Language Learning and the Brain:
An Evolutionary Perspective[*]

William S-Y. Wang
Chinese University of Hong Kong

In this essay I will discuss some aspects of the phylogeny and the ontogeny of language within an evolutionary perspective. The first part of the essay stresses the heterogeneity of language, a view proposed by Weinreich et al. (1968). In contrast to Generative Grammar, which represents linguistic structures in terms of deep derivations, I submit that these structures can be better studied in terms of shallow memorial processes, cf., Bolinger (1961), and Tomasello (2003). This view of shallow memorial processes is consistent with an observation von Neumann (1958) made when he compared the computer with the brain, and with the recent proposals toward flat structure by Culicover & Jackendoff (2005, Chapter 4). The second part begins with a very brief review of recent developments in cognitive neuroscience, and especially of EEG methods. Several experiments are then discussed which reveal the remarkably rich biological resources for learning that the infant brings to the challenge of language acquisition. These include its ability to imitate facial expressions almost immediately at birth, and to distinguish the phonetic features of its native language from non-native languages, etc. These developments are demystifying language acquisition, and steadily laying a solid empirical foundation upon which our understanding of language ontogeny can build.

Key words: evolution, brain, cognition

1. Phylogeny and ontogeny

In a panoramic survey of the origins of life, two eminent biologists divided the billions of years of life on earth by eight major transitions (Maynard Smith & Szathmáry 1999). They started at the very beginning, with the appearance of isolated molecules that could replicate themselves — the first condition for life surely being the ability to reproduce. After several billion years of evolution, during which life forms on earth made several momentous changes, including the formation of DNA, sexual reproduction, etc.,

---

[*] This essay is offered in celebration of Alain Peyraube, a dear friend of several decades, in appreciation for his numerous important contributions to Chinese linguistics. I hope he enjoys the position taken here.

the final stage of life came with the emergence of human language. Indeed, more than anything else, language is the mental instrument that contributes to the unique achievements of our species.

Compared with the other seven biological transitions in evolutionary time, the emergence of language is really an extremely recent affair. According to the best estimates from anthropology and population genetics, modern humans left Africa to colonize the world around 100,000 years ago (Cavalli-Sforza et al. 1988, and Cavalli-Sforza & Feldman 2003). Many behaviors which we consider human-like that clearly require language, such as ritual burials, coordinated hunting, art and music, and sea crossings, came dozens of millennia later.

Yet 100,000 years is much too short a time span for biological evolution to endow us with anything like a "language organ" (Anderson & Lightfoot 2002), when we consider the millions of years it took Nature to fashion much less demanding body parts, such as the heart and the eye. The central issue for understanding language, therefore, can be cast as two related questions on its emergence. Phylogenetically, the question is how our species made the transition in the first place, from having no language to having language; that is, how did language emerge? Ontogenetically, the question is how does the child manage to learn something as complex as a language, so reliably, and seemingly so without effort?

Pinker (1994) advocated the idea of a "language instinct", which he used as the title of his highly popular book. But as Tomasello (1995) rightly points out in his review of the book, Pinker's "instinct" is quite distinct from the way the word is usually used. An instinct typically refers to a behavior that is highly stereotyped; yet it is evident that a major hallmark of language is its great diversity in time and in space. Also, an instinct typically refers to a behavior which does not require cultivation. Yet, it is clear that without a linguistic environment, no language can be acquired.

In the same review, Tomasello raises a deeper objection to the book. These are his words:

"The problem with the book is that Generative Grammar and its concomitant nativism are presented to the reader as established scientific facts, without even the hint of a hint that there are fierce theoretical and empirical debates currently raging over almost every issue discussed. That many linguists, indeed the majority of linguists, do not believe in a Chomsky-like Universal Grammar is not acknowledged anywhere in the 430 pages of the book."

Indeed, in its early days the movement of Generative Grammar rolled over the landscape of linguistics in juggernaut fashion. Some attribute this movement to the personality, some to the political times — in any case we must wait for a historian of our field to eventually sort out the actual socio-dynamic elements involved.[1] Nonetheless, the dissenting voices from highly respected sources can no longer be simply ignored, as they have been for several decades.

One of the purposes of this essay is to bring together some of the other voices from both within linguistics and from scholars in other fields who comment cogently on language. Language is not the deeply derivational, computational instrument that some believe it to be, but a rich heterogeneity of massively stored patterns. I will also try to review some of the recent research on the powerful resources for learning that infants bring to the acquisition of language. That these numerous patterns can be acquired so uniformly and so effortlessly is largely due to the availability of these powerful resources which each infant is endowed with even at birth.

Returning for now to the Generative Grammar movement, many criticize it for its sectarian stance, insulating linguistics from healthy interactions with other disciplines, including even the data-rich area of language pedagogy. For example, Philip Lieberman decried the "hermetic disjuncture" the movement created, and started his discussion with this assessment: "*The major 'contribution' of generative grammar to cognitive science is negative.*" (2005:289).

From another academic discipline, the well-known neuroscientist Ramachandran compared the 'nativism' of Generative Grammar with the divine intervention of Alfred Russell Wallace, which both shocked and saddened Charles Darwin:

"Alfred Russell Wallace said the mechanism is so complicated it couldn't have evolved through natural selection at all and must have resulted from divine intervention.… Chomsky said something quite similar, although he didn't

---

[1] Sven Öhman, for instance, wrote that "Chomsky is not a scientist at all, but a political ideologue..." (2007:5). Indeed, Chomsky's meteoric rise in public visibility in the 1960s was clearly due in large part to his political involvement with the Vietnam War.

invoke God…. He almost says it's a miracle. Unfortunately, neither Wallace's nor Chomsky's theory can be tested." V. S. Ramachandran (2004:75).

A cartoon by the humorist Sydney Harris makes the point well; it is reproduced below as Figure 1. In a spirit similar to Ramachandran's observation, my commentary of 1984 on Derek Bickerton's "bioprogram" was titled *organum ex machina* by analogy with the *deus ex machina* used in Greek dramas. Indeed, attributing the human ability to learn language to some implausible 'organ', 'instinct', or 'bioprogram' is not really facing the challenge directly to look for an explanation.

The most detailed analysis of Generative Grammar from the inside is that by Paul Postal, who devotes the better half of a volume to dissect both the reports and the behavior of the reporters, calling his exercise *Studies of Junk Linguistics* (2004:205-338). He recently recapitulated some of his more trenchant arguments in the latest issue of the journal *Biolinguistics* (2009), in an article entitled *The Incoherence of Chomsky's 'Biolinguistic' Ontology*.
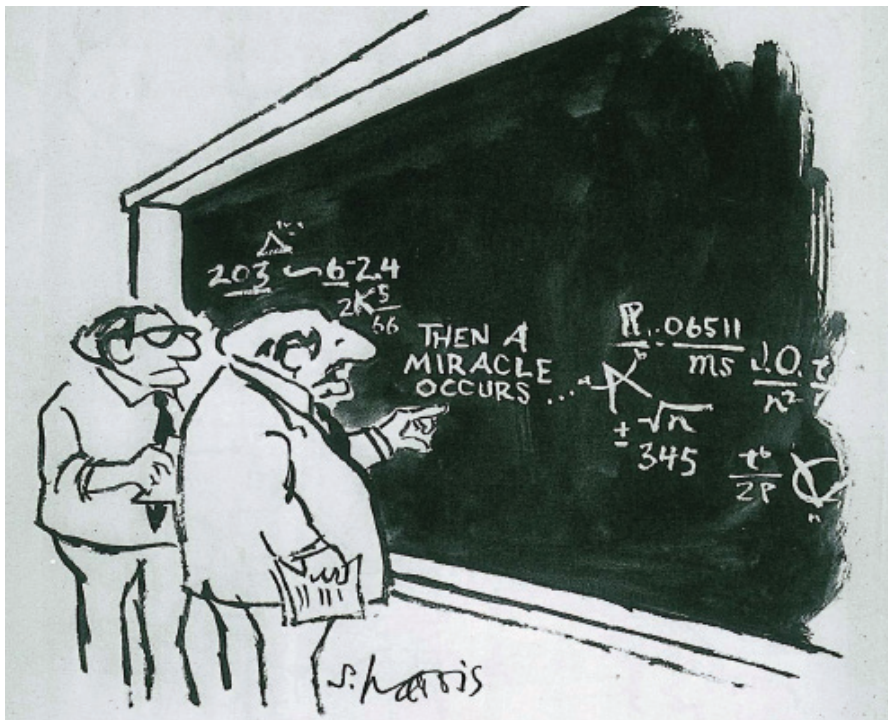
No doubt a major weakness of Generative Grammar can be traced to its early ethnocentric bias toward language structures of the European type, which carried over

to the Universal Grammar hypothesis. As late as 1980, given the tremendous wealth of linguistic diversity that has been reported from different parts of the world, it is incredible that anyone can still take a position of ethnocentric hubris expressed by the following words:[2] "*I have not hesitated to propose a general principle of linguistic structure on the basis of observation of a single language.*" Chomsky (1980:48).

