, listening to music and dancing.

Moreover, sexual selection of the ability to sing is more plausible than sexual selection of the ability to speak. Sexual selection requires only an inherited preference for singers of distinctive emotional melodies rather than good story telling - something that requires that language itself is first well understood.

However, it might be objected that this fails to explain why females would also sing and speak. It should be noted that, while it is true that in many song birds only males sing, females inherit genetically the abilities to sing - something that can be shown since female singing can be triggered by hormonal treatment. Therefore, it is evolutionarily possible that a small genetic change triggered hormonal changes so that singing by females became possible, after it had first been sexually selected for in males. From considering some tropical song birds, we might understand how song capacity of females might have been selected for, eventually after it arose in males by sexual selection first.

4.2.2.3. Music and bonding

Indeed, the situation whereby male song birds exclusively sing happens to be true only for temperate regions. In some species of tropical song birds, females as much as males can engage in singing. Moreover, unlike in temperate areas, where male song links to the defense of territory and attracting potential mates, in these tropical species male and female singing links to bond formation and bond maintenance. This becomes apparent from the following quote [\[88\]](#):

"In the tropics, although there are many species of birds the song of which is doubtless just as territorial in function as is usual in the temperate regions, the ornithologist is also struck by the number of examples where song appears much less aggressive in intent and where its function is apparently as a social signal, for maintaining pair and family bonds and as part of the sexual display, rather than a territorial one. Moreover, it is perhaps significant that most of the outstanding vocal imitators are found among tropical or subtropical species."

Thorpe & North [\[88\]](#) give the example of a pair of birds which communicated via a 15 note antiphonal duet. However when one bird died the survivor resumed the performance of the whole - something it had never done previously! They note of another case of duetting, reported elsewhere, that *'when the partners were*

absent, the remaining bird would use the sounds normally reserved for his partner, with the result that the partner would return as quickly as possible, as if called by name'. This strongly suggests that we witness here a real case where song is used meaningfully in social communication as a bond. On top of that, the vocal tract of these birds has attained such sophistication, that it enables them to imitate human speech.

Music has bonding function in close relatives of ours as well. As noted above, male and female Siamang sing (the male bitonally and without melody; the female monotonously) to establish and maintain pair-bonding and the social recognition of their territory [38].

The cue of the use of song as a bond strengthening means of communication, rather than song being a trait which has evolved by sexual selection alone, itself leads to some intriguing remarks with regard to the special 'sociological' case humans are among primates (and animals/mammals in general). We know, by comparing the social nature of humans with other apes, that we too have evolved an unique capacity to bond with each other. Indeed, it is also in the depth and complexity of our bonding that humans differ (apart from language) from other primates. From these observations and considerations we are tempted to conclude that musicality not only can explain how symbolic language evolved, but also that song, as a means to aid bond formation, can help to explain how the characteristic sexual-social relationships between humans became possible (see also 4.3.2).

4.3. Present and past social music

4.3.1. Various observations

What evidence exists for the key role of music in the lives of humans? Below we give a very limited excerpt of the functions and possibilities in human social life. All human cultures possess lullabies and use them to sing children to sleep. The music business is among the world's major industries. Going to war is so much more fun with a drum band marching along. Dancing to music can give people mystical trance experiences. Music brings up deep emotions such as hope, pleasure, comfort or sadness, and probably no other 'art' can do this as profoundly as music. From observations of currently existing 'premodern' societies, it is clear that music (and its counterpart, dance) must have played an even more important, pivotal role in early human societies. Music has a role, not only in rituals, but also in many practical activities. For instance, Australian aboriginals memorize the look of landscapes in songs. Although the music making of early humans has left no physical remains, it must have been a major part of their lives, as it still largely is an essential part of our lives.

4.3.2. Music and group identity

There is the observation that rituals, dance and song enhance group identity. With respect to territorial behaviour, it should be noticed that singing is indeed used for that purpose in close relatives of ours: "*In addition to the well-known territorial bird songs, some monkey species and all species of lesser apes have territorial songs.*" [91].

From what we know about ourselves as apes, increasing group identity could have put strong evolutionary pressure on singing behaviour. To understand this we must digress upon what has recently been found about our uniqueness as social apes. Humans, chimpanzees and presumably our earliest shared ancestors mix a life-style of belonging to a group, while separating into smaller parties during much of the lives. This is called an atomistic or fission-fusion social existence [72]. We, however, do so in a way that is unique because the bonds are robust and long-termed and allow for long periods of separation. Biological parents in all human societies form bonds with each other (though not necessarily monogamous ones). People form life-long attachments with friends and distant kin. We, moreover, usually form a life-long attachment with our 'identity

group' from the level of our extended family to that of our nation and religion.

Early humans faced the paradoxical problem of relying for survival both on a group and on the recurrent need to split-up. Anthropologists and historians identify the mechanism by which people create and sustain the required social attachments with rituals and group activities involving synchronised song and dance [54]. The need for sustained social bonds may have further selected (after possibly initial sexual selection and selection for stronger pair bonding (see 4.2.2.3)), for dance and song competence [Note 6].

Modern remnants of this ancient function of music might be the supporters' songs of sports teams, songs of any kind of club (e.g. students), war music and the national hymns, closely linked to the notions of territory and group identity. Indeed, music, singing and dancing still plays the central role in social life of all extant original bands. Ceremonies, rituals, and many other group activities (work-gangs, parties, festivals) all exploit the strong emotions which come with the ensemble of vocalisations and movement. Just think of the emotional bonding, the sense of belonging, the experience of 'together we are invincible' that accompanies marching songs, football stadium chants, National Anthems, camp-fire songs, hymns, corals, etc.

Increasing group identity exists in a nonmusical form in the collective intoning and synchronisation of bodily movements in religious prayers, petitions, supplications, orisons and worship. In modern societies, such synchronisation offers people a temporary sense of belongingness. In most cultures, they form an important part of rituals, ceremonies and other shared enjoyments which result in the affective togetherness that creates and sustains a society's collective existence [10, 54, 78].

4.4. Conclusion

Whatever the role of early singing was (territorial marking, courting, pair bond maintenance, enhancing group identity) it is clear that singing, musicality and dance had an important role to play in human social interactions, and that consequently musicality is plausibly selected for by good old natural selection. The development of a complex organ like the *Homo sapiens* vocal tract then can be understood to have been developed by natural selection more easily than in case we have to hypothesize that this natural selection occurred on the basis for selection of better speech [21, 67]. Only later on, these vocal abilities were used for speaking, and this view coincides with the proposition of Gould & Lewontin [34] that language is a spandrel or an exaptation: language was possible because of a preadaptation which developed for other reasons. While singing is an innate capacity, an instinct, speaking is a possibility emerging from singing and increased mental representational capacities. We could better speak of the song instinct than of the language instinct.

