

# Cognitive Critique



## BOOK REVIEW: MACNEILAGE'S *THE ORIGIN OF SPEECH*

**JAMES J. JENKINS**

*Department of Psychology,  
University of South Florida at Tampa*

*E-MAIL: j3cube@aol.com*

Accepted 28 May 2010

Following the much-celebrated year of the bicentennial of Darwin's birth, and the sesquicentennial of his masterwork, *The Origin of Species*, it is a privilege to review such an important work in the Darwinian tradition. The title of the book, *The Origin of Speech*, is deliberately and self-consciously chosen to evoke its inspiration. In this work Peter MacNeilage (2009) has integrated almost 50 years of his own research in psychology and linguistics and his extensive and critical reading of the research and theories of others to construct an account of the evolution of human speaking. In addition, although it is not his chief aim, he hints at the importance of this evolution as a first step in developing an account of the language faculty itself.

This is not a timid undertaking. MacNeilage gives fair warning in his opening chapter that he means to hew to the Darwinian line.

“It is my intention in this book, to give an account of the evolution of speech that unflinchingly adheres to a Neodarwinian perspective --- that contends, in short, that speech didn't just *happen* by means of a secular miracle but, instead, evolved by descent with modification in accordance with the principle of natural selection.” (p 17)

His foil throughout the book is what he calls the *Classical position* of Plato, Descartes, Saussure and Chomsky. He sees this position as asserting that speech and language are special forms, unique to humans. Although such forms are said to be genetically determined or innate in some unspecified manner, they are held to be without evolutionary predecessors. Thus, MacNeilage sets up two possible roots for the origin of speaking, one, Darwinian, (functionalist) and the other Classical (formalist). He takes as his aim the careful, detailed explication of the former and the rejection of the latter as non-scientific.

MacNeilage attributes the framework of his analysis to Tinbergen, the famous ethologist. With regard to any function Tinbergen (1952) asked:

1. How does it work? What are the mechanisms?
2. What does it do for the organism? How does it affect the organism's capabilities to survive and reproduce?
3. How does it get that way in development? What genetic and epigenetic factors guide its growth?
4. How did it get that way in evolution? How does the history of the species help us understand the structure of the trait?

MacNeilage believes that a basic biological orientation must be committed to finding serious answers to these questions. To shrug these questions off by retreating into the competence/performance distinction and ignoring them as *mere performance* problems, is simply unscientific and a regression into the age-old mind-body distinction.

He carefully outlines his argument and the structure of the work. The book is divided into seven parts, each consisting of two or three chapters. Part 1, as suggested above, sets out the philosophical issues between the two orientations and outlines the author's position. Chomsky is chosen as the antagonist to characterize the position of many modern linguists in refusing to come to grips with the evolutionary questions. MacNeilage concedes that Chomsky's position has softened a little (see Hauser, Chomsky and Fitch, 2002) but sees him and his followers as still rejecting serious consideration of evolutionary issues and finding no place for these issues in their view of syntax and phonology.

In Part 2, MacNeilage's characterization of speaking is laid out in detail. The basic unit is the syllable. This unit constitutes the frame into which the consonants and vowels are inserted as content. Thus, he calls his formulation the frame/content approach. The syllable carries the readily observed peaks of sonority due to the vowel which characterizes the most open position of the mouth. The consonants in turn result from the constrictions and changes of the glottis, lips, tongue and jaw preceding and following the vowel opening. The most prominent motor characteristic of the syllable is the oscillation of the mandible as the jaw opens and closes during speaking.

Whence comes this behavior in the *deep time* of evolution? Unfortunately, until the invention of sound recording, speaking left no historic traces. Consequently we must study current members of related species of primates that presumably departed from our family tree at earlier times. Here we may search for clues in their nervous systems and their behavioral characteristics.

There is a growing consensus that speech did not arise from primate vocal calls, many of which tend to be emotional and perhaps reflexive. This view, which MacNeilage accepts, looks instead to the motor behaviors involved in chewing, licking and sucking, all of which are biphasic cyclic activities which result in the subsequent communicative acts of lip smacks, tongue smacks, lip protrusion, tongue protrusion, and teeth chatters. This approach is congruent with the current orientation in psychobiology towards the importance of embodiment in the understanding of perceptual and cognitive abilities. (Embodiment holds that the systems of the brain out of which the mind arises are in large part generated by the experiences and capabilities of the body systems operating in the world.)