The recent contribution by Evans & Levinson (2009), surveying a wide spectrum of far flung languages and demolishing one putative universal after another, should go a distance toward stemming such hubris. Although some of their conclusions may need to be revised with future work,[3] it rightly stresses an empirical aspect of language research that has been slighted too long. Furthermore, even if we had full records of all the 6,000 languages spoken in the world today, these authors remind us, they are still but a small fraction of the totality of languages that have ever been used in the millennia since language emerged. Considering the innumerable languages that have not survived to this day, the total diversity must be much greater than we could ever discover.

Moving past misleading metaphors like 'organ', 'bioprogram', and 'instinct', linguists are now much more open to interaction with other disciplines, realizing what an enormously complex challenge language is. Ray Jackendoff recently criticized the excessively narrow perspective of Generative Grammar, especially for its counter-productive isolation from other disciplines, and for its excessive emphasis on syntax. He spoke for many when he concluded his huge synthetic volume with these final words:

> "But linguistics alone cannot sustain the weight of the inquiry. We need all the help we can get from every possible quarter. Above all, my hope for the present work is that it can help to encourage the necessary culture of collaboration." Jackendoff (2002:429)

In fact, in a later work reviewing various versions of Generative Grammar over several decades, Culicover & Jackendoff (2005:88) rightly point to several principal failings of this movement:

---

[2] Quoted in Evans, Nicholas & Stephen Levinson (2009).

[3] The study of linguistic diversity has typology as one of its main tools at present. However, Ernst Mayr urged the replacement of typological thinking with population thinking: "Typological (essentialistic) thinking is misleading when applied to organisms. What must be used instead is population thinking which realizes that in a biological population every individual is unique and differs from all others. The statistical mean value of a population is merely an abstraction. Dual causality as well as the uniqueness of every individual of a biopopulation characterize the world of living beings and are therefore characteristic for biology" (excerpted from his Walter Arndt Lecture). These remarks concerning research methodology certainly merit our reflection.

"First, it does not take the task of the learner to be central; rather, it develops out of a conception of human language as a 'perfect' system. Second, it assumes without argument that the correct characterization of linguistic knowledge takes the form of a derivation. Third, it imposes certain notions of economy on its formulations that appear to have little if any empirical motivation. And fourth, it relies heavily on the Uniformity methodology of earlier work."

Indeed, many linguists have been handicapped too long by the artificial and counter-productive stance of thinking of their discipline as "autonomous," taking refuge in such simple dichotomies as Saussure's *langue:parole*, or the *competence:performance* of Generative Grammar. How can language learning, whether by child or by adult, be anything but "central" to a discipline dedicated to language?

The child gathers its language samples from a diversity of sources — from its primary caregiver as well as from other members of the household, and later on from playmates as well as from other adults in school or in the neighborhood. As Gauchat (1905) demonstrated over a century ago, these language samples display a high degree of variation, even within a single family in an isolated Alpine village. We have longitudinal data here as well, thanks to the follow-up study of Hermann (1929). These works and many other similar studies have been put in a modern perspective on linguistic variation by Labov (1994). In order to construct language from these samples, the child does its best to impose some sort of order on this heterogeneity, recalling these terms from the insightful paper by Weinreich, Labov & Herzog (1968). But the end product of this process of construction is very far from "perfect".

The variation can be readily seen not only in the spoken and written language samples that people produce, but also in their judgments on grammaticality and acceptability. J. R. Ross reports on some experiments carried out with professional linguists, who were asked to make such judgments. The extensive variation he found among these subjects led him to remark:

"The view of language that seems most plausible to me is that the sentences of a language are points in an n-space… An idiolect is a vector in this n-space… And each speaker's vector, or path, through the space will, I expect, be as individual as his or her face — a linguistic fingerprint." (Ross 1979:160)

In a related vein, Edward Sapir once said that "all grammars leak." Natural languages by their very nature are basically different from the clean homogeneity of formal systems, and Procrustean efforts to model them as such have steered us in

counter-productive directions. In an insightful discussion of evolution, the geneticist François Jacob (1977) wrote:

> "Living organisms are historical structures, literally creations of history. They represent not a perfect product of engineering, but a patchwork of odd sets pieced together when and where opportunities arose."

Languages too are clearly historical structures, both phylogenetically and ontogenetically (Schoenemann & Wang 1996). Their evolution is opportunistic in much the same way, at both the level of our species from generation to generation, and at the level of the child as it pieces together the odd samples of language in the ambient environment to construct a language of its own. What Jacob calls 'patchwork' above I have elsewhere called 'mosaic', with the same scenario in mind. The situation is similar when an adult adds new words or constructions to his language, driven by fortuitous social needs.

Regarding the second point on derivational complexity made by Culicover & Jackendoff above, it is instructive to examine a sentence that has been discussed for many years in Generative Grammar, reproduced here from Culicover & Jackendoff (2005:99). The sentence is actually a very simple one: *Floyd broke the glass*. However, as can be seen in Figure 2, it is derived by a set of eight sentences marking the tense, aspect, transitivity, inchoativeness, etc., each successively adding to the depth of the derivation.

