Nils L. Wallin

BIOLOGICAL ASPECTS OF THE RELATIONSHIPS BETWEEN MUSIC AND LANGUAGE

Unesco and the International Council of Music have begun work on a musicological project of considerable extent, since it is a universal history of music in ten volumes. At present, the provisional title is *Music as a Language of Man: A World History of Music.*

For some time, however, the experts assigned to the preparation of the work have begun to express doubts about the validity of this title. If the project is global in scope, it is certain that there is not *one* music but a large number of *musics* in the world, a plurality that is not easily covered in a single title. The provisional title has thus been changed to *Music in the Life of Man: A World History.*

The hesitation, as we see, is not concerned with the fundamental problem, namely, if music is properly speaking a language. In general, the existence of a musical language is not questioned. Scientific works concerning musical structures are rather closely

Translated by Jeanne Ferguson

https://doi.org/10.1177/039219218303112201 Published online by Cambridge University Press

influenced by the semantic and linguistic methods. Music as a structure is considered to have an immanent communicative function. Also set aside are the artisanal aspect of music as a product and its bi-social role, different from that of language. The intelligence of a normal child isolated from all oral communication will be atrophied if his isolation ends only after the age in which the speech faculty would have reached its full maturity. On the other hand, a child whose musical gifts have not been developed will have no problem attaining what is called a complete vital capacity.

In saying this, I do not want to suggest that music is without importance for the flowering of the mind. I merely say that without doubt there is an ontological and genetic difference between music and speech. Perhaps Noam Chomsky was correct when he said there is no innate language. However, speech faculty is well-localized in the brain, as we shall see, from the functional point of view as well as anatomical. It is not the same with music. Music is a very complex ensemble of qualities that exist in varying degrees in individuals, differently from what occurs with the speech faculty; it may go quantitatively from zero to the ability to conceive and hear all the subtleties of the Great Fugue by Beethoven. A geneticist would thus ask himself if musical competence is a polygenetically conditioned quality.

By definition, language is a function of communication which means, according to the latest hypotheses of research in communication, that it is "double-tracked." The fact that the human brain has structures that are specifically oriented toward speech indicates that the two speakers, the sender and the receiver, are homo-typical at the time of sending and receiving. It also means that they have common frames of cognitive reference.

Is the same thing true for music? The salmon fisher when he approaches his favorite fishing spot hears far off in the forest whether the water of the stream is high or low. The river informs the fisher about its condition through the sound of the water. Music, a stream of sounds, tells the listener, a fisher of sounds, something of the condition of the composer. But what music tells depends on a genetic combination that cumulatively gives what we call "musicality." Senders and receivers are only analogous and in the best of hypotheses have nothing between them other than analogous frames of reference. In this case it is a matter of information, and "information" is not a synonym for "communication."

Music may be, after all, only an index system that we read the way a hunter reads the tracks of game or the doctor the symptoms of an illness. Music, as such, is not a "double-track" phenomenon.

Even so, it is a matter of semiology.

Plato said that man is a biped without feathers, to which Diogenes insolently retorted, brandishing a plucked rooster under the noses of Plato's disciples, "Is this a man?"

No, of course not, but today we may affirm, without the risk of shocking anyone, that they have a common basis. Let us then pluck the rooster, like Diogenes, and brandish this semiological rooster, this linguistic-musical hybrid; let us make a sort of "pre-semiological" experiment.

I. Two Suppositions and a Hypothesis

1. *The auditive system and its tonotopic organization*

Why must we continually return to the tonotopic* organization of the auditive system? To be sure, whether it concerns music or speech, we have a great deal of knowledge about acoustic phenomena, that is, about what goes into the ear, the input. Thanks to phonetics, linguistics and the theory of music, among other things, we are also fairly informed on what issues in the form of behavior, or the output. Is psychology able to deduce from this what happens inside the brain, that is, between input and output?

F.H.C. Crick (1979), one of the "fathers" of the theory of the double-spiral structure of the DNA molecule, puts the question with admirable clarity:

> "The basic problem is that almost any process we can study by observing overall behavior (reading, say) involves the complex interaction of many different regions of the brain, each with its own way of handling information. We know

* On the tonotopic organization of the auditive system see p. 7.

https://doi.org/10.1177/039219218303112201 Published online by Cambridge University Press

only in the barest outline how these distinct regions should be recognized and classified. Although our knowledge of how they are interconnected is growing rapidly, it is far from complete, both qualitatively and quantitatively. Moreover, we seldom know what operation each region is performing, that is, what relates the outputs to the inputs, and in some instances we do not have even the faintest idea of what is going on.

This is the main reason pure psychology is, by the standards of hard science, rather unsuccessful. It is not that it cannot be quantitative. The branch of it called, rather curiously, psychophysics is certainly quantitative, often quantitative in a sensible and imaginative way. The basic difficulty is that psychology attempts to treat the brain as a black box. The experimenter studies the inputs and outputs and tries from the results to deduce the structure and operation of the inside of the box. Such an approach is not necessarily a bad one. For many years genetics was a black-box subject. It tried, with some success, to deduce the structure and function of the genetic material by studying breeding patterns. Indeed, much good biology is done by the black-box method. This can happen at all levels. To the previous generation of biochemists an enzyme was a black box. Nowadays many enzymologists study the structure of an enzyme to try to correlate the structure with the enzyme's behavior. One man's black box is another man's problem.

The difficulty with the black-box approach is that unless the box is inherently very simple a stage is soon reached where several rival theories all explain the observed results equally well. Attempts to decide among them often prove unsuccessful because as more experiments are done more complexities are revealed. At that point there is no choice but to poke inside the box if the matter is to be settled one way or the other."