Comparing the role of song in some tropical song birds and in the siamang, one is tempted to state that song co-evolved with pair bonding, and thereby also helps to explain how the intriguing social and sexual characteristics of human life evolved.

5. Melody and language learning by children

5.1. Introduction: the semantics of spoken syntax

Do humans have a language acquiring device as Chomsky has proposed?

Most students of language easily accept that semantics is about linking mental representations of objects and concepts to the symbolic lexicon that happens to be used by a language. However, when it comes to syntax, most linguists seem to assume that there is only linguistic syntax. Above we have argued that all higher animals

possess some universal mental, thinking syntax (see [2.2.2](#)), while on the other hand it is tremendously difficult to discover any universality among the amazing diversity of spoken syntaxes (see [3.1](#)).

The problem with spoken syntax therefore boils down to the same problem of linking lexicon to mental representation: semantic meaning must be given to spoken syntactic entities by linking them to the mental syntax. We think this approach has been overlooked by most students of language. Then one must wonder how this can be achieved, since spoken syntax can be any kind, while mental syntax can be supposed to be largely alike among humans - and basically even among higher animals.

5.2. Music and language development

One of the big mysteries in speech acquisition is how children identify words. While the words on this page are divided by spaces, spoken words are not. Before you can identify words you have somehow to identify where and when they start and end. Failure to solve this hard problem holds back artificial speech recognition. The earliest voice recognition programs required that people spoke words slowly and in isolation. We suggest, backed up by a growing research, that infants solve this problem by listening to the rhythmicity and to the melody of stresses and tones in speech.

Even before children are born, their brains are familiar with the sounds that will surround them after birth. Newborns prefer the voice of their mother over that of strangers [\[23\]](#). If a mother repeats a short story twice a day for the last six and half weeks of her pregnancy, her newborn child will prefer hearing it to one she did not [\[24\]](#). The womb is an acoustic filter that preserves the intonations of a mother's speech. Thus, the brain is learning to hear speech as a melody from long before birth.

This is supported by other work upon newborns. It has been shown that newborns can discriminate the rhythm of multisyllabic stressed words suggesting they are already sensitive to the word-rhythm [\[74\]](#). Moreover, newborns already prefer infant-directed prosody stressing speech (motherese) over adult-directed speech [\[18\]](#). Complementary to this, mothers expand the intonation contours of their speech to their child as soon as it is born [\[28\]](#). Such motherese compared to adult-directed speech has emphasized prosody, namely higher overall pitch, wider pitch excursions, broader pitch range, increased rhythmicity, slower tempo, longer word durations and increased amplitude. Newborns moreover can distinguish their own language from a foreign one, something which must be due to the unique, prosodic cues of a language [\[55\]](#). This suggests they are increasingly able to focus upon the unique intonation aspects of their 'mother' tongue.

Children's own vocalisations, it should be noted, also start to be affected by these intonations:

"A cross-cultural investigation of the influence of target-language in babbling was carried out. 1047 vowels produced by twenty 10-month-old infants from Parisian French, London English, Hong Kong Cantonese and Algiers Arabic language backgrounds were recorded in the cities of origin and spectrally analysed. ... Statistical analyses provide evidence of differences between infants across language backgrounds. These differences parallel those found in adult speech in the corresponding languages." [\[22\]](#).

There is also the observation of the tremendous similarity of pronunciation within a slang. We all know the phenomenon that one can easily recognize the region where one comes from. Many people never succeed in speaking properly the standard language because of an uneradicable accent, which indicates the thorough imprinting which occurs: we do not only acquire lexicon, we mimic intonation almost exactly from our environment [\[Note 7\]](#).

The previous paragraphs lead us to suggest that some auditory equivalents to Rizzolatti-cells (see [Note 1](#))

must exist. Rizzolatti-cells and equivalents may be an important cue to understanding mimicking, to link the behaviour of genetically encoded cells to copyable observable (visual, auditory) behaviour of animals.

Intonation provides cues to how words are structured in sentences [59]. Words are not said uniformly but are intonation phrased. Spotting this intonation structure facilitates children to grasp how words are syntactically put together. Children use the intonational cues that tend to identify word beginnings [62]. These cues vary with language: stress for example in English, syllable in French and mora in Japanese. Children in all these languages develop a sensitivity to the intonational beat provided by these cues that mark off word separation.

Let us take a famous Chomskyan example which relates to inversion of the word order of statements in order to turn these into a question. Children with English speaking parents readily adopt that 'The man is here.' becomes a question by reversal of noun and verb: 'Is the man here?'. But how does one turn the slightly more complex sentence: 'The man who is tall, is here.' into a question? One might expect a child, who has just mastered the simple example to place the first 'is' in front of the sentence, to say: 'Is the man who tall is here?'. But children never make this mistake. Do we need Chomskyan theory, borrowed from mathematics and logic? Linguists developed rather complicated theories (like X-bar theory) whereby humans use 'null' elements to cope with this and related problems (see Smith and Szathmary [83] for a brief explanation).

What if we adopted the answer that children simply hear which of the two verbs is the main verb. Say the complex sentence to yourself and listen how the intonation on the second 'is' is different from the first 'is'. Now, try to reverse intonations. It requires a little exercise to do so, since it is experienced as a very 'unnatural' (we should actually say 'uncultural') thing to do, which by itself provides circumstantial evidence on the importance and the strict use of intonation. Children just hear which verb is the one which goes along with the man, because of the intonation of the main verb. Remark that the pitch of the main verb in the complex sentence is exactly the same pitch the main verb carries in the simpler sentence AND in the question. Once this has been acquired, children can generalize this principle to any similar sentence they meet. The intonation recognition capacity is one which stems from our innate musicality (naturally selected recently for other reasons than language itself), while the generalization capacity is part of the mental syntax capacity which we have inherited from animals (naturally selected for still other reasons). Bringing the two together one can have something like syntactic symbolic language.

Thus, children start off experiencing language as a kind of music. Parents and others respond to this sensitivity by making their language to them more musical - motherese. The rhythms of speech, which are heightened by motherese, provide the child with a means to use their sense of rhythm to spot the words and sentence structure. Memetic ontology thus replicates memetic phylogeny. In other words: music is both the answer to the phylogenetic and to the developmental origin of language.

Children, before acquiring the language spoken around them can distinguish phonetic categories of foreign languages they have not heard [27, 86, 90], only to lose this ability at around ten months [4]. One wonders why children should have this ability, in case language was naturally selected for, since this would require only the evolution of recognition of a limited phonological set. While explaining this from a 'natural selection for language' point of view is a real conundrum, it becomes triviality when adopting an innate sensitivity for melodizing.