As a result, a very simple sentence is represented by a monstrously complex structure, reminiscent of Rube Goldberg machines, created for satire rather than function. It is hard to imagine that such a heavy syntactic structure ever comes into play in the mind, either when the child learns the sentence or when anyone uses it in a real communicative situation. The derivational theory which produces such structures has clearly drifted too far away from our goal of understanding language in a realistic way.
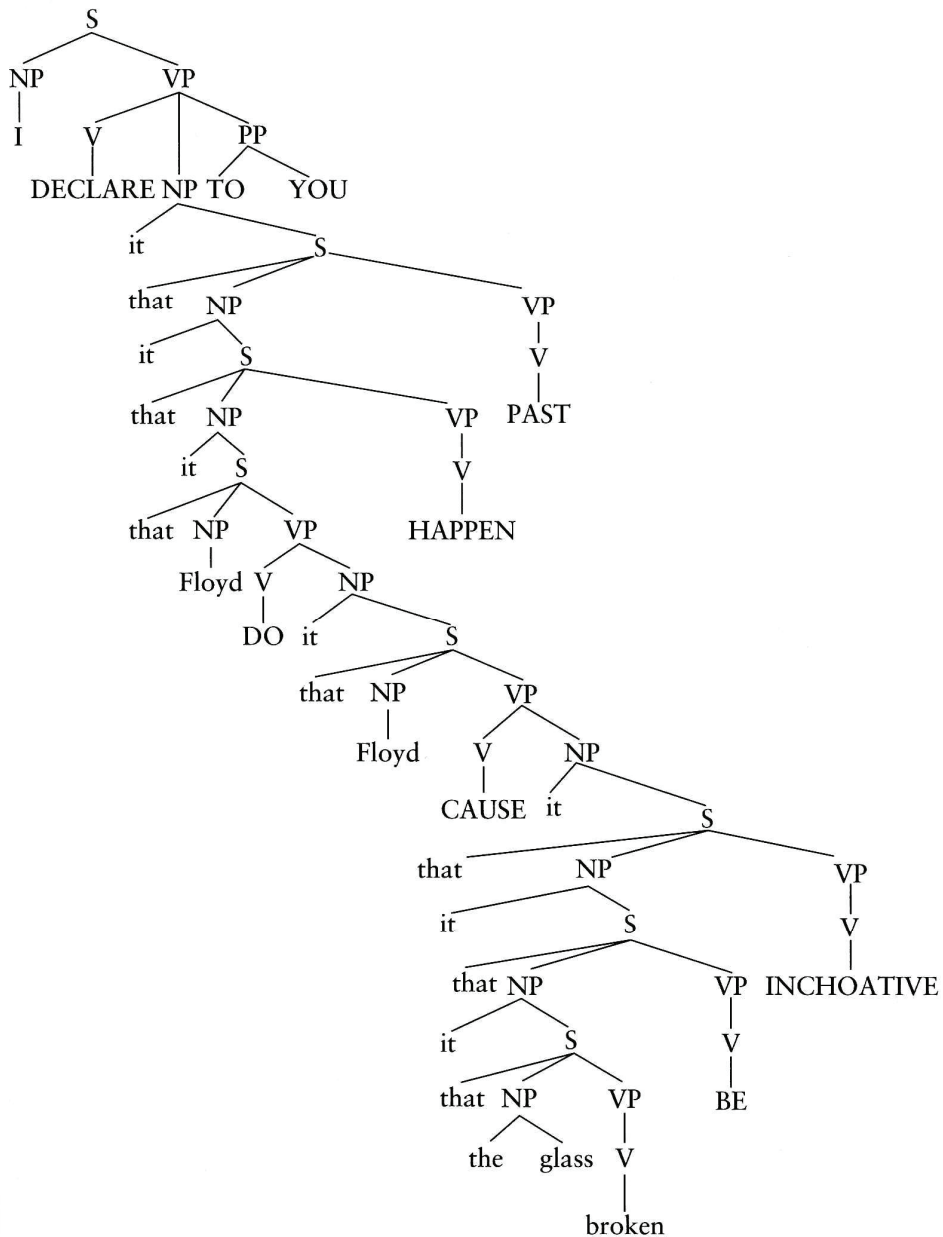
In addition to the various arguments advanced by Culicover & Jackendoff for flat structures in syntax, a general point to be made here has to do with an important distinction between how brains and computers process information. A comparison of

the nerve cell with the electronic transistor was first made by John von Neumann (1958), whose pioneering design for computer architecture remains centrally relevant today.[4] His book has been recently reprinted (2000), with a substantive foreword by Paul and Patricia Churchland.

In terms of current knowledge in these fields, transistors are both much faster and much more accurate than neurons. This fact was dramatically revealed by the mass media in 1997, when IBM's Deep Blue Supercomputer won the world chess championship beating Grandmaster Garry Kasparov. On the other hand, the immense power of the brain resides in its massive number of synapses, several orders of magnitude greater than the number of transistors in any supercomputer. It is the parallel and distributed activities of these innumerable synapses that enable the brain to outperform the computer in many tasks. As the Churchlands put it:

> "Conjointly, these two severe limitations — one on speed, and the other on accuracy — drive von Neumann to the conclusion that whatever computational regime the brain is using, it must be one that somehow involves a minimum of what he calls 'logical depth'. That is, whatever the brain is doing, it cannot be sequentially performing thousands upon thousands of sequentially computational steps... Given the slowness of its neuronal activities, there isn't enough time for the brain to complete any but the most trivial of computations. And given the low accuracy of its typical representations, it would be computationally incompetent even if it did have enough time."

Taking the above into consideration, it would seem that the assumptions behind the syntactic analysis in Figure 2, leading to excessive derivational depth, are not justified. Just as natural languages are basically different from formal systems, the parallel processing human brain works very differently from the serial processing computer. In fact, the derivational approach of Generative Grammar, challenged by Culicover & Jackendoff, was questioned much earlier by Dwight Bolinger (1961), who asked:

> "Is grammar something where speakers 'produce' (i.e., originate) constructions, or where they 'reach for' them, from a pre-established inventory, when the occasion presents itself? If the latter, then… constructions are not produced one from another or from a stock of abstract components, but filed side by side, and their interrelationships are not derivative but mnemonic."

---

[4] A more recent comparison of real and artificial neural networks has been made by Crick (1989), who discusses some of the more important differences between them.

Looking at Bolinger's question[5] from the viewpoint of language acquisition, Tomasello (2003) similarly opts for a memorial approach, as stated in these words:

> "This means that in many cases children's comprehension and production of relatively complex utterances are based on a simple retrieval of stored expressions, whereas in other cases they are based on the cutting and pasting together of stored linguistic schemas and constructions of various kinds and degrees of abstraction."

2. Evolutionary perspective

To move forward on our inquiry, we must put language back into an evolutionary framework, connecting it securely to its biological and social roots. The geneticist T. Dobzhansky (1973) once famously said that "*Nothing in biology makes sense except in the light of evolution.*" The basic concepts of evolution, variation and selection, apply no less to linguistics, except for the fact that whereas organisms typically evolve along one tract, language evolves along two tracts — the biological and the cultural.

The major course of how an organism will develop is determined at the instant when the sperm and the egg fuse to produce a new cell, which contains the entire genetic blueprint necessary for the organism to grow biologically. In contrast, there is no such blueprint for language. The child must depend on its biological resources, constantly sample the surrounding world of sights and sounds, people and objects, comforts and frustrations, and construct a language on its own. Since each child has its own unique history of development, the language it constructs will be no less than a *linguistic fingerprint*, in the sense of Ross quoted earlier.

If this ability to construct a language seems miraculous and mysterious, it is because such biological resources have been evolving for millions of years, the power of which is something we are only now beginning to appreciate. This perspective is described vividly in Jacob's famous metaphor that evolution works like a *tinkerer* (1977), giving old parts new uses.[6] The mechanism is so central to evolution that Gould & Vrba (1982) proposed a new word to highlight it, which is '*exaptation*'. Recently, Ramachandran (2004:76) put language squarely within such an exaptative perspective:

---

[5] Bolinger's question has been recently addressed in neuroscience with respect to arithmetic, cf. Grabner et al. (2009). Perhaps similar methods can be used for language.

[6] Some readers will be familiar with the Tinkertoy, very popular with children before electronic toys. It is discussed by Dewdney (1993).

"I suggest… that it is the fortuitous synergistic combination of a number of mechanisms which evolved for other purposes initially that later became assimilated into the mechanism that we call language."

The tinkerer metaphor is descriptive and vivid, though I came to know Jacob's 1977 paper only many years later.[7] In reaching for a metaphor toward explicating language evolution in a series of lectures I gave in India in 1978, I opted for the word "mosaic", as in the paragraph below, and also in Wang (2007).

"Many of these abilities are present to different degrees in other animals (witness the instances of tool making and problem solving in chimpanzees). Most of them probably emerged much earlier than language in hominid evolution. Gradually and piece by piece, these abilities were increasingly made accessible for use in the elaboration of language, much as adding pieces to a **mosaic**. In parallel fashion, these abilities have also been made accessible to several other elaborate human institutions, most notably mathematics and music."

Independently of my usage, Jim Hurford also settled on the mosaic metaphor in a recent paper (2003). The theme that is recurrent across these metaphors, *mosaic*, *tinkerer*, *patchwork*, *cut-and-paste*, is that language emerged fortuitously as a complex adaptive system, not as a new instinct or a new organ (Tzeng & Wang 1983b). Language grew gradually as various domain general abilities were tinkered together.[8] Clearly, many regions of the brain must participate to make language possible. But in the words of Dick et al. (2001:760),

"... it is no more appropriate to refer to a participating region as a language zone or grammar zone than it would be to refer to the elbow as a 'tennis organ'."