The way we perceive the pitch of sounds is as important for the comprehension of speech as it is for music. There are several theories on this point. So far, none has succeeded in covering this immense area, which becomes increasingly large, especially through the progress made in ethnomusicology.

The vibrations of sounds that penetrate as far as the inner ear and the basilar membrane are transformed into electric energy.

https://doi.org/10.1177/039219218303112201 Published online by Cambridge University Press

Through the acoustic nerve this energy stimulates the neuronal nuclei of the central nervous system, mutually disposed relative to their different roles, which are to modulate information in its many aspects and direct it toward the auditive regions of the neocortex.

Up to this point, the processes take place mainly within the specific auditory system, i.e., within the pathways and nuclei which principally project auditive stimuli. Afterward, at the level of the auditive cortex, a first generalization occurs at the moment in which the information is directed toward the periphery of this region before being sent off toward those cortical regions where specific information (visual, olfactory, auditive, tactile) is coordinated with information controlling the mechanisms of memory, attention, emotions and metabolic changes. The brain then proceeds, as an integrated structure, to a complete evaluation of the information received and releases at the same time the definitive response to the stimulus, a response that will be translated into a relevant behavior.

A relevant behavior is, first of all, one that contributes to survival. The lateral system of the fish, in which many believe to see the origin of the auditive system of animals, is in the beginning only a hydrodynamic system of equilibrium and orientation important, among other reasons, for the individual behavior of the fish within the collective behavior of the shoal. Some labyrinth fishes develop inside their branchial cavities a sort of sac-shaped ampullae that allows them to perceive aerial oscillations and serves at the same time as a sort of additional respiratory organ. In the frog, the auditive system is already paired with phonation in a retroaction (feedback) analogous to the system of interconnection between hearing, breathing and vocalization found in man. The cat has an extremely sensitive auditive system that is quite similar in function to that of the higher primates and thus to ours.

The localization of sounds in space is of vital importance to nocturnal predatory animals. The cat has a prepotency for sound over light. This localization is also very important for man. The exact appreciation of the time required for sounds to reach our ears, as well as the slight variations in the intensity of the sounds, is a useful additional data with regard to visual and tactile information. Organic sounds, that is, those emitted by animals and consequently specific to each species (the species-bound key excitations, according to Konrad Lorenz) are probably first of all perceived by their amplitude, their register and their place in the acoustic spectrum, their harmonic composition, the nature of their on-set and their affective aura. Frequency discrimination enters as a general factor, but the need to determine the exact pitch of a sound is, in my opinion, only a secondary element from the point of view of global perception.

Exact pitch enters only exceptionally also into the identification of un-organic sounds. In this case, it would rather be a matter of distinctions between "high" and "low," "right" and "left," "near" or "far off," in other words, of orientation in space, whether it is a question of sounds with determined frequencies or of noises.

J.L. Flanagan (1955) claimed that in the normal register of the human voice, man does not distinguish between the frequencies of formants when the difference between them is below 60 Hertz. The discontinuity in delivery, the prosodic line of the phrase, the variations in intensity between the different frequential bands in the discourse expose phonetic clues which facilitate the discrimination and identification of linguistic elements in a continuous sequence. These indicators intersect and thus modify the phonetic qualities that characterize the isolated segment. Philip Lieberman (1975) describes what happens when a speaking subject wants to pronounce the word "bat":

> "A human speaker in producing this syllable starts with his supralaryngeal vocal tract in the shape characteristic of [b]. He does not maintain this articulatory configuration but instead moves his tongue, lips and so on, toward the position that would be attained if he were instructed to produce an isolated sustained [ae]. He never reaches these positions because he starts toward the articulatory configuration characteristic of [t] before he reaches the "steady-state" (isolated and sustained) [ae] vowel. The articulatory gestures that would be characteristic of each isolated sound are never attained. Instead they are melded together into a composite characteristic of the syllable."

Even in sonic African languages, in which the relationships of the pitch of sounds within the phrase have a semantic value (J.H.K. Nketia, 1974), these relationships are far from being uniformly coded, in the absolute sense of the term. In conclusion, we may therefore say that in linguistic perception the directional tendency of the pitches of sounds plays a greater role than does the discrimination of their exact pitch.

I have just mentioned the threshold of discrimination (difference limen) for frequencies of formants included within the normal register of speech. The threshold of discrimination in a musician may descend to around 0.2 Hertz at 1000 Hertz. Here we are confronted with a paradox that impels us to describe, although summarily, the way the brain functions with regard to the auditive field. It is what I have elsewhere called the "tonotopic paradox" (Wallin, 1982).

The primary area of the auditive field, that is, the region in which auditive information first reaches the neocortex, consists of three-dimensional modules, disposed in a linear system. In each module are found cells and groups of cells that, at a certain prefatory intensity and at a certain temporal relationship between identical stimuli to the two ears, respond maximally to a certain frequential excitation of the basilar membrane ("best frequency, BF"). The neurons and these clusters of neurons also differ in the sense that some only, or best, respond to an excitation coming from only one ear (monaural) while others respond to an exitation coming from both ears (binaural). The organization of the system is tonotopic,* which means that each point in the cochlea of the inner ear transmits its information to a corresponding point of the auditive zone of the cortex, and that the points which, in the basilar membrane, are in functional relationship with each other are represented at the level of the auditive cortex by elements placed in rapport with each other according to a system of overlapping segments.