Also, there are the numerous reports on the application of Music Intonation Therapy [2] to treat language disorders, as is exemplified by the quotes below:

"In order to develop a useful communication system, a 3-year-old, non-verbal autistic boy was treated for 1 year with a Simultaneous Communication method involving signed and verbal language. As this procedure

proved not useful in this case, an adaptation of Melodic Intonation Therapy (signing plus an intoned rather than spoken verbal stimulus) was tried. With this experimental language treatment, the patient produced trained, imitative and, finally, spontaneous intoned verbalizations which generalized to a variety of situations." [56]

"We examined mechanisms of recovery from aphasia in seven nonfluent aphasic patients, who were successfully treated with melodic intonation therapy (MIT) after a lengthy absence of spontaneous recovery." [7].

"In patients with brain lesion, a pre-verbal, emotionally-focussed tonal language almost invariably is capable of reaching the still healthy sections of the person. Hence, it is possible for music therapy to both establish contact with the seemingly non-responsive patient and re-stimulate the person's fundamental communication competencies and experience at the emotional, social and cognitive levels." [43].

Furthermore, it has been shown that music is not only important for developing linguistic skills, but also serves as a memory aid [65] and plays a role in the development of motoric skills [12].

Strong suggestions for the existence of a music acquiring device comparable to the hypothetical Chomskyan language acquiring device have been made by others. We claim that this MAD is our LAD:

"Full-term infants' performance in detection of melodic alterations appeared to be influenced by perceptual experience from 6 months to 1 year of age, and an experiment with infants born prematurely supported the hypothesis that experience affects music processing in infancy. These findings suggest parallel developmental tendencies in the perception of music and speech that may reflect general acquisition of perceptual abilities for processing of complex auditory patterns." [51].

"This indicates the existence of a partly innate and partly acquired competence to judge what is acceptable and what is not, within the tradition of Western popular or classical music. This seems to indicate the existence of some deep structure of tonality, comparable with Chomsky's deep language structure. Asians who have not been much exposed to this kind of music find the task very difficult." (Kalmus & Fry [44], reporting on experiments whereby subjects were asked to evaluate some characteristics of Western classic music).

The last sentence from the previous quote is again a strong indication for the importance of the tonality of the language and the music of a child's culture in moulding its innate recognition capacities. Depending on the culture, one's experience of what sounds acceptable and what is not, is completely different (by itself again an indication against a universal spoken grammar and natural selection of language). This is nicely illustrated by the fact that (Western) MIT therapy has to be adopted when it is used in an Asian country. When applying MIT for use with Japanese patients, the authors report that basic changes were necessary, because of the completely different 'pitch' of Japanese language [79].

Furthermore, Simmons & Baltaxe [80], studying adolescent autistics with linguistic impairments, suggested that:

"... perception of prosodic features may be crucial for decoding and encoding linguistic signals. Autistic children may be lacking in this ability."

6. Final conclusions

6.1. Summary

We argue that a combined genetic and memetic explanation is needed to understand what language is about and how it developed originally and develops with almost every new human.

According to the point of view presented here, symbolic, spoken language emerges from the (coincidental) combination of complex representational capacity with intonation recognition/reproduction capacity (which itself develops in close connection with singing capacity). As such, it is claimed that it is not language itself which has been naturally selected for. Language is considered as a cultural phenomenon very well comparable to bird song culture, only more sophisticated (variable, flexible, more symbolic, syntactic) just because of the more sophisticated mental representation capacities of higher apes. In summary, birds did not develop symbolic language to the extent that humans did, because of more limited mental representation capacities, chimpanzees did not because of lack of singing capacities. Humans simply happened to combine both characteristics. How language then can develop by memetic evolution, might partially be answered by work presently being done with interacting robot agents [\[84\]](#), and is the subject of further work [\[82\]](#).

Once the preference for sound variety has been selected for, something which may happen for various reasons and which has occurred independently in different animal taxa, individuals which can produce any kind of primitive song may be reproductively more successful through sexual selection. Moreover, the group of singing and dancing individuals as a whole, whatever the genetic make-up of the individuals, may become more successful because of the increased group identity awareness which makes its members cooperate more efficiently or which may make the members lose their individuality to some degree, resulting e.g. in more fierce, aggressive behaviour with regard to non tribe members. Indeed, another typical characteristic of humans is our long tradition of warfare and genocide [\[25\]](#).

6.2. How does this approach compare to the hypotheses put forward by Chomsky [\[16, 17\]](#), Pinker [\[67\]](#) and Deacon [\[21\]](#)?

With respect to the development of language in children, one can agree with Chomsky that humans have special abilities to adopt language and syntax very spontaneously early in childhood and this can be called an innate language acquiring device. Still, it probably might best be understood as an innate music acquiring device, which enables to link any possible syntax of spoken language - the one used by the adults which happen to raise the child or by other children which happen to grow up with the child - to the universal mental syntax, of which we share the general basic possibilities for categorization and for generalization of causal rules with animals.

We do not agree with the Chomskyan suggestion, taken for granted by Pinker [\[67\]](#), but thoroughly criticized by e.g. Allott [\[3\]](#), Deacon [\[21\]](#) and Tomasello [\[89\]](#), that there is such a thing as universal linguistic grammar.

Furthermore, the explanation of the origin of language in evolution and during individual development, as proposed here, has nothing in common with the adaptationist explanations of Pinker (see [3.2](#)). Not only Allott [\[3\]](#) and Tomasello [\[89\]](#) point to different shortcomings of this kind of reasoning, but also Deacon [\[21\]](#) has clearly indicated several flaws. Several other criticisms are possible [\[82\]](#). What Pinker [\[67\]](#) calls a 'boring conclusion', is simply a completely erroneous conclusion.

We can largely agree with Deacon [\[21\]](#) that we are a symbolic species, and his evolutionary reasoning is much more relevant than that of Pinker. However, Deacon [\[21\]](#), like Pinker [\[67\]](#), relies on long term (2

million years) gradual evolution through selective advantage offered by the use of symbols, while instead proposing Baldwinian evolution (evolution by genetic assimilation of behavioural characteristics).

Both our approach and - to a certain degree (because of the pivotal role of symbolic gestures and sounds) - that of Deacon could be called 'memetic'. The difference is that in Deacon's approach gestures and symbolic sounds come into play already 2 million years ago (at the stage of *Homo habilis*) and reshape the brain by genetic assimilation. In our approach natural selection for better general mental abilities and, only recently (possibly with the advent of *Homo sapiens sapiens*), natural selection for musicality explains the reshaping of the brain and the vocal tract and we claim that it is from the combination of increased intelligence and vocal flexibility that language emerges as a cultural process, while we dismiss natural selection or Baldwinian evolution guided by the advantages brought along by the linguistic capacity - as is proposed by either Pinker [67] or Deacon [21].