It is only natural that we share these domain general abilities to varying extents with other species, especially with the chimpanzee, our closest relative, separated from us some six million years ago. Recent research in primate behavior has revealed a whole series of interesting discoveries here, shedding much light on our homological development. However, as Bolhuis & Wynne (2009) remind us, a lot can be learned as

---

[7] I thank Tom Schoenemann for calling my attention to Jacob's paper, as well as to many other important works in evolution and in neuroscience over the years.

[8] Over the last two decades there have been numerous computer simulations of language emergence through the random interactions of populations of artificial agents. Gong (2009) provides an overview of these efforts as well as a model of his own.

well from more distant species; birds in particular have evolved remarkable vocal skills by convergent evolution.

Chimpanzees have a complex social structure that permits them to hunt in cooperative groups, mothers teaching skills to the young, such as how to crack nuts, as well as the transmission of various cultural traditions which are specific to regional populations. Comparative studies of the two species of chimpanzees, *Pan paniscus* and *Pan troglodytes* reveal very interesting differences between them, both in group structure and in social behaviors, which are radically different from each other. F. De Waal has produced a series of useful reports of our knowledge in this area (1998, 2005a, and 2005b). Evidence is also accumulating that chimpanzees can plan ahead, such as gathering stones for future throwing, cf. Osvath (2009).

Here is a succinct description by the biologist Christophe Boesch (1991) of an instance of a chimpanzee mother teaching her son the proper way to crack a nut. The observation is all the more valuable since it took place in the wild rather than in a zoo or primate center, where the animal's behavior may be influenced by humans.

> "... After successfully opening a nut, Sartre replaced it haphazardly on the anvil in order to gain access to the second kernel. But before he could strike it, Salome took the piece of nut in her hand, cleaned the anvil, and replaced the piece carefully in the correct position. Then, with Salome observing him, he successfully opened it and ate the second kernel."

Given our close genetic relations with the chimpanzee, such studies provide an important though approximate baseline toward understanding the biological resources the child can have at his disposal toward language acquisition. The recent experiments by Herrmann et al. (2007) are especially instructive in comparing the human child with the chimpanzee and the orangutan, exploring what they call the Cultural Intelligence Hypothesis. They find that whereas all three apes have similar cognitive skills dealing with the physical domain, the human child, at age two and a half, performs much better with those skills dealing with the social domain. We are much better endowed genetically for social cognition than the other apes.

It is important to remember, of course, that six million years of independent development separating us from the chimpanzee can bring about major changes, even given the slow pace at which biological evolution operates. As noted earlier, two million years of separation between the two species of chimpanzees has already produced radical differences in their societies. Our understanding here has increased significantly now that the genomes of both the human and the chimpanzee have been completely sequenced, and powerful algorithms have become available to compare these masses of data to hunt for significant differences between these genomes.

Hidden among these billions of nucleotides are short stretches of human DNA which have mutated at much faster rates, which presumably have been favored by natural selection. One such stretch discovered early in this research, consisting of 118 nucleotides, has been called HAR-1 — human accelerated region 1. This stretch appears to underlie some general abilities which have special implications for the emergence of motor movements which we have exapted to produce speech. An accessible discussion of these exciting new developments is Pollard (2009). Rapid progress is being made in identifying more HARs which greatly enhance our understanding of the uniqueness of our species.

However, in interpreting these findings, it is important to keep in mind that just as there is no 'language organ' *per se*, it makes even less sense to speak of a 'language gene.' Rather, a balanced statement of where we stand at present in our understanding of the problem is the following by Dick et al. (2004:226):

"The most parsimonious account of language evolution is one where incremental, quantitative changes in primates' vocal tract, fiber pathways, and neural anatomy converge with social and cultural developments. From this convergence arises the framework upon which the complex language skills could build."

3. Language in the brain

Although language disorders have been observed since antiquity, cumulative interest in the relation between language and the brain began in the 18th century, when various researchers observed the special role the left hemisphere of the brain seems to play. Injured people who are paralyzed on the left side of the body often keep their language intact, while paralysis on the right side of the body typically is accompanied by language impairment. Since it was known that the right side of the body is controlled by the left hemisphere, the inference was made that an injury there could be responsible for the language impairment as well.

Systematic study of language and the injured brain began in the middle of the 19th century, with the French scholar, Paul Broca (1824-1880), who has given his name to a type of language disorder, Broca's Aphasia. Patients with this disorder typically exhibit difficulties in producing fluent speech. Broca identified a region in the frontal lobe in the left hemisphere which is correlated with this syndrome. Broca's discovery was soon followed by a German scholar, Carl Wernicke (1848-1904), who identified a language region in the temporal lobe in the left hemisphere. Patients with Wernicke's Aphasia can speak fluently with normal rhythm and intonation. However, the speech they

produce is often semantically meaningless, with non-words mixed in with words, resulting in what has been called 'verbal salad.'

Although these two discoveries were of fundamental importance, the science that they needed for a foundation was not yet available. The realization by the great Spanish anatomist, Santiago R. Cajal (1852-1934), that the human brain was made up of many billions of unconnected neurons was to come only at the very end of the 19th century. Nonetheless, their influence has been pervasive, especially in clinical work in helping aphasics.

The Harvard neurologist, Norman Geschwind (1926-1984), enhanced their influence by high-lighting a fiber tract connecting these two areas in the left hemisphere, called the Arcuate Fasciculus because of its bend around the lateral sulcus that separates the frontal lobe from the temporal lobe. Although nowadays neuroscientists are much more aware of many other parts of the brain involved in the use of language, cortical as well as subcortical, there is little disagreement that these three regions constitute a 'core' area.

The last quarter of the 20th century saw an explosive growth of interest in the brain, together with revolutionary advances in the technology of studying it non-invasively in the normal brain. No doubt this progress was much facilitated by George Bush, in a Presidential Proclamation (#6158) on July 17, 1990, which named the period 1990-2000, Decade of the Brain to stimulate research funding. We can now measure the movement of blood and changes in electromagnetism in the head while the subject is doing various tasks, as well as trace the fiber pathways from region to region by a method called Diffusion Tensor Imaging. A recent investigation by Marco Catani et al. (2005) using the latter method shows the nerve fibers of this core area of language in vivid colors (see their Figure 3), thus giving the Broca-Wernicke-Geschwind model a solid empirical interpretation.

Amid this wealth of technology for imaging the brain's activities, including MRI, PET, MEG, and EEG, some are better at spatial resolution while others are better at temporal resolution. There is now a sizeable literature on the application of each of these methods in language research.

Let us first consider an early MRI study on bilingualism by Kim and his group in New York (1997). MRI experiments can yield very precise spatial information on which area of the brain is active during a specified event. Theirs is a functional MRI study on 12 bilingual subjects, involving a spectrum of languages. Six of the subjects grew up bilingual, and six subjects learned their foreign language as young adults. While in the MRI machine, each subject was asked to describe some events silently by internal speech, in one of his two languages.

Two interesting and closely related results emerged from their experiment that can be roughly summarized as follows. One is that approximately the same posterior region

14

of the left hemisphere is active regardless of the language involved, and regardless of the age of language acquisition. This region is assumed to correspond to Wernicke's area. The other result has to do with an anterior region of the left hemisphere, assumed to correspond to Broca's area. It turns out the age of acquisition makes a critical difference here. The subjects who are bilinguals from childhood also activate approximately the same region, regardless of the language in which they were describing the events while in the machine. On the other hand, the bilinguals who learned their other language as young adults activated separate and distinct portions of the anterior region for the two languages.

These results may be interpreted as saying that the semantic structures have a lot in common across various languages, and they draw upon the same neural resource. On the other hand, the phonetic forms of the words across languages involve motor skills the acquisition of which is much more age sensitive. This reminds us of authors who can write eloquently no matter how late they have mastered a foreign language. But they give away their native language the minute they start to speak.