Such a structure should permit a very fine selective discrimination. It is not that simple, however. Each cell having a BF is surrounded by a field with a maximal response at a central point and a peripheral zone whose response is more and more feeble as distance from the center increases. These fields overlap in one way or another at the surface of the module, but they are also found stratified in deep layers, partially covering the neighboring fields

https://doi.org/10.1177/039219218303112201 Published online by Cambridge University Press

situated on the same plane. If the six layers of the neocortex are penetrated, which is precisely what axons and dendrites do, aggregates of neurons are found, in horizontal as well as vertical positions. Some of the composing cells are monaural, others binaural, still others have a narrower zone of information and others broader. The extent of the receptive fields traversed varies according to the angle of penetration. If, taking into account an adequate excitation (intensity, for example), we look for the region that best responds to a given frequency, we see that the corresponding representation covers the largest part of the auditive neocortex, for the majority of frequencies. It is difficult to understand how this structure, which offers such a rich base for the localization of sounds, can at the same time serve as a basis for an extensive selective discrimination of pitch.

This is, however, what it does.

"Formalized" or sophisticated music is only a small part of what we call music. It is the result of a systematization conditioned by the cultural milieu and the historical moment, but it is also a continuous process making the musical structure more autonomous as a product of human ingenuity, or an artefact. It is strongly imbued with cognitive notions that through language indirectly link this music to mathematics and applied technology (scale systems, construction of instruments) but also to cosmological speculations. It makes a rigorous selection among frequencies while music, on the whole, is less demanding both in the degree of autonomy and the severity of selection, preferring to entrust itself to approximative values (I deliberately avoid the often used term "tolerance") that better correspond to the functional auditive system I have just outlined. The primary melodic configurations of all civilizations have a broad latitude in their acceptance of different pitches and intervals.

Such a structure may be called *emmelic* (Greek: "according to the melody," Wallin, 1982) as opposed to formalized music, which could be qualified as *melic*, for the purposes of the discussion. *Emmelic* could not be confused with some secondary aspect of the term "primitive." It is a matter of a certain musical behavior that could just as well be inscribed within a highly differentiated cultural context, whether it is archaic or whether it expresses a living tradition or a new tendency. This behavior stresses other expressive qualities than those that characterize musical behavior that is more systematized and autonomous (*melic*). It is typical of musical expression of the small child (Michel Imberty, 1981) in all civilizations. It is also largely the musical expression of the "musically uneducated" person, to whatever civilization he belongs. Further on in this article we will see an analogous method of classification of standard and non-standard pertaining to speech.

In a different context, I proposed that this paradoxical image of a "blurred" system could be dissolved if the auditive system were not considered as a linear and closed or deterministic system having a certain number of deviations from the fixed points that the "ear tolerates," but if:

> the selectivity of frequency discrimination responds to a standardized level that may be considered as "sufficient" relative to the habitual needs of the system concerning the cellular description of the stimulant, in other words, relative to the primordial finality of the system. The "sufficient" level may be raised or lowered according to need... The auditive field is dispositive with regard to socio-cultural needs. It is conceived for a maximal disposability on the condition that all the driving components (biological, psychological and socio-cultural) are reunited. I imagine that the neuro-biologists and psycho-acousticians of the future will one day be able to integrate all these factors into a mathematical model (Wallin, 1982).

But how do we arrive at the maximum disposability of the system? When we have resolved the paradox we will have answered the question, it seems.

Apparently, the auditive system responds differently to sounds of high and low frequencies that it receives through the acoustic nerve that leads the impulse from the ear into the central nervous system. Apparently, the acoustical field of low frequencies to which music and speech belong corresponds to temporal structures in the patterns of neural responses throughout the system, and these structures are coordinated in such a way that they correspond to types of oscillation that may be expressed in terms of relationship between multiple integrals such as 1:1, 1:2, 1:3, and so on..., 2:3, 3:4, 4:5, and so on: in other words, a coordination of oscillations corresponding to the economical principle

https://doi.org/10.1177/039219218303112201 Published online by Cambridge University Press

of the greatest cooperation possible through the greatest possible simplicity. This coordination of oscillations is obtained within the framework of a temporal analysis, conducted throughout the entire auditive system according to patterns of monaural as well as binaural responses and in an interaction between activating and inhibiting processes, a highly selective activity.

The field of low frequencies in the auditive system seems particularly sensitive to nuances of the stimulus. Music has a structure that cultivates repetitive figures (*Gestalt*) whose constitutive elements (duration, intervals, accents) seem to respond to the principle of economy mentioned above (contrarily to what happens with speech and its constantly changing structures dictated by semantics). As far as musical structures are concerned, the auditive system may thus have evolved in the course of time by acquiring or adopting, due to a continually more subtle coordination of vibrations, a variable selectivity in the form of codes that decrease redundance (Pribram). The system thus becomes dispositive.

The notion itself of a "dispositive system," however, is a socio-cultural notion. "Interaction between activating and inhibiting processes" (see above) is so far a truism that does not allow us to solve the tonotopic paradox. However, the coordination of vibrations is a notion of thermodynamic physics or of biological physics. A possible solution would be, beginning with the temporal and rhythmic coordinations we can objectively disclose in auditive phenomena, to refer to the thermodynamic theory of Prigogine *et al.* concerning "dissipative structures," a theory that has recently shown its worth in its applications to various domains in biological research, from microbiology to the macrogenetics of the populations of the world (Sinz, 1980). It is a problem to which we will briefly return at the conclusion of this article.

2. *The two hemispheres: function and localization*

Research on aphasia, that is, on the central disturbances of speech understanding and speech production, have long been prompted by the idea that complex mental phenomena are functionally linked to determined regions of the brain. Numerous clinical observations have permitted the refinement, if not the confirmation, of this theory of cerebral localizations. Since disturbances of the linguistic function at this central level very often also affect musical functions, a number of researchers have been able not only to present interesting observations on individual cases of central musical disturbances, amusia, but also to insert the study of these cases into a general description of different types of aphasia.