Once humans combined mental capacity and musicality, we rely on genetically encoded flexibility of the brain to explain how symbolic sounds - memes - could develop and restructure brain mapping in a nongenetically inheritable manner. In other words, genes provide general capacities like brain flexibility, vocal dexterity, intonation recognition and reproduction capacity, while memes - through interaction with the developing brain - strongly influence the rewiring of the neuronal connections which make up a brain.

Although we date the influence of symbolic sounds much later than Deacon [21], we claim that once they originate, further changes occur in an almost purely memetic manner. The example below of the differences between literate and illiterate persons indicates how influential the means of communication are with respect to our mental abilities.

Our musical language origin theory coincides best with "the idea that removal of vocal limitations released untapped linguistic abilities which has been a major theme of a number of language origin theories (most notably argued by Philip Lieberman, in a number of influential books and articles)[49, 50]" (quote from Deacon [21], page 354). Deacon however considers this as an oversimplification and states that: "... *the development of skilled vocal ability was almost certainly a protracted process in hominid evolution, not a sudden shift.*" ([21], page 354), whereupon we disagree, backed up by the archeological record (see 4.1). Our hypothesis provides strong support for the insights of Lieberman [49, 50] (see also the [postscript](#)).

6.3. Why has the musical origin of language hypothesis been overlooked?

There is a further intriguing question, in case our hypothesis - which we will defend also on grounds of a more linguistic and neurolinguistic approach [82] - turns out to be a major key in understanding the origin of language. Indeed, one keeps wondering why this obvious, straightforward, and with hindsight even trivial approach to explaining the origin of language has been overlooked by linguists during the last decades. This is even more astounding, first because some of the earliest theories posed that musicality had to lay at the origin of language [19, 73, 94] - even Darwin [19] pointed to the resemblance and second because the importance of rhythm, intonation, melody, etc. in every day life, in language therapy and in child language (as briefly reviewed above) is so overwhelming, and is well studied.

Several explanations can be thought of. First, there is of course the adaptationist paradigm which keeps us thinking in terms of function, usefulness, and which makes us overlook that usefulness is a posthoc consideration which can only serve as an explanation once the necessary events leading to the existence of some characteristics have taken place. Natural selection can explain why something still exists, but not how it came into being. The necessary variation is not a matter of natural selection, it is a matter of contingency, coincidence, mutation, recombination, symbiosis, evolution of characteristics for other reasons than the ones for which they eventually are useful now (preadaptation, exaptation).

Second, and closely linked to the previous considerations, there is the fact that we all are impressed by the explanatory power of natural selection of genetic characteristics in general, which makes us forget that natural selection is just a special case of selection (see [Note 2](#)). Therefore, there is a tradition of trying to explain everything with genes only.

Third, with respect to language, another important bias may exist. It appears that most linguists depart for their considerations from the present form of language, which needs a sophisticated grammar because much of communication is in the form of written code, which lacks the intonation characteristic of spoken language. E.g. writing down a joke may be experienced as an insult instead of as the tongue in cheek remark it was meant to be. In oral communication this will in most instances be clear, because of the facial expression and the intonation. Using written code, we need question marks, exclamation marks or ": -) " (the smile-sign as used in e-mail discussions) to indicate that what we write is meant as a question, an important remark or a joke. Written code, lacking intonation and eye-contact, compensates grammatically for the absence of a shared context with the listener, and finally influences more and more the way we speak, as becomes clear from studies comparing cognitive linguistic capacities between literates and illiterates.

Illiterates - when compared with literates of the same background - have been found to show cognitive difficulties in nonreading tasks such as phoneme awareness [\[8, 61\]](#), repeating nonwords (phoneme sequences that do not pronounce a familiar word) [\[60\]](#), memorising pairs of phonologically related words compared to semantically related ones, and difficulties in generating words which start with a common phoneme sound or which are the names of animals or furniture [\[69\]](#). Several other studies lead to the suggestion that learning to read and write might not only challenge how people process oral language but also does change the organisation of people's brains [\[14, 95\]](#). This was already suggested upon nonpsychological and nonneurological grounds [\[64\]](#).

However, most linguists start from the current situation (a literate world) and extrapolate and/or impose our way of thinking, living, interacting, communicating to the illiterate societies in which the original humans lived at the time language originated (see [3.2.5](#) for a comparable bias), thereby forgetting how different we are because of the completely different memes which populate our brains and because of the fact that the environment we have to cope with is incomparably different to the natural environments in which language first evolved.

6.4. Could large vocabularies alone be sufficient for the development of syntax?

It is important to quote here recent work of Bates & Goodman [\[6\]](#), which indicates that syntax abilities parallel very tightly vocabulary size over a wide variety of ages. Thus, though children may vary widely with respect to the size of the vocabulary at a certain age (some children acquire words more easily than others), the degree of grammatical competence they acquire is strictly linked to the lexical stage at which they are. This means that two children - one 3-year-old and one 5-year-old, but each with a vocabulary of 200 words, will have both the same stage of syntax.

Bates & Goodman [\[6\]](#) point out the implications of this for language in chimps. Chomskyans make it a slogan that 'animals cannot learn grammar' and hence that 'grammar is unique to the human species'. Bates points out however that chimps taught language in fact attain the level of syntactical competence you would expect from human children with the same size of vocabulary. Bates and Goodman [\[6\]](#) state that, if chimps lack syntax, it is not because they lack a human competence for syntax, but because their vocabularies are too limited.

This becomes apparent from the following quote:

"These differences between grammar and vocabulary are usually interpreted to reflect a qualitative difference in the language-learning abilities of non-human primates (that is, they have lexical abilities, but they lack a 'grammar acquisition device'). That may well be the case; after all, they are not human. However, the data that we have presented here suggest another interpretation: Because the animals studied to date apparently find it difficult to produce more than 200-300 words, symbols or signs, we should not be surprised to find that they also have very restricted abilities in expressive grammar. Consider the developmental relationship between grammar and vocabulary size that we have observed in human children. From these figures, it is clear that children with vocabularies under 300 words have very restricted grammatical abilities: some combinations, a few function words in the right places, the occasional bound morpheme, but little evidence for productive control over morphology or syntax. Viewed in this light, the difference between child and chimpanzee may lie not in the emergence of a separate grammar 'module', but in the absolute level that they are able to attain in either of these domains. Chimpanzees do not attain the 'critical mass' that is necessary for grammar in normal children; instead, they appear to be arrested at a point in lexical development when grammar is still at a very simple level in the human child. Hence, the putative dissociation between lexical and grammatical abilities in nonhuman primates may be an illusion".