Perhaps the best known illustration of this dissociation is the famous Polish author Joseph Conrad, who did not speak English until his twenties. He wrote elegant and perceptive English, as in his novel *Heart of Darkness,* but always spoke with a strong Polish accent. In my personal experience, the most vivid example of this dissociation of semantics from phonetics was the great linguist Roman Jakobson, who was once introduced at the University of Michigan as being able to speak twelve languages fluently, "all in Russian."[9]

A more recent MRI study I will review here, by Siok et al. (2009), concerns the famous Sapir-Whorf hypothesis, which speculates that our behavior is influenced by the language we speak, even when language is not overtly involved, cf. Kay & Kempton (1984). In a series of studies by Paul Kay and his associates over the last several years, this hypothesis has been examined from the viewpoint of color perception. These studies make use of the fact that our visual field can be divided into the left visual field, which projects more strongly to the right hemisphere of the brain (RH), and the right visual field, which projects more strongly to the left hemisphere (LH). By appropriate experimental conditions, such as centering a fixation point to minimize eye movement and presenting visual stimuli very briefly on the screen, we can control which hemisphere will receive the stimuli first.

The subjects are presented with 12 color patches, arranged in a ring around the fixation point. Eleven of these patches are identical in color; they are called 'distracters';

---

[9] The story of Jakobson's linguistic prowess and of his strong Russian accent is recorded for posterity in Lamb (1999:41). Lamb's volume also provides a good discussion of the cognitive neuroscience of language from the viewpoint of a linguist.

one patch has a different color and may appear in any position in the ring, it is called the 'target.' The task is for the subject to focus on the fixation point and report as quickly as possible whether the target appeared on the left side of the ring, or on the right side of the ring. A key element in this experiment lies in how the target color differs from the distracter color — whether the difference crosses a linguistic boundary. Physically, the two colors in the experiment are always the same distance apart in terms of their wave lengths. However, they may both fall within the same linguistic category, for instance 'green'. Or, they may fall in different linguistic categories, for instance 'green' for one and 'blue' for the other.[10] The purpose is to examine how linguistic categories, not overly involved in the task, influence the perceptual judgment, and to find the brain regions which reflect this influence.

Brain activities were monitored by event-related functional magnetic resonance imaging (fMRI) during the experiment. Consistent with earlier studies on color perception, it is found that LH activities are stronger and initiate earlier when the subject responds to color differences if they fall across linguistic categories, i.e., when different color names are involved. More interestingly, activities in the V2/3 area of the brain, which is known for color perception, coincide in time with those of the temporoparietal language region in the LH.

This synchronization between two separate brain regions provides the empirical basis for confirming the Sapir-Whorf hypothesis in color perception, where linguistic categories play a facilitating role even when language is not overtly involved while subjects perform their assigned task. We may hope that in the years to come, more and more distinct brain regions will be identified in their operation to support various linguistic processes. These are the distinct pieces that evolution has tinkered together to make the language mosaic.

4. Language and brain waves

I will now turn to another technology, the EEG, or **E**lectro-**E**ncephalo-**G**raphy. Specifically I will focus on the recent use of the EEG to probe into various distinctive aspects of language acquisition in the case of very young children.

One important advantage of EEG is its fine temporal resolution — on the order of just milliseconds following the brain's activity. The brain is monitored simply by an

---

[10] It is well-known that languages differ in how they divide up the color spectrum. For example, in contrast to English, Japanese does not use separate lexical labels for 'green' and 'blue', whereas Russian has two separate lexical labels for different shades of 'blue'. Winawer et al. (2007) report on color perception in the Russian case.

ensemble of electrodes attached to the scalp; hence another advantage is that the method is completely non-invasive and free from problems of radiation associated with some other methods. The subject can move about relatively freely during an experiment, and can carry the EEG machinery constantly with him.

Such is the case, for instance, in the growing field of research on Brain Computer Interface - BCI, where the paraplegic patient controls the movement of his wheelchair by his brain waves. The method is also especially useful for young children who will not stay still for any length of time. Also, the subject is not required to perform any prescribed judgments or motor responses, which are often not possible for prelinguistic children or various types of apes.

Brain waves were first noted by a German psychiatrist, Hans Berger (1873-1941), some two decades after Cajal discovered the neuron. His 1929 paper showed the alpha wave, with a frequency within the range of 8 to 12 Hz; many more brain waves have been discovered since, covering a wide range of frequencies. We now know that these oscillations in electric potential picked up at the scalp are due to the electro-chemical signals the neurons send to each other when they communicate across synapses. In particular, when certain groups of neurons are aligned in parallel and fire synchronously, they will produce a more stable EEG measurement.

There are trillions of synapses in the human brain, with a numerous quantity active at any given time, reporting different sensory stimuli, processing a multitude of diverse information, giving commands to various muscles, etc. The waves we pick up with the electrodes attached to the scalp are a jumbled up summation of all these neural activities, filtered through the intervening membranes covering the brain and the bony skull which encase them. It is a monumental challenge to decode these jumbled waves, and interpret them according to the particular linguistic task the subject is performing at that time. With the help of powerful computational equipment and sophisticated statistical methods, we are beginning to make some headway in face of this challenge.

A standard method of research in EEG is to mark the time of an event with respect to the brain waves which are associated with it. As this event is repeated a large number of times, the waves of electric potential associated with the event will summate, while other incidental unrelated waves will tend to cancel each other out. The result is a wave form with distinctive properties of polarity, latency, and shape, which are presumed to be specifically related to the event. Such waves are called ERPs, or Event Related Potentials. Thus the P300 is an ERP that represents a positive potential that peaks around 300 ms after the event, and the N400 is an ERP which is negative and peaks around 400 ms after the event.[11] These two are among the handful of ERPs which have

---

[11] The abbreviation 'ms' stands for 'millisecond(s)'.

attracted a great deal of attention in the past decades. A recent analysis of the P300 is Polich (2007), and of the N400 is Lau et al. (2008). An overview of EEG research in language is Kutas et al. (2007).

I was first exposed to ERP research on language early in the 1980s, through collaboration with the lab members of Emanuel Donchin, a pioneer in electro-physiological studies, who was then at the University of Illinois. See his historical note on EEG in Donchin (2006). We were inspired by exceptional irregularity of spelling in English, caused largely by the large influx of French words due to the Norman Conquest. Our stimuli consisted of four lists of word pairs. The Ro words rhyme, and are written alike, such as *plea* and *flea.* The R words also rhyme, but their spelling is quite different, such as *make* and *ache.* The Wo words are written as though they should rhyme but they do not, such as *said* and *paid*; as in this example, the Wo words differ only in the initial consonant. Lastly, there is the control list of words which are not related either in rhyme or in written form, such as *mind* and *wall*. See Polich et al. (1983).

These lists of word pairs were presented visually under two conditions. In one condition the pair of words were separated by a very brief ISI (**I**nter**S**timulus **I**nterval) of 50 ms. In the other condition, they were separated by an ISI of 600 ms. The subjects performed two tasks. Both the reaction times of their responses and ERP data were taken during the experiment. Since the reaction time requires a motor response, it involves a longer latency time. The reaction times ranges from 600 to 1,000 ms, while the ERP latency ranges narrowly from 530 to 620 ms.

In the rhyme task, respondents judge whether words rhyme or not. To make this judgment, the subject must retrieve the presented visual words from his mental lexicon, access their phonetic forms, and compare them. The Ro words encourage a fast 'yes' response since the visual forms already suggest that the words rhyme. For R words, like *make* and *ache*, the visual form is actually interference, urging a 'no' response when they do rhyme, a judgment that is called a 'false negative'. For Wo words, like *said* and *paid* the interference is in the other direction, urging a wrong 'yes' response, which is a 'false positive'. In fact, these are the words that produced the greatest latency in the rhyme task.