From the end of the 19th century, however, some neurologists have been questioning the simplistic idea of a specific localization of such complex functions in such or such portion of cellular tissue and prefer the idea of level of construction, of a functional system hierarchized according to a dynamic process. This is the case with the English neurologist, Huglings Jackson. Later, progress in brain surgery and the development of a considerable technical arsenal following the two World Wars have amply verified this more dynamic view. I have therefore chosen to consider principally the results presented by one of the most famous specialists in this field, R.W. Sperry, the author of *Forebrain Commissurotomy and Conscious Awareness* (1977), who received the Nobel Prize in 1981 for his research.

***

The right and left hemispheres of the brain in mammals are normally separated from each other but linked by nerve fibers that intersect and are found in bundles, the commissures of the brain, especially the one called the corpus callosum which, in man, comprises more than two hundred million nerve fibers. Another junction occurs at a phylogenetically very ancient level, the brain stem. Experiments made on animals during the last thirty years have shown that even if the bridge between the two hemispheres is surgically cut, each can still fulfill its functions to some degree at the level of superior activities such as perception, learning and memory. C.B. Trevarthen has shown that with an optical system of filters polarizing light, the two separated hemispheres can be induced to perceive two different objects simultaneously occupying the same place in space. The conclusion is that the conscious experience of each hemisphere is distinct from the conscious experience of the other. The absence of any conscious communication between the two hemispheres seems confirmed by the impossibility for an animal whose brain has been thus split to integrate a sensorial information that has been simultaneously given to it, partly in one and partly in the other cerebral hemisphere.

A cerebral commissurotomy is sometimes performed on subjects suffering from severe epilepsy, cutting the passage between the two hemispheres via the corpus callosum. Experiments then made with all the sensorial lateralized modal ties (that is, visual, auditive, olfactory and tactile) confirm the observations made of monkeys operated on in the same manner: each hemisphere is unaware of what the other perceives.

It has however been shown that if the operated subjects manifest a clear tendency to concentrate their attention either on the input from the left or the input from the right, each at the other's expense, the direction of the attention towards what happens here or there does not necessarily exclude all conscious experience of what is happening elsewhere: the two hemispheres can at the same time be effectively and continually occupied with different, indeed, rival, tasks and be fully aware of it.

Thus the subject who has undergone cerebral commissurotomy, "the split-brain man," seems to be able to discern two totally distinct things occupying simultaneously the same area in space, which the normal brain rejects. In conformity with the theory extolled by Gestaltism, each hemisphere endeavors automatically to complete the partial stimulus it receives so as to restore a bisymmetrical perception. Cerebral commissurotomy thus seems not only to cut the brain in two but also the mind:

> "Two separate realms of subjective awareness are apparent: one in each disconnected hemisphere, and each in itself seems to be remarkably whole, unified, and capable of supporting behavior comparable in many respects to that of the combined intact system."

Affective coloration of perception seems to be the only exception to the rule that each isolated hemisphere remains unaware of the conscious experience of the other. The emotions, in fact, are transmitted immediately from one hemisphere to the other, probably through the intermediary of the brain stem, which is not

https://doi.org/10.1177/039219218303112201 Published online by Cambridge University Press

affected by the operation. The same is true for some states of consciousness and vigilance, as well as for certain sensorial qualities. Sperry summarizes, in conclusion:

> "With the foregoing in mind, and with some further qualifications to be mentioned below, we can accept the general conclusion that brain bisection yields two conscious minds or selves within the one cranium."

He states that the two separated hemispheres function not only as though each had an independent consciousness but also had its own qualities that are not found in the other. Here we find the old argument of cerebral localization of human faculties.

Sperry brings up facts that show the existence of a profound asymmetry on the level of speech, cognition, attention, memory and sensory-motor mechanisms. The most striking of these asymmetries has reference to the localization of speech, writing, reading and calculating in the left hemisphere, while these functions are almost absent in the right hemisphere that on the contrary has other aptitudes, non-verbal and non-mathematical, but which, on the contrary

> "mostly involve spatial and visualizing abilities in which a single mental image is more effective than a long series of words. Geometrical discriminations of topological forms, for example, are performed at a high level by the right hemisphere but seem to be extremely difficult or impossible for the left hemisphere."

The right hemisphere is not, however, completely mute. It can emit isolated words or certain series of words, especially those having a metric form that can be sung or that compose a succession of spontaneous exclamations released by an emotion. The fact is confirmed by the study of observed cases of aphasia or amusia. An aphasic can sometimes sing a phrase that he cannot pronounce in speech. It also happens that he abruptly regains speech in the form of an interjection under the effect of surprise, for example, if during a conversation with his doctor a bird flies against the window (Ustvedt, 1937).

The idea that cerebral commissurotomy causes the appearance of two conscious unities that are largely independent of each other has not gone unnoticed, but opinions are divided on the conclusions to be drawn from it. Some think that the totality of the conscious self has been preserved in the patient around a center located in the left hemisphere or in some meta-system at the level of the brain stem, which has remained intact. As new facts appear, the problem changes. If it is difficult to maintain today that the right hemisphere is totally deprived of consciousness, the belief persists that it lacks self-consciousness.

To go farther into the problem, Sperry imagined methods permitting the revealing of the presence or absence in the right hemisphere of self-consciousness or social consciousness:

> "The kinds of emotional reactions that were generated and the selectivity of responses to follow-up questions of the examiners and to vocal cues from the subjects' own comments showed that true identifications were made in the right hemisphere accompanied by appropriate cognitive and conative associations."