From these considerations, it appears that to explain the rise of syntax, the problem is not how to explain any 'syntax' module arose peculiar to humans, but the problem is to explain why large vocabularies arose. If you can explain that, you can explain the rise of syntax. The solution to the problem of how a large vocabulary could arise, follows from what we suggest: humans are originally musical primates. Once humans gained the neurological abilities to control vocalisation needed to sing, they gained the abilities to create vast vocabularies of words. Although a large vocabulary on its own may be sufficient for syntactical ability to develop, as Bates & Goodman [6] suggest, we think that it helps when you have a MAD, a well developed intonation recognition/reproduction device, at your disposal. Musical ability may explain the rise of a large vocabulary and at the same time may be an extra gain to create and acquire linguistic grammar.

A 'musical origin of language' theory enables to bring together the ideas of Deacon [21], Lieberman [49, 50] and Bates & Goodman [6] (among many others). One could say that at some point, quantity (increased intelligence/mental syntax capacity, increased vocal flexibility, increased vocabulary) may change into (or emerge as) quality (linguistic syntactic ability). The basic difference between humans and animals then can be explained almost exclusively by the usage of symbolic/syntactic language. Of course, the explosive cultural evolution which became possible - once symbolic information processors like modern human brains arose - at first sight justifies the claim that at least one qualitative difference must distinguish humans from animals. It should be kept in mind that a minor additional trick sometimes can make a large difference. Moreover, one of us has previously briefly argued that the widely spread human need to claim human uniqueness can itself be explained from the need for continued self confirmation, which again follows from adding symbolic memes to the emotional - animal - being we are in the first place [92].

6.5. Musicality may also explain other typically human characteristics

Finally, it should be emphasized again that song as a powerful means for pair bonding, as it appears to function in some animal species, can very easily explain another intriguing and far reaching characteristic of (modern?) humans. Human musicality can explain how the typically strong human pair bonding could have evolved. As such, song could explain not only speech, but also could help to understand the typically sexual and social behaviour of humans.

Postscript.

After this manuscript was accepted for publication, Lieberman [Lieberman, D.E. 1998. Sphenoid shortening and the evolution of modern human cranial shape. *Nature* 393: 158-162] argued to consider *Homo sapiens sapiens* (modern man) as a separate species from '*H. sapiens neanderthalensis*', because of clear facial differences with other hominids, incl. neanderthals. Lieberman suggests that these changes may be related to the ability of speech. These considerations coincide with the claims - embraced in this article (see [4.1](#)) - for a late origin of language, while the essential facial morphological characteristics of modern man may have been selected for by singing ability, enabling speech, but not for speech.

Acknowledgements

MV is indebted to the FWO Flanders for an appointment as a research director.

Notes

Note 1

In essence, the original memes (as used among animals) can be defined as behaviours which can be mimicked. Dawkins [\[20\]](#) referred to bird songs as memes. However, one reviewer remarked that only humans can imitate in an observable manner. If only this kind of conscious imitation counts for memes, than only humans produce memes and washing sweet potatoes by Japanese macaques (see [3.2.4](#)) would not be caused by imitation and thus not be memetic. One may object that there is strong evidence for unconscious imitation underlying learning in animals, as becomes apparent from the work of Rizzolatti et al. [\[71\]](#):

"In area F5 of monkey premotor cortex there are neurons that discharge both when the monkey performs an action and when he observes similar actions made by another monkey or by the experimenter. We report here some of the properties of these 'mirror' neurons and we propose that their activity 'represents' the observed action. We posit, then, that this motor representation is at the basis of the understanding of motor events."

Finally, it should be noted that conscious imitation itself might be a secondary consequence of the development of language, which makes possible reflexive awareness. If one could show that conscious imitation is a consequence of reflexive awareness (i.e., consciousness), this kind of imitation could be considered itself largely as explained once one has explained language.

Note 2

It is essential here to reflect on the definitions of selection and natural selection. Selection is a general principle: whenever there is variation on a theme, selection by the environment will occur, since none, one, more or all variations (configurations) may fit for existence in this environment. Natural selection is a special case which follows from the fact that selection takes place among variants on the theme of self-replicating systems, i.e. cells. The survival of the information processor (the cellular enzymatic machinery) is intrinsically linked to the information itself and vice versa. While differential survival of the information processors (the cells and the multicellular colonies) determines the reproductive success of the information molecules (the genes), the (genetic) information in turn determines the survival rate and reproductive success of the

information replicators.

We could speak of a closed semantic circle (present in a metabolically open system).

However, in cultural-memetic selection, the information processors (animals, humans, copy machines, presses, computers) can die or stop functioning while the instantiations of information (memes, habits, knowledge) continue to flourish, and vice versa some instantiations of information can be lost or gained - for different reasons - without influencing the survival and/or activity of the information processors. As such, selection of behavioural/memetic/cultural information is basically different from the 'special case' of natural selection, although the general principles of evolution (change over time) and selection can be applied.

Note 3

Consider the following experiments:

A chimpanzee named Panzee first saw a keeper hide food in one of two locked boxes. When a second keeper entered, Panzee learned to point the second keeper in which of two cages the food was hidden in order to obtain the food. The next experiment however seems definite proof of the fact that the chimp knows which knowledge is in the mind of the attendants and which knowledge it should add to get the food: keeper 1 hides the food, locks the box and gives the key to keeper 2, while leaving. After keeper 1 left, keeper 2 hides the key and leaves. Keeper 1 then returns without knowing where the key is hidden. If the chimp had learned by trial and error alone, she would still point to the box where the food was hidden. Instead, on her first try, she pointed to where the key was hidden. The chimp showed she could fathom the working of another mind: she knew that keeper 1 did not know where the key was.

(after Mills [\[57\]](#))

Note 4

This leads to the remark that 'The Language Instinct' as the title of a book claiming a Darwinian approach to the problem of the nature and the origin of language, would have been disapproved by Darwin [\[19\]](#) himself.

Note 5

There might be some other resemblance between the song capacities of song birds and humans, although this is not really essential to the hypothesis put forward here. The front limbs in birds have been specially adapted for repetitive motor behaviour, flight, and Calvin [\[15\]](#) has proposed that special motoric capacities in humans, through e.g. natural selection for better throwing capacities, led to increased brain capacities in humans. Analogously, song birds are among the most intelligent birds. However, Calvin [\[15\]](#) and/or others seem to claim that these motor capacities by themselves are sufficient explanation for the linguistic capacities of humans, while it is argued here that these were only preadaptations which enabled singing, which itself then forms the essential preadaptation to speech. Thus, one could propose that for birds the flying capacity was a useful preadaptation for the possibility of song capacity, like for humans specific motoric capacities - needed for e.g. throwing - prepared for the possibility of singing.