In the other task, the visual task, the subjects judge whether or not words have similar spellings. Like in the rhyme task, the fastest latency is with the Ro words; even though no reference to the phonetics was required. The fact that the words rhymed facilitated the judgment of their visual similarity, and the judgment comes earlier than the Wo words. Since the words in both these lists differ only in their first consonant, speech recoding apparently kicked in automatically to speed up the judgment for Ro words.

Another interesting observation can be made on these data with respect to the two settings of the ISI, i.e., 50 ms and 600 ms. This has to do with the consistent greater

latency for the shorter ISI across both tasks and all four lists, on average about 25 ms. The subject can only make a comparative judgment, of course, after the second word is presented. In the case of the 600 ms ISI, we can assume that the first word has been fully retrieved and analyzed from the mental lexicon, and the comparison can be accomplished as soon as the second word is retrieved and analyzed. On the other hand, an ISI of 50 ms may be too brief for the full analysis of the first word, and this process continues into later stages, thus delaying the judgment. Whether this is a correct interpretation must await future experiments.

The brain must process languages written with alphabets differently from those written with other orthographic principles. In written language as well as in spoken language, we must resist the easy temptations of ethnocentrism. In particular, investigations on how the brain processes Chinese writing,[12] with its complex shapes and dual construction of semantic and phonetic components, can tell us a great deal about language in the brain. An overview of research in this area is available in the comprehensive anthology compiled by Ping Li et al. (2006), where a third of the volume is devoted to language and the brain. Rapid progress is being made in this area through the ERP experiments of Chiaying Lee and her group in Taiwan (Lee et al. 2006, 2007, and Lee 2008), and the fMRI experiments of Lihai Tan and his group in Hong Kong (Siok et al., 2004, 2008, 2009, and Tan et al. 2005, 2008). Hopefully, before long we will have a more complete understanding of how the brain supports reading and writing in our species; (cf. Tzeng & Wang 1983a).

Another exciting frontier in applying ERP methods is to provide information on children in assessing their neural development for language by comparing their brain waves with those generated by adults. The working hypothesis is that when the children's brain waves become the same as those of the adult, they will have developed the biological resources for that linguistic task. Valerie Shafer (2007) and her group have been investigating this issue, and provide some rough landmarks on this issue.

> "Synaptogenesis is creation of synaptic sites on which axons can connect. Subcortical and primary cortical areas are relatively mature in newborns and reach a peak in synaptogenesis by 3 months of age. Association cortex, such as prefrontal regions, does not reach peak synaptogenesis until around 3 years of age. These peaks are followed by loss of synaptic sites and loss of neurons related to absent or weak connections...." (Shafer et al. 2007:32)

---

[12] A historical account of literacy in Chinese is Wang et al. (2009).

The general principle here is approximately ontogeny recapitulating phylogeny. We share many subcortical parts of the nervous system with other species, and construction of these parts is relatively complete at birth. It is our cortex that has evolved much more elaborately than that in other species; this takes place much later phylogenetically, as well as ontogenetically.

The primary cortical areas which are responsible for sensori-motor systems are constructed comparatively early. The association cortex, which is active in linking and integrating sensory modalities, is completed much later. This is another region of the cortex that Norman Geschwind focused on, especially the Angular Gyrus, situated close to the junction of three hemispheric lobes — temporal, parietal and occipital. It is of special importance for language, since language is the associator *par excellence* of the various modalities. All languages appear to have words that are shared across modalities, e.g., soft light, soft sound; sharp taste, sharp edge. Primates without language have much greater difficulty making cross-modal associations.

An early discovery in this area is the ability of newborn infants to imitate adults' facial expressions, the youngest surveyed being less than one hour old (Meltzoff & Moore 1977). The vision of infants that young is still quite underdeveloped, and of course they have yet to see themselves in a mirror to understand their body parts. Yet they are able to imitate a gesture made by an intruding face. This issue of imitation has been recently discussed (Meltzoff & Decety 2003) in the context of mirror neurons (Rizzolatti & Craighero 2004, and Arbib 2006) and the theory of mind that every child will develop. On the whole, this is a very exciting and active area of research, and in the years ahead we may expect significant advances in our understanding here.

Closer to speech, the very early ability to structure perception to serve language on the basis of communicative sounds is certainly an important part of the biological resource that a child brings to his task. Similar abilities have been reported for other species as well for species-specific vocalizations, though in humans they are considerably more refined. Kuhl (2004) provides a valuable overview of our knowledge here, including a developmental time-table of how perception relates to production, as well as the child's use of statistical information in locating word boundaries.

Several ERP studies have been reported on the question of when infants begin to detect phonetic contrasts which are found in their native language, as opposed to those found in a foreign language. For one study in Seattle, Rivera-Gaxiola et al. (2005) prepared three sets of syllables which differed in their Voice Onset Times (VOT). Spanish speaking infants have a pre-voiced /d/ with a VOT of -24 ms in contrast with a /t/ with a VOT of +12 ms. On the other hand, a stop consonant with a +12 ms VOT would be heard as a /d/ by English speaking infants. For this latter group, a /t/ is typically strongly aspirated, with a VOT of +46 ms.

The ERP finding of these researchers is that both the English-speaking infants and the Spanish-speaking infants can readily distinguish native contrasts from non-native contrasts at as early as seven-months old. Furthermore, in the same experiment they found an intriguing difference in the 11-month old infants they tested, which were divided into two groups. While both groups maintained the ability to distinguish native and non-native contrasts, the polarity of the ERP in one group was positive, while it was negative in the other group. Such a finding would be obscured if an average ERP were taken across all 11-month olds, as Kuhl (2004) also observes. Given this finding however, it would be of great interest to follow them longitudinally to see how the difference in polarity affected their linguistic abilities.

The Seattle study was followed by an ERP study on a similar issue, reported by a collaborative team consisting of Angela Friederici and her colleagues based in Leipzig and in Paris (2007). The stimuli they constructed were the disyllables which they transcribed as /ba:ba/ and /baba:/, where the colon indicates a long vowel. Their subjects were four-month old infants, still younger than the Seattle study. The German-speaking infants were assumed to feel native with the German pattern of word initial stress, suggested by the long vowel in the first syllable. The same went for the French infants with the French pattern of word final stress, by the long vowel in the second syllable of the stimuli. On the basis of the ERP patterns these stimuli elicited in the two groups of infants, they suggest that neural representations of word forms are already in place in the infant brain at as early as four-months of age.

Working with infants still younger, Chao He et al. (2007) studied the EEG of three groups. They monitored two types of mismatch responses in two-month-olds, three-month-olds, and four-month-olds. One type of response, a left-lateralized positive wave (MMP) is strong in two-month olds, weaker in three-month olds, and insignificant in four-month olds. In contrast, the other type of response, a faster, adult-like negativity (MMN), lateralized to the right hemisphere begins to appear in the two-month olds and becomes earlier and stronger as age increases. The dissociation between these two ERP components is a very interesting finding on the maturational time-tables in infant perception in this early age range. The stimuli these investigators used are changing pitches in piano tones. Hopefully, future studies will reveal brain responses to other age ranges, as well as to other forms of auditory stimuli more closely related to speech sounds.

5. Concluding remarks

As the studies reviewed above show, we are just at the beginning of laying an empirical foundation toward understanding how language learning takes place. Building

upon the legacy of pioneers like Broca, Wernicke and Geschwind, we now have a much more quantitative knowledge of brain functions which make language possible. The experiments are getting increasingly focused on the biological resources which make language possible.

The coverage in this essay is neither balanced nor comprehensive, leaving out many topics that are directly relevant to our concern. Among these are the area of children with Selective Language Impairment (SLI) compared with those of Typical Language Development (TLD), discussed by Shafer et al. (2005), and the area of the language of neurologically intact adults compared with aphasic patients, discussed by Dick et al. (2001). Nonetheless, it is easy to see the impressive progress that has been made over the last decade or so. In large part, this new empirical foundation is due to the convergence of expertise and interest from many disciplines, ranging from developmental psychology, to genetics, to animal behavior, as well as the explosive advances in the cognitive neurosciences.