At the present stage of the problem and considering observations gathered from cases of aphasia, it seems that current interjections such as "Oh no!," "Dear me!," or "Good Lord!" come from the right hemisphere, as do some very frequent isolated words such as "Yes/No," "Uh-hu" or "Good!" and even some words that have been introduced into the conversation by the consulting physician and thus only employ a mechanism of repetition by the patient, for example, "past," "soon," "state," "animal," and so on:

> "With each hemisphere mentally searching for the answer, the right hemisphere, needing a correct word or name to express what it has recognized visually and the left needing something more specific as a focus for the vague mental *aura*** that transfers, a correct oral cue could have instant resolving effects in both hemispheres which would be rapidly finalized."

** In the detailed presentation of each experiment, the authors use the term "aura," a term I would willingly propose for musicology, to designate the mental and emotional state that is the "*contenance*" of the work.

As to knowing the nature of this *aura*, its function and way of propagating, the report concludes that it seems to spread rapidly, at least for an important part, from one hemisphere to the other, undoubtedly through the brain stem. Aside from an overall affective charge, this transfer seems to permit certain cognitive effects that allow subtle distinctions between the general categories of thought, such as "public" and "private," "familiar" and "strange," "historical" and "episodic," etc. The experiment permits the belief that these affective and conative *auras* play a critical role in the normal functioning of the brain, especially in the mechanism of memory.

As far as the latter is concerned, it is believed that the phenomenon may be located at the level of the limbic system of the brain stem; it would be released either by a voluntary act or by an emotional shock, probably most often by a combination of the two factors.

Surgical intervention evidently provokes a certain independence between the two hemispheres, the content of consciousness proper to each being mostly unknown by the other. However, we must not forget certain fundamental factors concerning the course of the most important sensorial trajectories. A sound which has just struck the ear is registered and transmitted unilaterally from the left ear to the ipsilateral hemisphere (left) and from the right ear to the ipsilateral hemisphere (right); it is, however, also transmitted bilaterally from the left ear to the right hemisphere and inversely. The sane, healthy, normal subject thus probably registers a perfect bilateral perception. In this case, according to Sperry, we may speak of a doubling of the region of self-consciousness. If the mental activities are identical in the two hemispheres, nothing in the behavior proves that they are separate. At the most we may cautiously infer it, through hypothesis and deduction.

A hypothetical model of evolution to account for the mutual relationships between music and speech

According to present scientific data, the decisive step in the evolution of primitive vertebrates was taken when the neocortex took the relay of the inferior nervous centers that until then had the essential functions of control of sensory-motor reflex action.

In the rat, the pathway that leads from the motor region of the brain to the different muscular organs is limited, while in the cat the number of nerve fibers that take the same pathway is around 186,000 and in man around 1,200,000. Such a development requires more and more neuronal space, which explains why the human brain is larger than that of hominid monkeys. One reason for this enlargement is probably due to the speech faculty, whose development makes man a reflective being endowed with a critical sense and the power to reason, capable of controlling the results of the actions that are necessary to his survival and unceasingly improving their efficacity by making tools that are constantly bettered through a methodical reflection. That the brain of *Homo loquens* possesses not only a consciousness but a *self*-consciousness is one of the great enigmas of existence. The sound gesture of the higher animals, which we still use when we mutter, whistle or hum to express the spontaneous affective reactions that act as signal and index now serves another purpose, which is to *name*, to *describe*, and to *argue*. The genetic code registered this new cerebral function.

Neanderthal man used tools and observed ritual practices when burying his dead: these two phenomena indicate that his speech faculty was fully developed. In modern man the brain is functionally asymmetric with a clear linguistic dominance in the left hemisphere. Scarcely ten years ago Norman Geschwind and some others proved that this functional asymmetry corresponds to an anatomical asymmetry: on the one hand, the scissure of Silvius of the left hemisphere is longer and oriented somewhat differently; on the other hand, the posterior part of the left planum temporale that forms part of the Wernicke region so important for the cognitive functions of speech, is more developed on the left than on the right.

Later it was realized that this asymmetry exists in the fetus and consequently is not the result of a development of the speech faculty in the infant during its early years but of a congenital divergence, whose trace was found by Marjorie L. LeMay in the fossil crania of the Neanderthal: a bony protruberance at the edge of the inside of the cranium indicates the implacement of the Silvius scissure, and the configuration of the cranial cavity was the same then as it is today.

What that brain could perform is not however encoded. T. Dobzhansky, among others, tells us:

> "Genes make the origin of culture possible, and they are basic to its maintenance and evolution. But the genes do not determine what particular culture develops where, when and how. An analogous situation is that of language and speech—genes make human language and speech possible, but they do not ordain what will be said. ... There is no such thing as a gene for self-awareness, or for consciousness, or for ego, or for mind. These basic human capacities derive from the whole of the human genetic endowment, not from some kind of special genes."

Homo erectus appeared more than 250,000 years ago, with a cranium whose increase in volume is estimated at 4.6% in 100,000 years. His successor in the evolutionary chain, Neanderthal Man, had a cranial capacity which increased 7.5% in the same length of time, after which the volume of the brain decreased with ca. 100 cubic centimeters in modern man, either because the muscular mass to be moved is smaller today or a volumetric expansion of the brain preceded a differentiation in the species. It is possible that the rapidity of growth corresponded to the development of the speech faculty and its correlative anatomical modifications.