The Classic view is, of course, that there are innate distinctive features that are the basic units of speech. MacNeilage regards features as convenient units of taxonomic analysis but rejects them as innate units of mental structure or behavioral atoms. He points out that language surveys show no evidence of converging on a fixed number of distinctive features. Instead, such surveys reveal an astonishing variety of speech sounds distributed continuously along various parameters. Fifty percent of speech sounds occur in only one language and no single sound appears in all languages. (See Ladefoged, 2006, and Ladefoged and Maddieson, 1996.) This is an unlikely outcome if distinctive features are supposed to be innate. Other evidence in studies of speech errors strongly suggests that

speech sounds move as phonetic units to corresponding places in adjacent syllables, thus supporting both the notion of the syllable frame and the functional unity of the phoneme.

MacNeilage further notes that the generative approach has no time dimension in either evolutionary time or in developmental time. The mental systems are said to appear full blown at some one point in the species and to manifest themselves in infant development as soon as performance capabilities permit. Current genetics knows no parallel to such phenomena in complex behaviors or traits and current biological thinking is quite incompatible with such a concept. Recent views on development focus on the organism acting in its environment and emphasize dynamical systems and principles of self-organization. The genetic heritage is seen not as a blueprint that specifies everything in advance, but as a recipe in which the relevant ingredients intermix in sequences that interact with each other and with environmental events throughout the course of development. Gene with gene, gene with organism, and gene with environment interactions are everywhere present.

In Part 3, MacNeilage spells out his view of the developmental sequence in the child. The basic issue is seen to be Lashley's (1951) old problem of serial order in behavior. MacNeilage answers this with the frame which he sees emerging in babbling. The preferred form is consonant-vowel (CV) here just as it is the most common syllable in most languages of the world. He argues that babbling is already somewhat mimetic i.e. imitative. Research shows that infants match speaking faces with heard vowels, and spontaneously imitate tongue and mouth gestures. Further, they show a lower rate of nasals in babbling than would be expected by chance, echoing the low frequency of nasals in speech inventories and use.

Data analysis shows that in babbling and first words the pairing of consonants and vowels is not independent: labial consonants tend to go with central vowels (*ba ba*); coronal consonants go with front vowels (*dee dee*) and velars are paired with back vowels (*go go*). This correlation is also true of VC pairings. MacNeilage regards the labial consonant-central vowel as the *pure frame* because it involved only mandibular oscillation. The other two frames follow rapidly, one with the tongue forward and one with the tongue back. These popular frames are taken as evidence that this is a stage of pure frames, not yet the assembly of independent units.

When frames are assembled in sequences as the next stage begins, it is argued that such constructions are easier when they start

with a pure frame such as the labial consonant-central vowel syllable. Data from first words support this claim. Finally, there is a shift away from the high frequency of reduplicative syllables to variegated syllables characteristic of adult speech. This shift develops as the learner begins to acquire more words and is forced to achieve the more varied syllabic production. To achieve this, MacNeilage appeals to a general purpose mimetic capacity in humans. (See Donald, 2001.) The argument is that word growth begins with traditional baby talk (mama and papa as the canonical forms) and then proceeds to elaborate via mimesis under the pressure to develop more word possibilities.

The origin of words themselves is believed to arise from important social behaviors such as vocal grooming (see Dunbar, 1996). This in turn leads to the pairing of sounds with already existing concepts. There is ample evidence of the existence and use of concepts in chimpanzees and gorillas so it is a question of environmental pressures, the general purpose mimetic abilities and the ability to produce varied utterances that made the first words possible. MacNeilage suggests that the big breakthrough took place in the family unit as a result of these abilities and the increased time of nurturance required by the neonate, compared to related species. Following Jakobson (1960) and Murdock (1959) he regards the terms for mother and father as candidates for first words and first contrasts; nasal stops for mother and oral stops for father. There is abundant evidence that these terms are omni-present in today's languages and, indeed, it appears that they are frequently reinvented as languages change over time.