Note 6

We focus on song here, because the aim of this paper is linking it to speech. However it is clear that song and dance go together. Many societies are known not to distinguish song from dance [\[78\]](#). In most circumstances where singing and dancing have not been professionalised and so are done by all members of a group, when

people sing they dance (or make other collective bodily movements), and when they dance they sing. Dance does not require vocal control but it can be suggested that the processes which modulate vocalisation are not restricted purely to the vocal tract but extend to incorporate other aspects of the body. Indeed, research indicates a close linkage between speech and gestures [29]. We suggest that part of the evolution of vocal modulation included the ability to incorporate with vocalisation other patterns of movement.

Note 7

With respect to the 'environment', it should be noted in passing that children learn more readily from other children than from their parents and that they are more profoundly influenced by the habits (including language) of other children than by the habits of their parents (personal observations). A possible reason may be that they need to adopt the behaviours and habits of their play mates to get accepted in this social group.

References

- [1] Aitchison, J. 1997. *The Seeds of Speech, Language Origin and Evolution*. Cambridge University Press.
- [2] Albert, M.L., R.W. Sparks, and N.A. Helm. 1973. Melodic intonation therapy for aphasia. *Archives of Neurology* 29: 130-131.
- [3] Allott, R. 1997. Pinker's language instinct: gradualistic natural selection is not a good enough explanation. <http://www.percep.demon.co.uk/pinker.htm>
- [4] Aslin, R.N., D.B. Pisoni, B.L. Hennessy, and A.J. Perey. 1981. Discrimination of voice onset time by human infants: new findings and implications for the effects of early experience. *Child Development* 52: 1135-1145.
- [5] Baker, R.R. 1996. *Human Sperm Competition: Copulation, Masturbation and Infertility*. Chapman & Hall, London.
- [6] Bates, E., and J. Goodman. 1997. On the inseparability of grammar and the lexicon. *Language and Cognitive Processes* 12: 507-586. (The paragraph quoted about chimpanzee language is pp. 544-545.)
- [7] Belin, P. P. Van Eeckhout, M. Zilbovicius, P. Remy, C. Francois, S. Guillaume, F. Chain, G. Rancurel, and Y. Samson. 1996. Recovery from nonfluent aphasia after melodic intonation therapy: a PET study. *Neurology* 47: 1504-1511.
- [8] Bertelson, P., B. de Gelder, L.B. Tfouni, and J. Morais. 1989. Metaphonological abilities of adult illiterates: New evidence of heterogeneity. *European Journal of Cognitive Psychology* 1: 239-250.
- [9] Bickerton, D. 1990. *Language and Species*. University of Chicago Press, Chicago.
- [10] Blacking, J. 1974. *How Musical is Man?* University of Washington Press, Seattle.
- [11] Botha, R.P. 1989. *Challenging Chomsky*. Blackwell, Oxford.
- [12] Brown, J., C. Sherrill, and B. Gench. 1981. Music may have importance to the development of general motoric skills rather than to language alone: effects of an integrated physical education/music program in changing early childhood perceptual-motor performance. *Perception and Motor Skills* 53: 151-154.

- [13] Burling, R. 1993. Primate calls, human language, and nonverbal communication. *Current Anthropology* 34: 25-53.
- [14] Castro-Caldas, A., K.M. Petersson, A. Reis, S. Stone-Elander, and M. Ingvar. In press. The illiterate brain: learning in childhood determines the functional organization of the adult brain.
- [15] Calvin, H. 1983. A stone's throw and its launch window: timing, precision, and its implications for language and hominid brains. *Journal of Theoretical Biology* 104: 121-135.
- [16] Chomsky, N. 1957. *Syntactic Structures*. Mouton, The Hague.
- [17] Chomsky, N. 1988. *Language and the Problems of Knowledge*. MIT, Cambridge.
- [18] Cooper, R.P., and R.N. Aslin. 1990. Preference for infant-directed speech in the first month after birth. *Child Development* 61: 1584-1595.
- [19] Darwin, C. 1871. *The Descent of Man*. Murray, London.
- [20] Dawkins, R. 1976. *The Selfish Gene*. Oxford University Press, Oxford.
- [21] Deacon, T.W. 1997. *The Symbolic Species: The Co-evolution of Language and the Brain*. W.W. Norton & Cy, New York.
- [22] De Boysson-Bardies B., P. Halle, L. Sagart, and C.J. Durand. 1989. A crosslinguistic investigation of vowel formants in babbling. *Child Language* 16: 1-17.
- [23] DeCasper, A.J., and W.P. Fifer. 1980. Of human bonding: newborns prefer their mother's voices. *Science* 208: 1174-1176.
- [24] DeCasper, A.J., and M.J. Spence. 1986. Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior and Development* 9: 133-150.
- [25] Diamond, J.M. 1997. *Guns, Germs, and Steel: The Fates of Human Societies*. Norton, New York.
- [26] Diamond, J.M. 1997. The language steamrollers. *Nature* 389: 544-546.
- [27] Eimas, P.D., J.L. Miller, and P.W. Jusczyk. 1987. On infant speech perception and the acquisition of language. Pp. 161-195, In *Categorical Perception* (S. Harnad, Ed.). Cambridge University Press, New York.
- [28] Fernald, A., and T. Simon. 1984. Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology* 20: 104-113.
- [29] Feyereisen, P. and J.-D. de Lannoy. 1991. *Gestures and Speech: Psychological Investigations*. Cambridge University Press, Cambridge.
- [30] Fisher, S.E., F. Vargha-Khadem, K.E. Watkins, A.P. Monaco, and M.E. Pembrey. 1998. Localisation of a gene implicated in a severe speech and language disorder. *Nature Genetics* 18: 168-170.
- [31] Geschwind, N. 1980. Some comments on the neurology of language. In: *Biological Studies of Mental Processes*, ed., D. Caplan. MIT Press.