As far as knowing what happened during protohistory or prehistory, extreme prudence must be observed, considering recent tragic excesses. Racist theories of the 1930's collapse before the evidence of anthropology and modern genetics, which prove the fundamental unity of the human species.

This basic unity, however, is submitted to the phenomena of natural selection and genetic evolution, which are seen on the macroscopic scale in the statistically disclosable variations of intensity and orientation in human topography and on the microscopic scale by observable differences in the behavior of individuals subject to the same fundamental genetic system. This "program" does not only govern the physical aspect of development. Our DNA structure also decides the forms our hereditary behavior will take and the way in which we will adopt this code of conduct coming to us through tradition, in other words, through a culture.

In this case we cannot speak of heredity in the strict biological sense of the word except to the degree that the apprenticeship of adaptation and acculturation in the broadest sense of the term admit a genetic component, if it is true that the genome of the organism governs its capacity to acquire new knowledge. The interaction between culture and genetical heredity contributes to a program of instruction (conserving) and selection (changing).

In the preceding pages, I have mentioned the auditive mechanism of the fish and the frog. In the fish, the lateral system seems primarily to have a function of orientation and localization in space. (The definition of the lateral system as an auditive system has been disputed by some). Morphological resemblances to an auditive system are, however, striking. In the frog as in i.e. the field cricket (*gryllus campestris*), it is the functional correlation between the auditive system and the system of production of acoustic signals which mainly draw our attention. In both cases, there is an auditive selection that adapts to the sound spectrum of each species and determines the relationship between the distinctive key-excitations of the species and the releasing mechanism. While the lateral system of the fish seems devoted to its function of localization and orientation, the homologous system of the cricket and the frog have this function of orientation but add a supplementary function that concerns audition combined with phonation.

The distinctive "prosodic repertory" of the frog and the cricket is poor in comparison with the sound gestures of the evolved mammals that have, within more or less restricted limits, an audio-oral register specific to each species. It seems that this register is largely governed by the control centers of the brain stem, in liaison with the reticulate system, thus with the oldest parts of cerebral development, comprising vigilant attention, motor activity and emotions. There again we find a coordination between impulses emanating from the specific disposition of the stimuli and their decoding at reception by the auditive system. This coordination occurs through the intermediary of the neocortex. In man is found an analogous retroactive mechanism that plays an important role in the acquisition of speech and *emmelic-melic.*

We know that linguistic functions are mainly located in the

https://doi.org/10.1177/039219218303112201 Published online by Cambridge University Press

left hemisphere. At the level of the planum temporale it presents a relative hypertrophy with respect to the right temporal lobe that corresponds to this localization. Geschwind points out that such anatomical asymmetries are not peculiar to man, that they are found in the brain of the large monkeys, and that they are perhaps to some degree normal in the entire animal kingdom. In man, phenomena of interaction are produced at the level of the neocortex between the different sensorial modalities, but we have no proof that a direct correlation exists on this plane. However, the zones of convergence of sensory data in the large monkeys (in *both* hemispheres of the brain) are generally considered as homologous to the region that in man, in the left hemisphere *only*, harbors speech faculty. It must be remembered that quite near the Wernicke region are found the two zones in which are localized auditive and visual perception.

The development of the Wernicke region in the left hemisphere of the human brain would thus explain how man has been able to multiply the exchanges between the ways of perception, indeed, to furnish himself with a properly linguistic super-modality whose pathways were prepared by the anatomical evolution of several mammals, especially the large monkeys.

On this background, we can imagine an evolution starting from the fundamental biological requirements of the system—in response to the functions of localization and orientation in space—and diversifying in selective correlation with the sound production mechanisms characteristic of each species and progressive modifications of the organic anatomy, structures and functions which finally end in a cognitive and supramodal perception. For von Economo (1931) the development of the frontal lobes is a specifically human attribute. According to others who have studied the ontogenetic development of the lateral functions, the child arrives rather late at the cognitive specialization which is furnished him by the asymmetry of the left hemisphere: speech, and speech faculty would be ontogenetically presented in a bilateral way. We can see in this the trace of a bilateral expressive system, phylogenetically anterior to speech.

On the basis of known facts and suggested hypotheses I will say that, beginning with this pre-verbal system, branch off on the one hand articulated speech and on the other the archaic music

https://doi.org/10.1177/039219218303112201 Published online by Cambridge University Press

I have called *emmelic*, each profoundly rooted in its particular hemisphere but with very strong ties between them. The partition may have occurred during an evolution toward the specialization of the tasks between the two hemispheres, ending in an asymmetry that was anatomical as well as functional. The left hemisphere develops pre-eminently a supramodal perception, analytically cognitive, while the right hemisphere retains and develops a holistic and more appositional way of perception depending on associations by contiguity, strongly colored with affective values, corresponding to the oldest connections of the cerebral system.

The two expressive systems do not correspond exactly on the space-time level and do not quite use in the same way the resources of the dispositive field. Musical perception is relatively free from the semantic servitude imposed by the rapid identification of external stimuli that are constantly changing. Slower than linguistic perception, it uses in a particular manner the temporal processes of the relays and ways of transmission. During an evolution covering millennia, it has adapted so as to respond to its specific type of stimulus by a spatio-temporal conduct or control (oscillatory) of cellular structure. Linguists and neurologists agree on emphasizing the distance between the conditions for emission of vowels and those for explosive consonants. While the latter are exclusively, or almost so, identified by the left hemisphere, vowels are much less clearly localizable. In my opinion, this goes back to the difference between acoustic behavior of the sounds: the speed of modifications of the formants in explosive consonants is very high, while the formants of the vowel are steady, which could in the end go back to a different spatio-temporal behavior as far as concerns the way the ear perceives in the dispositive system.