Paradoxically, linguistics has not been at center stage in the progress that has been made, as one might expect of a discipline whose central focus is on language. This is in part due to the unfortunate "hermetic disjuncture", in Philip Lieberman's words, that Generative Grammar has created for several decades between linguistics and other disciplines. A great deal of time and energy has been spent by many linguists in deriving very complex syntactic trees, with little regard for the relevance these diagrams may have for the language user, despite the early doubts raised by eminent scholars, like the late Dwight Bolinger (1907-1992).

We can only hope that in the decades to come, more and more linguists will take the evolutionary perspective to heart, and commit more positively to the inter-disciplinary research on language. Language is a gift that natural selection has endowed our entire species with, and it is the business of linguists to know how it varies among the numerous diverse populations in space and time. Knowledge of such variation is of course an indispensable ingredient for inter-disciplinary research if we are ever to truly understand the nature of this precious gift that is uniquely our own.

References

Anderson, Stephen R., and David W. Lightfoot. 2002. *The Language Organ: Linguistics as Cognitive Physiology*. Cambridge & New York: Cambridge University Press.

Arbib, Michael A. (ed.) 2006. *Action to Language via the Mirror Neuron System*. Cambridge & New York: Cambridge University Press.

Boesch, Christophe. 1991. Teaching among wild chimpanzees. *Animal Behaviour* 41.3:530-532.

Boesch, Christophe. 2005. Joint cooperative hunting among wild chimpanzees: taking natural observations seriously. *Behavioral and Brain Sciences* 28.5:692-693.

Bolhuis, Johan J., and Clive D. L. Wynne. 2009. Can evolution explain how minds work? *Nature* 458:832-833.

Bolinger, Dwight L. 1961. Syntactic blends and other matters. *Language* 37.3:366-381.

Call, Josep, and Michael Tomasello. 2008. Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences* 12.5:187-192.

Catani, Marco, Derek K. Jones, and Dominic H. ffytche. 2005. Perisylvian language networks of the human brain. *Annals of Neurology* 57.1:8-16.

Cavalli-Sforza, Luigi Luca, and Marcus W. Feldman. 2003. The application of molecular genetic approaches to the study of human evolution. *Nature Genetics Supplement* 33:266-275.

Cavalli-Sforza, Luigi Luca, Alberto Piazza, Paolo Menozzi, and Joanna L. Mountain. 1988. Reconstruction of human evolution: bringing together genetic, archeological and linguistic data. *Proceedings of the National Academy of Sciences* 85.16:6002-6006.

Chomsky, Noam. 1980. On cognitive structures and their development: a reply to Piaget. *Language and Learning: The Debate between Jean Piaget and Noam Chomsky*, ed. by Massimo Piattelli-Palmarini, 35-52. Cambridge: Harvard University Press.

Churchland, Paul M., and Patricia S. Churchland. 2000. Foreword to 2nd edition of *The Computer and the Brain*. New Haven: Yale Nota Bene.

Crick, Francis. 1989. The recent excitement about neural networks. *Nature* 337:129-132.

Culicover, Peter W., and Ray Jackendoff. 2005. *Simpler Syntax*. Oxford & New York: Oxford University Press.

De Waal, F. B. M. 1998. *Chimpanzee Politics: Power and Sex among Apes*. Baltimore: Johns Hopkins University Press.

De Waal, F. B. M. 2005a. *Our Inner Ape: A Leading Primatologist Explains Why We Are Who We Are*. New York: Riverhead Books.

De Waal, F. B. M. 2005b. A century of getting to know the chimpanzee. *Nature* 437:56-59.

Dewdney, Alexander Keewatin. 1993. *The Tinkertoy Computer and Other Machinations*. New York: W. H. Freeman and Company.

Dick, Frederic, Elizabeth Bates, Beverly Wulfeck, Jennifer Aydelott Utman, Nina Dronkers, and Morton Ann Gernsbacher. 2001. Language deficits, localization, and grammar: evidence for a distributive model of language breakdown in aphasic patients and neurologically intact individuals. *Psychological Review* 108.4:759-788.

Dick, Frederic, Ayse Pinar Saygin, Suzanne Moineau, Jennifer Aydelott, and Elizabeth Bates. 2004. Language in an embodied brain: the role of animal models. *Cortex* 40.1:226-227.

Dobzhansky, Theodosius. 1973. Nothing in biology makes sense except in the light of evolution. *The American Biology Teacher* 35.3:125-129.

Donchin, Emanuel. 2006. fMRI: Not the only way to look at the human brain in action. *APS Observer* 19.11.

Evans, Nicholas, and Stephen C. Levinson. 2009. The myth of language universals: language diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32.5:429-492.

Friederici, Angela D., Manuela Friedrich, and Anne Christophe. 2007. Brain responses in 4-month-old infants are already language specific. *Current Biology* 17.14:1208-1211.

Gauchat, Louis. 1905. L'unité phonétique dans le patois d'une commune. *Aus Romanischen Sprachen und Literaturen: Festschrift Heinrich Morf*, 175-232. Halle: Niemeyer.

Gong, Tao. 2009. *Computational Simulation in Evolutionary Linguistics: A Study on Language Emergence*. Taipei: Institute of Linguistics, Academia Sinica.

Gould, Stephen Jay, and Elisabeth S. Vrba. 1982. Exaptation—a missing term in the science of form. *Paleobiology* 8.1:4-15.

Grabner, Roland H., Daniel Ansari, Karl Koschutnig, Gernot Reishofer, Franz Ebner, and Christa Neuper. 2009. To retrieve or to calculate? Left angular gyrus mediates the retrieval of arithmetic facts during problem solving. *Neuropsychologia* 47.2: 604-608.

He, Chao, Lisa Hotson, and Laurel J. Trainor. 2007. Mismatch responses to pitch changes in early infancy. *Journal of Cognitive Neuroscience* 19.5:878-892.

Hermann, Eduard. 1929. Lautveränderungen in der Individualsprache einer Mundart. *Nachrichten der Gesellschaft der Wissenschaften zu Göttingen* 11:195-214.

Herrmann, Esther, Josep Call, María Victoria Hernàndez-Lloreda, Brian Hare, and Michael Tomasello. 2007. Humans have evolved specialized skills of social cognition: the cultural intelligence hypothesis. *Science* 317:1360-1366.

Hurford, James R. 2003. The language mosaic and its evolution. *Language Evolution*, ed. by Morten Christiansen & Simon Kirby, 38-57. Oxford & New York: Oxford University Press.

Jackendoff, Ray S. 2002. *Foundations of Language: Brain*, *Meaning*, *Grammar*, *Evolution*. Oxford & New York: Oxford University Press.

Jacob, François. 1977. Evolution and tinkering. *Science* 196:1161-1166.

Jacob, François. 1982. *The Possible and the Actual*. Seattle: University of Washington Press.

Johnson, Jacqueline S., and Elissa L. Newport. 1989. Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology* 21:60-99.

Kay, Paul, and Willett Kempton. 1984. What is the Sapir-Whorf hypothesis? *American Anthropologist* 86:65-79.

Kim, Karl H. S., Norman A. Relkin, Kyoung-Min Lee, and Joy Hirsch. 1997. Distinct cortical areas associated with native and second languages. *Nature* 388:171-174.

Kuhl, Patricia K. 2004. Early language acquisition: cracking the speech code. *Nature Reviews Neuroscience* 5.11:831-843.

Kutas, Marta, Kara D. Federmeier, Jenny Staab, and Robert Kluender. 2007. Language. *Handbook of Psychophysiology* (3rd edition), ed. by John T. Cacioppo, Louis G. Tassinary & Gary G. Berntson, 555-580. Cambridge & New York: Cambridge University Press.

Labov, William. 1994. *Principles of Linguistic Change*, Vol. 1: *Internal Factors.* Oxford: Blackwell.