Part 4 is devoted to brain organization and the evolution of speech. This section begins with a tutorial on the brain that will help the uninitiated follow the discussion. Current literature suggests, contrary to earlier beliefs, that related primates have predominantly left hemisphere dominance for controlled routine motor behaviors. The great apes appear to be right handed and perhaps right footed. Examination of the literature on our primate relatives suggests that the area governing vocal calls in other primates is not located in a homologous area to that which governs human speech. However, communicative non-emotional behavior in primates is in the left hemisphere both for production and perception of the various oral-facial gestures (lip smacks, tongue smacks etc.). Further, recent neurological evidence has revealed that there are mirror neurons

in this region involved in both ingestive behaviors and visuofacial communicative behaviors. (See Rizzolatti and Craighero, 2004)

In further discussion of the neurological capacity of humans, MacNeilage finds a candidate area for the generation of frames, the supplementary motor area, located above the sensory-motor strip in the left hemisphere. When this area is galvanically stimulated in brain explorations in humans, it yields repetitive, cyclical, motor and speech behaviors which last beyond the duration of the stimulation. No other area of the brain is known to have such response to stimulation. These findings and the steadily increasing capacity for general purpose mimetic ability furnish the neurological foundation for the emergence of the articulatory skills found in speech.

Part 5 is devoted to a critique of generative phonology and its inability to deal with either the development of speech in the human child or the origins of speech. The shortest form of the argument is that phonology is basically descriptive but unscientific, looking for regularities and then using the regularities as *rules* to *explain* the same data. He notes again the lack of real cross-language solutions to the nature and number of distinctive features and comments on the poverty of the notions of markedness. He suggests that linguists accept phonetic data when it confirms their beliefs and reject such data as *mere performance data* when it disagrees with their *rules*.

Part 6 tackles questions concerning the nature of sign language. Can the existence of sign language be taken as evidence that the externalization of language is modality independent as Chomsky asserts? MacNeilage examines the non-parallel characteristics of vocal-auditory language and the manual-visual form and concludes that they are fundamentally different. The speaking code is linear and sequential. The manual system is simultaneous. Although both are babbled under appropriate circumstances, they are not synchronous in onset or in progressive development, nor is their recognition based on the same rhythmic structures. Further, comprehension of sign language seems to depend to a greater extent on right hemisphere properties than does speech.

Part 7 assays a review of the argument in terms of Tinbergen's fourth question: How did speech get there phylogenetically? First, MacNeilage points to the unlikelihood of direct genetic control of any aspects of universal grammar or any other specific gene-to-particular-phenomenon of language. Although the gene FOXP2 was briefly considered to be an instance of such a connection, it now

appears to have a much more general sphere of influence, general motor control. No other candidates are in sight.

Bird song gives evidence for innateness in the selection and vocalization of specific songs, but there is no evidence that humans have anything like a parallel language-specific innateness. A sparrow will never sing like a lark, no matter what its environment or training. But no one believes that a Japanese baby would not learn English if raised by an English-speaking family. Birdsong does, however, show a frame/content organization which suggests some form of convergent evolutionary device to solve the problem of sustained, repeated vocalization. Humans do, in addition, show innate imitation of facial expressions and, particularly, movements of the tongue and mouth.

Over evolutionary time the oral-facial gestures and phonation are believed to serve in vocal grooming, facilitate infant-parent vocal interaction and labeling, and eventually lead to the coupling of sounds and concepts. This was the monumental social discovery that ultimately was transmitted as part of culture (a meme) and replicated itself through the general mimetic capacity of the species, perhaps even giving impetus to the enhancement of working memory capacity through the phonological loop (Baddeley, 1986). The overall picture is one of bodily functions that permit certain kinds of actions being recruited in the service of social needs and consequent selection advantages in adaptation. All of this leads to further interactions of genes and memes and to the eventual result of language.

Finally, MacNeilage concludes:

“I hope the Darwinian approach to the evolution of speech I have presented here will become part of the framework enabling the phonological component of speech to enter the mainstream of modern science where it deserves to be, considering its importance in getting us to be who we are.” (p. 334).

## EVALUATION

Why should a cognitive scientist be interested in this book?

First, it is a serious scholarly work; a study integrated across many areas in linguistics, psychology, ethology, neurology, genetics, and epigenetics. It includes a valuable set of references in these fields (30 pages with approximately 500 citations) to which the reader is directed for further information and evidence concerning the author's claims. In this reviewer's opinion the book is much more soundly based in data than most works in *evolutionary psychology* that are being offered today.