II. Examination of Different Models of Perception

In 1971 two Americans, J.E. Bogen and H.W. Gordon, presented interesting observations connected with surgical interventions they had performed on the brain. So that the operation might not damage the linguistic functions of the patient, it was necessary to localize them beforehand. The fact of being left- or right-handed is not always sufficient indication for a contralateral localization of the speech function. Attempting to localize the latter, they profited from the occasion to control musical function. The original idea was that the ability to sing is not lost, even if the left hemisphere is ablated, and that a hemiplegia paralyzing the right hemisphere does not alter linguistic function. An injection of amobarbital in the right or left carotid artery brings about a temporary inhibition of the corresponding hemisphere. In all the subjects participating in this experiment, the ability to sing was found to be strongly disturbed when the injection was given on the right, while the left side of the body was paralyzed; the faculty of speech was affected only to the degree in which the articulation of words was slower and monochord, while intonation and pronunciation and the ability to participate in a conversation did not suffer. Singing was levelled to the point that variations in frequency became purely fortuitous. Tonal memory and the sense of melodic space had completely disappeared. Rhythm, on the contrary, seemed less affected. Bogen and Gordon concluded from this that the mute hemisphere is dominant as far as tonal capacity is concerned, unless inversely tonal capacity is not a bilateral tributary function to a collaboration between the two hemispheres. This latter is what I have elsewhere called "musical relativism" (Wallin, 1981).

For several years, Bogen had been associated with works that had given similar results. The right hemisphere is dominant when it is a matter of combining different cubic configurations, copying or reproducing object-tests such as Necker's cube, the Greek cross or the swastika, for example. Jerry Levy (1971) restudied the question, prompted by a test requiring only a rudimentary motor skill but revealing a complex faculty, based on spatial transfers, of interpreting and manipulating spatial relationships with intersecting modalities. Levy had thirteen series of wooden pieces made, small enough to be held in the palm of the hand, each series having three pieces that were *almost identical*, the difference between the pieces consisting in differences in form, surface structure or marking. One piece of the puzzle was placed in the left or right hand of the patient, so as to determine if identification by touch comes from the right or left hemisphere. A screen hid both the hand and the object to be identified. A

https://doi.org/10.1177/039219218303112201 Published online by Cambridge University Press

little later, the subject was shown a card whereon were three models of the wooden "forms" to be used to make cardboard boxes that he had seen open in only two dimensions. The two hemispheres were able to see the bi-dimensional models, but only the hemisphere corresponding to the hand holding the real model could know the right answer. The subject was asked to designate the figure corresponding to the piece of wood he had in his hand, which required that he either try to refold the visual models or unfold the piece he had in his hand.

The result showed that the "mute" hemisphere—in other words, the right hemisphere—was twice as efficient in the test as the left hemisphere, considered dominant. When the left hand touched a piece of the puzzle, the response was immediate and silent. When it was the right hand, the response occurred with about 45 seconds of delay, hesitant and accompanied by stammering commentaries that showed the effort necessary to arrive at a logical solution of the problem, which proves that the two hemispheres work in a different way.

An entire series of similar tests confirm this result: as soon as a linguistic process of any kind enters into perception, the response is dominated by the left hemisphere, but when it is a matter of a visual operation, tending to establish a correspondence with a form or model, the right hemisphere is dominant. This is valid especially for forms that resist a verbal transcription. Even when words only serve as acoustical stimuli the right hemisphere prevails over the left hemisphere, on the condition, however, that the object perceived is not the semantic content of the message, which has the effect of transferring the task from the one to the other. Confirmed at the same time is my thesis that the sung text is perceived differently from the spoken text.

Simultaneous experiments in "dichotic" listening prove the specific character of the right hemisphere. Dichotic listening simply means that different stimuli are simultaneously presented to both ears, right and left, according to a method introduced in the early 50's by an English neurologist, Donald E. Broadbent. The ear opposite the dominant hemisphere prevails, taking into account the intensity of the given impulse. Doreen Kimura, at the end of a long-term research in Canada, gives us more information on the asymmetry of the functions of the brain, beginning with similar dichotic experiments. She very soon proved that normal auditive asymmetry can only be shown by a dichotic presentation of stimuli. A series of numbers rapidly pronounced from one ear to the other resulted in only a feeble and insignificant dominance in favor of the right ear (left hemisphere), and when the stimuli were given to one ear only, no difference was perceptible between the right and the left. The slight dominance of the right ear, in the case of an alternating and rapid presentation, Kimura explains by the fact that this type of presentation puts the two ears into competition, which is not the case with monaural audition. Subcortically, there is a partial parallelism between ipsilateral conduct and contralateral conduct, in which the second absorbs, and thus practically supresses, all or part of the impulses transmitted ipsilaterally. Kimura concludes from this that there must be, in addition, a similar central absorption, an idea corroborated by experiments concerning the binaural processes of the neocortex (Wallin, 1982).

In a normal dichotic test, the subject is supposed to verify all the responses he hears, in the order he himself chooses. In one particular experiment, Kimura asked her subject to react only to stimuli that struck one or the other ear—allowing to choose which one—independently of the fact that the stimuli were presented dichotically. The subject then had to report what he had heard through one or the other ear in a series of stimuli composed in this case of monosyllables (consonant and vowel) presented monaurally, that is, without simultaneous competition with the other ear. The outcome was that the subjects who had chosen to give the responses from the right ear obtained better results than those who had opted for the left ear. Kimura saw in this a rivalry in character between the two hemispheres, that is perceptual rather than sensorial.