Lamb, Sydney M. 1999. *Pathways of the Brain: The Neurocognitive Basis of Language*. Amsterdam & Philadelphia: John Benjamins.

Lau, Ellen F., Colin Phillips, and David Poeppel. 2008. A cortical network for semantics: (de)constructing the N400. *Nature Reviews Neuroscience* 9.12:920-933.

Lenneberg, Eric H. 1967. *Biological Foundations of Language*. New York: Wiley.

Lee, Chia-Ying, Jie-Li Tsai, Yu-Chin Chiu, Ovid J. L. Tzeng, and Daisy L. Hung. 2006. The early extraction of sublexical phonology in reading Chinese pseudocharacters: an event-related potentials study. *Language and Linguistics* 7.3:619-635.

Lee, Chia-Ying, Jie-Li Tsai, Wen-Hsuan Chan, Chun-Hsien Hsu, Daisy L. Hung, and Ovid J. L. Tzeng. 2007. Temporal dynamics of the consistency effect in reading Chinese: an event-related potentials study. *Neuroreport* 18.2:147-151.

Lee, Chia-Ying. 2008. Rethinking of the regularity and consistency effects in reading. *Language and Linguistics* 9.1:177-186.

Li, Ping, Li Hai Tan, Elizabeth Bates, and Ovid J. L. Tzeng. (eds.) 2006. *The Handbook of East Asian Psycholinguistics*, Vol. 1: *Chinese*. Cambridge & New York: Cambridge University Press.

Lieberman, Philip. 2005. The pied piper of Cambridge. *The Linguistic Review* 22.2-4:289-301.

Maynard Smith, John, and Eörs Szathmáry. 1999. *The Origins of Life: From the Birth of Life to the Origin of Language*. Oxford & New York: Oxford University Press.

Meltzoff, Andrew N., and Jean Decety. 2003. What imitation tells us about social cognition: a rapprochement between developmental psychology and cognitive neuroscience. *Philosophical Transactions of the Royal Society B: Biological Sciences* 358:491-500.

Meltzoff, Andrew N., and M. Keith Moore. 1977. Imitation of facial and manual gestures by human neonates. *Science* 198:75-78.

Neumann, John von. 1958. *The Computer and the Brain*. New Haven: Yale University Press.

Öhman, Sven. 2007. *The Essence of Language: A Philosophical Problem: Why Noam Chomsky was Never a Linguist*. Nora, Sweden: Nya Doxa.

Osvath, Mathias. 2009. Spontaneous planning for future stone throwing by a male chimpanzee. *Current Biology* 19.5:R190-R191.

Pinker, Steven. 1994. *The Language Instinct*. New York: Morrow and Company.

Polich, John. 2007. Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology* 118.10:2128-2148.

Polich, John, Gregory McCarthy, William S-Y. Wang, and Emanuel Donchin. 1983. When words collide: orthographic and phonological interference during word processing. *Biological Psychology* 16.3-4:155-180.

Pollard, Katherine S. 2009. What makes us human? *Scientific American* 300:44-49.

Postal, Paul M. 2004. *Skeptical Linguistic Essays*. Oxford & New York: Oxford University Press.

Postal, Paul M. 2009. The incoherence of Chomsky's 'Biolinguistic' ontology. *Biolinguistics* 3.1:104-123.

Ramachandran, Vilayanur S. 2004. *A Brief Tour of Human Consciousness: From Imposter Poodles to Purple Numbers*. New York: Pi Press.

Rivera-Gaxiola, Maritza, Juan Silva-Pereyra, and Patricia K. Kuhl. 2005. Brain potentials to native and non-native speech contrasts in 7-and 11-month-old American infants. *Developmental Science* 8.2:162-172.

Rizzolatti, Giacomo, and Laila Craighero. 2004. The mirror-neuron system. *Annual Review of Neuroscience* 27:169-192.

Ross, John Robert. 1979. Where's English? *Individual Differences in Language Ability and Language Behavior*, ed. by Charles J. Fillmore, Daniel Kempler & William S-Y. Wang, 127-163. New York: Academic Press

Sapir, Edward. 1929. The status of linguistics as a science. *Language* 5.4:207-214.

Schoenemann, P. Thomas, and William S-Y. Wang. 1996. Evolutionary principles and the emergence of syntax—Commentary on Müller: innateness, autonomy, universality. *Behavioral and Brain Sciences* 19.4:646-647.

Shafer, Valerie L., Mara L. Morr, Hia Datta, Diane Kurtzberg, and Richard G. Schwartz. 2005. Neurophysiological indexes of speech processing deficits in children with specific language impairment. *Journal of Cognitive Neuroscience* 17.7:1168-1180.

Shafer, Valerie L., and Karen Garrido-Nag. 2007. The neurodevelopmental bases of language. *Blackwell Handbook of Language Development*, ed. by Erika Hoff & Marilyn Shatz, 21-45. Malden & Oxford: Blackwell.

Siok, Wai Ting, Charles A. Perfetti, Zhen Jin, and Li Hai Tan. 2004. Biological abnormality of impaired reading is constrained by culture. *Nature* 431:71-76.

Siok, Wai Ting, Zhendong Niu, Zhen Jin, Charles A. Perfetti, and Li Hai Tan. 2008. A structural-functional basis for dyslexia in the cortex of Chinese readers. *Proceedings of the National Academy of Sciences* 105.14:5561-5566.

Siok, Wai Ting, Paul Kay, William S-Y. Wang, Alice H. D. Chan, Lin Chen, Kang-Kwong Luke, and Li Hai Tan. 2009. Language regions of brain are operative in color perception. *Proceedings of the National Academy of Sciences* 106.20:8140-8145.

Tan, Li Hai, Alice H. D. Chan, Paul Kay, Pek-Lan Khong, Lawrance K. C. Yip, and Kang-Kwong Luke. 2008. Language affects patterns of brain activation associated with perceptual decision. *Proceedings of the National Academy of Sciences* 105.10: 4004-4009.

Tan, Li Hai, John A. Spinks, Guinevere F. Eden, Charles A. Perfetti, and Wai Ting Siok. 2005. Reading depends on writing, in Chinese. *Proceedings of the National Academy of Sciences* 102.24:8781-8785.

Tomasello, Michael. 1995. Language is not an instinct. *Cognitive Development* 10.1: 131-156.

Tomasello, Michael. 2003. *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge: Harvard University Press.

Tzeng, Ovid J. L., and William S-Y. Wang. 1983a. The first two R's: the way different languages reduce speech to script affects how visual information is processed in the brain. *American Scientist* 71.3:238-243.

Tzeng, Ovid J. L., and William S-Y. Wang. 1983b. Search for a common neuro-cognitive mechanism for language and movements. *American Journal of Physiology* 246: R904-R911.

Wang, Feng, Yaching Tsai, and William S-Y. Wang. 2009. Chinese literacy. *The Cambridge Handbook of Literacy*, ed. by David R. Olson & Nancy Torrance, 386-417. Cambridge & New York: Cambridge University Press.

Wang, William S-Y. 1978. Explorations in language evolution. Diamond Jubilee Lecture. Osmania University. First published in 1982 as Supplement to Vol. 8 of *Osmania Papers in Linguistics*.

Wang, William S-Y. 1984. Organum ex machina. *Behavioral and Brain Sciences* 7.2: 210-211.

Wang, William S-Y. 2007. The language mosaic and its biological bases. *Journal of Bio-Education* 2.1:8-16.

Weinreich, Uriel, William Labov, and Marvin Herzog. 1968. Empirical foundations for a theory of language change. *Directions for Historical Linguistics: A Symposium*, ed. by Winfred P. Lehmann & Yakov Malkiel, 95-188. Austin: University of Texas Press.

Winawer, Jonathan, Nathan Witthoft, Michael C. Frank, Lisa Wu, Alex R. Wade, and Lera Boroditsky. 2007. Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences* 104.19:7780-7785.