- [32] Gopnik, M. 1990. Feature-blind grammar and dysphasia. *Nature* 344: 715.
- [33] Gopnik, M., and M.B. Crago. 1991. Familial aggregation of a developmental language disorder. *Cognition* 39: 1-50.
- [34] Gould, J., and R.C. Lewontin. 1979. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist program. *Proceedings of the Royal Society of London* B205: 281-288.
- [35] Grant, P.R., and B.R. Grant. 1997. Genetics and the origin of bird species. *Proceedings of the National Academy of Sciences USA* 94: 7768-7775.
- [36] Greenfield, P.M. 1991. Language, tools and brain: the ontogeny and phylogeny of hierarchically organized sequential behaviour. *Behavioral Brain Science* 14: 531-595.
- [37] Griffin, D.R. 1992. *Animal Minds*. University Chicago Press, Chicago.
- [38] Haimoff, E.H. 1981. Video analysis of siamang (*Hylobates syndactylus*) songs. *Behaviour* 76: 128-151.
- [39] Harris, R.A. 1993. *The Linguistic Wars*. Oxford University Press, Oxford.
- [40] Herman, L.M., D.G. Richards, and J.P. Wolze. 1984. Comprehension of sentences by bottlenosed dolphins. *Cognition* 16: 129-219.
- [41] Jespersen, O. 1922. *Language, its Nature, Development and Origin*. Allen & Unwin, London.
- [42] Jonker A. 1987. The origin of the human mind. A speculation on the emergence of language and human consciousness. *Acta Biotheoretica* 36: 129-77.
- [43] Jochims, S. 1994. Establishing contact in the early stage of severe craniocerebral trauma: sound as the bridge to mute patients. *Rehabilitation (Stuttg)* 33: 8-13.
- [44] Kalmus, H., and D.B. Fry. 1980. On tune deafness (dysmelodia): frequency, development, genetics and musical background. *Annals Human Genetics* 43: 369-382.
- [45] Kluender, K.R., R.L. Diehl, and P.R. Killeen. 1987. Japanese quail can learn phonetic categories. *Science* 237: 1195-1197.
- [46] Kuhl, P., and J. Miller. 1975. Speech perception by the chinchilla: Voice-voiceless distinction in alveolar plosive consonants. *Science* 190: 69-72.
- [47] Lee, R.B. 1979. *The !Kung San: Men, Women, and Work in a Foraging Society*. Cambridge University Press, Cambridge.
- [48] Lehrdahl, F., and R. Jackendoff. 1983. *A Generative Theory of Tonal Music*. MIT Press, Cambridge, MA.
- [49] Lieberman, P. 1984. *The Biology and Evolution of Language*. Harvard University Press, Cambridge, Ma.
- [50] Lieberman, P. 1991. *Uniquely Human: the Evolution of Speech, Thought and Selfless Behaviour*. Harvard University Press, Cambridge, Ma.

- [51] Lynch, M.P., L.B. Short, and R. Chua. 1995. Contributions of experience to the development of musical processing in infancy. *Developmental Psychobiology* 28: 377-398.
- [52] Maddieson, I. 1981. UPSID UCLA phonological segment inventory database. *UCLA Working Papers in Phonetics* 53: 1-243.
- [53] Matsuzawa, T. 1991. Nesting cups and metatools in chimpanzees. *Behavioral Brain Science* 14: 570-571.
- [54] McNeill, W.H. 1995. *Keeping together in time: dance and drill in human history*. Harvard Univ. Press.
- [55] Mehler, J., P. Jusczyk, G. Lambertz, N. Halsted, J. Bertoncini, and C. Amiel-Tison. 1988. A precursor of language acquisition in young infants. *Cognition* 29: 143-178.
- [56] Miller, S.B., and J.M. Toca. 1979. Adapted melodic intonation therapy: a case study of an experimental language program for an autistic child. *Journal Clinical Psychiatry* 40: 201-203.
- [57] Mills, C. 1997. Unusual suspects. *The Sciences*, July/August: 32-36.
- [58] Milo, R.G, and D. Quiatt. 1993. Glottogenesis and anatomically modern *Homo sapiens*: The evidence for and implications of a late origin of vocal language. *Current Anthropology* 34: 569-598.
- [59] Morgan, J.L. 1996. Prosody and the roots of parsing. *Language and Cognitive Processes* 11: 69-106.
- [60] Morais, J. 1987. Phonetic awareness and reading acquisition. *Psychological Research* 49: 147-152.
- [61] Morais, J., L. Cary, J. Alegria, and P. Bertelson. 1979. Does awareness of speech as a sequence of phones arise spontaneously? *Cognition* 7: 323-331.
- [62] Nelson, D.G., K. Hirsh-Pasek, P.W. Jusczyk, and K.W. Cassidy. 1989. How the prosodic cues in motherese might assist language learning. *Journal of Child Language* 16: 55-68.
- [63] Noble, W., and I. Davidson. 1996. *Human Evolution, Language and Mind: A Psychological and Archaeological Inquiry*. Cambridge University Press, Cambridge.
- [64] Ong, W.J. 1982. *Orality and Literacy: The Technologising of the Word*. Methuen, London.
- [65] Peretz, I., M. Babai, I. Lussier, S. Hébert, and L. Gagnon. 1995. Musical excerpts: indices relating to familiarity, age of acquisition and verbal associations. *Canadian Journal Experimental Psychology* 49: 211-239.
- [66] Pinker, S., and P. Bloom. 1990. Natural language and natural selection. *Behavioral and Brain Sciences* 13: 707-784.
- [67] Pinker, S. 1994. *The Language Instinct: How the Mind Creates Language*. Morrow, New York.
- [68] Provine, R.R. 1996. Laughter. *American Scientist* 84: 38-45.
- [69] Reis, A., and A. Castro-Caldas. 1997. Illiteracy: A cause for biased cognitive development. *Journal of the International Neuropsychological Society* 3: 444-450.

- [70] Ridley, M. 1993. The units of selection, Chapter 12. Pp. 303-322, In *Evolution*. Blackwell, Boston.
- [71] Rizzolatti, G., L. Fadiga, V. Gallese, & L. Fogassi. 1996. Premotor cortex and the recognition of motor actions. *Brain Res. Cogn. Brain Res.* 3: 131-141.
- [72] Rodseth, L., R.W. Wrangham, A.M. Harrigan, and B.B. Smuts. 1991. The human community as a primate society. *Current Anthropology* 32: 221-254.
- [73] Rousseau, J.J. 1852/1966. On the Origin of Languages. In *On the Origin of Language (translated essays of Rousseau)* (J.H. Moran, and A. Gode, eds). University of Chicago Press, Chicago.
- [74] Sansavini, A., J. Bertoncini, and G. Giovanelli. 1997. Newborns discriminate the rhythm of multisyllabic stressed words. *Developmental Psychology* 33: 3-11.
- [75] Sataloff, R.T. 1992. The human voice. *Scientific American* 267: 64-71.
- [76] Savage-Rumbaugh, E.S., and R. Lewin. 1994. *Kanzi: the Ape at the Brink of the Human Mind*. John Wiley, New York.
- [77] Seckel, H.P. 1960. *Bird-headed Dwarfs*. Karger, Basel.
- [78] Seeger, A. 1994. Music and dance. Pp. 686-705, In T. Ingold, (Ed.), *Companion Encyclopedia of Anthropology*. Routledge, London.
- [79] Seki, K., and M. Sugishita. 1983. Japanese-applied melodic intonation therapy for Broca aphasia. *No To Shinkei* 35: 1031-1037.
- [80] Simmons, J.Q., and C. Baltaxe. 1975. Language patterns of adolescent autistics. *Journal Autism & Child Schizophrenia* 5: 333-351.
- [81] Skoyles, J.R. In press. Human evolution expanded brains to increase expertise capacity, not IQ. <http://www.users.globalnet.co.uk/~skoyles/brain.htm>.
- [82] Skoyles, J.R., and M. Vaneechoutte. In preparation.
- [83] Smith, J.M., and E. Szathmari. 1995. *The Major Transitions in Evolution*. Freeman, Oxford.
- [84] Steels, L. 1996. Interacting robot agents. Synthesising the origins of language and meaning using co-evolution, self-organisation and level formation. <http://www.heise.de/bin/tp-issue/tp.html?artikelnr=6001&mode=html>
- [85] Steinscheider, M., J. Arezzo, and H.G. Vaughan. 1982. Speech evoked activity in the auditory radiations and cortex of the awake monkey. *Brain Research* 252: 353-365.
- [86] Streeter, L.A. 1976. Language perception of 2-month-old infants shows effects of both innate mechanisms and experience. *Nature* 259: 39-41.
- [87] Symons, D. 1979. *The Evolution of Human Sexuality*. Oxford University Press, Oxford.
- [88] Thorpe, W.H., and M.E. North. 1966. Vocal imitation in the tropical bou-bou shrike *Lanarius aethiopicus major* as a means of establishing and maintaining social bonds. *Ibis* 108: 432-435.