Second, it provides a plausible and persuasive account of the origin, evolution and development of speaking. Many current linguists and psychologists (following Chomsky) ignore the challenge of accounting for the origins of language or discount the problem as being completely and permanently beyond investigation. MacNeilage argues that we must take a biological approach and concern ourselves with these questions. In his view it is a question of whether we are going to win a place for our fields in modern science or languish in the role of describers and classifiers in the old Linnaeus tradition, leaving others to explain the regularities that we find. Recent advances in genetics, microbiology and brain sciences are revolutionizing our understanding of behavioral matters. The fields are being massively rewritten every decade. Current literature can scarcely keep up with the discoveries being made in these fields. We must not fall behind.

Third, this book explicitly challenges much of current linguistic thought and practice at a basic level. It argues that linguistic explanation is often fundamentally circular, a process that many psychologists and linguists have objected to for years. The linguist searches for regularities and, having found them, appeals to his generalization in the form of a rule as the explanation for the data. The observation and the generalization are of key interest, of course, but the subsequent *explanation by rule* is non-causal. MacNeilage argues against the casual acceptance of distinctive features and markedness, as if they were universal realities of speaking. In MacNeilage's opinion, surveys of the world's languages fail to confirm the universality hypothesis. He finds little real support for modern phonology and recommends more attention to phonetics and less to abstract, supposedly innate, categories that are sometimes only distantly related to observable data.



Many linguists and psycholinguists will object strenuously to MacNeilage's treatment of phonology, especially as regards distinctive features. That, it seems to this reviewer, is a matter of data. No one, including Ladefoged himself, has been able to advance a universal set of features that will fit the languages for which we have the necessary descriptions. Future work may settle the questions. Many will also protest MacNeilage's conclusion that sign language is fundamentally different in kind from spoken language. From MacNeilage's point of view, the burden falls on those arguing for the overall similarity. His argument is that there are important differences in surface features of the behaviors involved and that speaking and signing are not automatically evidence for a common underlying faculty of language.

Parts of the book, though interesting, seem to stray away from the central argument. Parts Five and Six are directed at counter-arguments that the reader may or may not be concerned with. For the reader interested in the evolutionary account, the first four parts are the crucial ones.

All accounts of evolutionary development are to some extent *Just So Stories*. Rigorous proof is not possible in most cases of complex traits. Every story must assemble what evidence it can find into a plausible account. This book does a masterful job of assembling and interpreting all of the evidence we have concerning the evolution of speaking. In the long run it may not be the final word, but until we have a better story, this is the one that must be the prime contender.

## ACKNOWLEDGMENTS

The substance of this review appeared earlier on the Linguist List, an electronic server, and is reprinted here with minor changes by permission of the List

**REFERENCES**

- Baddeley AD (1986) Working memory. Clarendon Press, London UK
- Donald M (2001) A mind so rare. Norton, New York, NY
- Dunbar RIM (1996) Grooming, gossip, and the evolution of language. Harvard University Press, Cambridge, MA
- Hausser MD, Chomsky N, Fitch WT (2002) The faculty of language: what is it, who has it, and how did it evolve? *Science* 298: 1569-1579.
- Jakobson R (1960) Why “mama” and “papa”. In: Caplan B, Wapner S (eds) Essays in honor of Heinz Werner. International Universities Press, New York, NY pp124-134
- Ladefoged P (2006) Features and parameters for different purposes. <http://www.linguistics.ucla.edu/people/ladefoged/PLfeaturesParameters.pdf>
- Ladefoged P, Maddieson I (1996) The sounds of the world's languages. Blackwell, Oxford, UK
- Lashley KS (1951) The problem of serial order in behavior. In: Jeffress LA (ed) Cerebral mechanisms in behavior. Wiley, New York, NY pp 112-136
- MacNeilage PF (2008) The origin of speech. *Studies in the evolution of language*, no. 10. Oxford University Press, Oxford, UK
- Murdock GP (1959) Cross-language parallels in parental kin terms. *Anthropological Linguistics* 1: 1-5
- Osgood CE, Sebeok TA (1954) Psycholinguistics a survey of theory and research problems. Waverly Press, Baltimore, MD
- Rizzolatti G, Craighero L (2004) The mirror-neuron system. *Annu Rev Neurosci* 27: 169-192
- Tinbergen N (1952) Derived activities: their causation, biological significance, origin and emancipation during evolution. *Q Rev Biol* 27: 1-32