- [89] Tomasello, M. 1995. Language is not an instinct. *Cognitive Development* 10: 131-156.
- [90] Trehub, S.E. 1976. The discrimination of foreign speech contrasts by infants and adults. *Child Development* 47: 466-472.
- [91] Ujhelyi, M.J. 1996 Is there any intermediate stage between animal communication and language? *Theoretical Biology* 180: 71-76.
- [92] Vaneechoutte, M. 1993. The memetic basis of religion. *Nature* 365: 290. At <http://www.sepa.tudelft.nl/webstaf/hanss/nature.htm>
- [93] Vaneechoutte, M. 1997. Bird song as a possible cultural mechanism for speciation. *Journal of Memetics-Evolutionary Models of Information Transmission* 1. At http://www.cpm.mmu.ac.uk/jom-emit/1997/vol1/vaneechoutte_m.html
- [94] von Humboldt, W. 1836/1988. *On Language: the Diversity of Human Language Structure and its Influence on the Mental Development of Mankind*. Cambridge University Press, Cambridge.
- [95] Zatorre, R.J., E. Meyer, A. Gjedde, and A.C. Evans. 1996. PET studies of phonetic processing of speech: Review, replication, and reanalysis. *Cerebral Cortex* 6: 21-30.

© JoM-EMIT 1998

Additional links

<http://www.livescience.com/culture/091105-baby-language.html>: Children listen to melody, already in the womb.

http://scienceblogs.com/notrocketscience/2009/03/why_music_sounds_right_-_the_hidden_tones_in_our_own_speech.php

[Cross, I.](#) 1999. Is music the most important thing we ever did ? Music, development and evolution. In *Music, Mind and Science*, Ed. Suk Won Yi, Seoul: Seoul National University Press.

http://scienceblogs.com/notrocketscience/2009/03/why_music_sounds_right_-_the_hidden_tones_in_our_own_speech.php

<http://www.oxy.edu/departments/psych/CHAPMAN/PHYSIOLOGICAL/topics/language.htm>

<http://groups.yahoo.com/group/language-origins/links>

<http://groups.yahoo.com/group/language-origins/links>

http://www.absw.org.uk/Briefings/Social_behaviour.htm

<http://perso.club-internet.fr/tmason/WebPages/LangTeach/CounterChomsky.htm>

<http://students.washington.edu/dschruth/musicevolution.shtml>

<http://members.telocity.com/~hydra9/marcaat2.html>

Robin Allott: <http://members.aol.com/rmallott2/origin.htm>

Herder: <http://www.percepp.demon.co.uk/herder.htm>

<http://www.sciencedaily.com/releases/1999/02/990216135800.htm>

<http://www.massey.ac.nz/~ALock/webdck/origin.htm>

<http://www.hinduonnet.com/thehindu/2001/10/11/stories/08110007.htm>

<http://www.mus.cam.ac.uk/~ic108/lithoacoustics/>

http://www.infres.enst.fr/confs/evolang/actes/_actes65.html

<http://is.gseis.ucla.edu/impact/f01/Papers/Vasche/is209paper.htm>

<http://perso.club-internet.fr/tmason/WebPages/LangTeach/CounterChomsky.htm>

<http://lied.org/index.html>

<http://serendip.brynmawr.edu/biology/b103/f01/web1/wang.html>

http://homepage1.nifty.com/NewSphere/EP/b/lang_00.html

<http://www.ling.ed.ac.uk/evolang2002/ABSTRACTS/skoyles1.txt>

http://www.geocities.co.jp/Technopolis-Mars/8080/p/lang_01.html

<http://www.netgo.co.il/sites/ai/origin.html>

<http://www.liv.ac.uk/researchintelligence/issue14/language.html>

Andrew Lock. 1997. On the recent origin of symbolically-mediated language and its implications for psychological science.

S. Lea and M. Corballis (Eds) Evolution of the Hominid Mind. Oxford: Oxford University Press.

First draft: February: <http://www.massey.ac.nz/~alock/webdck/origin.htm>

[The Mozart Effect](#): music training improves verbal memory. Science 301, 914. 2003.

A mathematical model for distinguishing sweet sound from sour noise:

<http://news.sciencemag.org/sciencenow/2011/09/how-the-ear-distinguishes-sweet-.html?ref=em&elq=bef69c84e3fb443484fd761e85161d9f>

[Home](#)

Further reading

AT Tierney cs 2011 PNAS 108:15510-5. The motor origins of human and avian song structure

Human song exhibits great structural diversity, yet certain aspects of melodic shape (how pitch is patterned over time) are widespread:

- a predominance of arch-shaped & descending melodic contours in musical phrases,
- a tendency for phrase-final notes to be rel. long &
- a bias toward small pitch movements between adjacent notes in a melody (D.Huron 2006 "Sweet Anticipation: Music and the Psychology of Expectation" MIT).

What is the origin of these features?

We hypothesize that they stem from motor constraints on song production (ie, the energetic efficiency of their underlying motor actions) rather than being innately specified.

One prediction of this hypothesis is that any animals subject to similar motor constraints on song will exhibit similar melodic shapes, no matter how distantly related they are to humans.

Conversely, animals who do not share similar motor constraints on song will not exhibit convergent melodic shapes. Birds provide an ideal case for testing these predictions: Their peripheral mechanisms of song production have both notable similarities & differences from human vocal mechanisms (T.Riede & F.Goller 2010 Brain Lang 115:6980).

We use these similarities & differences to make specific predictions about shared & distinct features of human & avian song structure, and find that these predictions are confirmed by empirical analysis of diverse human & avian song samples.

014736