Previously, during experiments prompted by Seashore's tests, following the application made by B. Milner on patients having undergone a lobotomy, Kimura had noted a relative superiority in the left ear (thus the right hemisphere) concerning tonic structures. Consequently, there is a line of demarcation in the form of auditive asymmetries in response to the same stimuli. These asymmetries reflect functional differences between the right and left hemispheres.

https://doi.org/10.1177/039219218303112201 Published online by Cambridge University Press

However, according to Kimura it does not suffice to say that words are verbal stimuli and that verbal stimuli are elaborated in the right hemisphere, since nothing is less clear than the composition of verbal activity. Things might be advanced by directing research toward series of syllables devoid of meaning, since by definition they cannot be conceptualized, even though formed of all the elements used by meaningful speech. Kimura's first experiment showed that these syllables are received principally in the dominant hemisphere. We are reminded of the conclusions formulated in 1961 by A.M. Libermann and his colleagues from laboratory experiments that tended to prove that a large part of of our speech is perceived with the aid of our articulatory experience. In other words, we find here the kinetic memory associated with a "sense of effort" (Wallin, 1982). Thus it is not primarily content that distinguishes the sounds of articulated speech from other sounds foreign to speech but their articulatory qualities. Later experiments using dichotic techniques have shown that asymmetry directed toward the left to identify meaningless syllables is first applied to consonant-vowel combinations, while vowels are received indifferently by both hemispheres. If the left hemisphere perceives sounds with the aid of their articulatory qualities, the latter must necessarily be specific. Later, we will constantly find this opposition between the consonant-vowel group on one hand and the isolated vowel on the other. I have already had the occasion to say on this subject that the decisive quality, in my opinion, is to be looked for in the temporal aspect of sounds as far as attack, acoustic stability and the potential relative to duration are concerned.

The surgically-split brain offers unique possibilities for the experimenter to study the specificity of the two hemispheres and the competence of each, since the limits of this competence are clearly determined and at the same time the field of interhemispheric cooperation is better defined. Some researchers, having different aims and methods, have explored the way in which we perceive the non-meaningful consonant-vowel combinations (CV) with regard to our perception of vowels (V). The results are well known, and I will only refer to some studies that throw light on our discussion on relationships between speech and music as biological phenomena. I shall begin by citing some experiments of Eran Zaidel, Roger Sperry's student and assistant, who, starting from the "split-brain" model, made essential contributions to the study of the perceptual lateralization of the brain (Zaidel, 1977).

As a point of departure, Zaidel took three basic hypotheses, which we have already encountered in Kimura: 1) the left hemisphere has gradually become specialized in linguistic processes; 2) the ipsilateral signal of the left ear to the left hemisphere is subcortically suppressed; and 3) the signal that from an ipsilateral ear (left) reaches the right hemisphere to pass on, via the corpus callosum, to the left hemisphere competes with the input arriving from the contralateral ear (right) and mixes with this input. Using these premises, Zaidel tried to find to what degree these results apply as well to CV syllables and if the ipsilateral inhibition is really subcortical. We have seen that Kimura remained attached to the idea of a cooperation of the central processes.

Six CV composed of voiced and voiceless explosive consonants combined with the vowel [*a*] were taped dichotically in the series [*pa*] - [*ta*] - [*ka*] - [*ba*] - [*da*] - [*ga*] and presented to a group of subjects who had undergone commissurotomy as well as to a control group of right-handed subjects. These CV differ by pairs because of their acoustic qualities (voiced, voiceless); by threes because of the point of occlusion in the oral cavity (labial, dental, guttural); or because of a combination of the two variations (for example, guttural sonorous, labial surd or the inverse). Earlier experiments had shown a certain predilection in favor of the right ear (left hemisphere), a predilection more noticeable in the case of the doubling of two CV's having no phonetic character in common than in the case where the pair had a common quality, thus a greater correlation for the group [*pa*] - [*ga*] than for the group [*pa*] - [*ba*], for example. Compared to the control group, the group of operated patients subjected to the same dichotic experiments showed a strong predilection for the right ear/left hemisphere, while for monaural experiments, the percentage of correct responses for this group was higher whichever ear was tested. Apparently the ipsilateral route of information is never so certain as when it does not compete with information that takes the contralateral route, according to Kimura. The results registered by the control group showed that in the normal subject the corpus callosum transferred most of the input from the left ear via the right hemisphere to the left hemisphere, and it was this route of signals that competed with the input of the right ear. Furthermore, it appeared that the impediment of the left ear was three times greater for the pairs of CV having two phonetic differences than for those having only one, for example, [*pa*] - [*ga*] against [*pa*] - [*ba*], or [*ba*] - [*ka*] against [*ba*] - [*ga*], while the percentage of correct responses for the right ear remained identical in both cases. This supported the initial hypotheses: Zaidel concluded that the suppression of the ipsilateral conduit is a function of hemispheric specialization and that nothing confirms the hypothesis of the subcortical inibition advanced by Kimura.

Other experiments have shown that the right hemisphere, in spite of its ability to decode speech, has more difficulty perceiving these differences between the pairs of consonants. The perception of these consonants is thus lateralized in the left hemisphere, not because of the linguistic processes as such but to mitigate the inadequacy of the right hemisphere that has trouble in grasping them. The reason for this may be that the right hemispere has only a poorly developed immediate memory; now, we have seen that the rapid changing of the formants in explosive consonants requires much of the immediate memory, called "echoic."

Using these results, Zaidel points out that the right hemisphere has only a weak aptitude for transmitting information on categorical *phonetic* qualities, and that it instead perceives speech by continually opposing models to each other, which would be the *acoustical* listening of *Gestalt*. The left hemisphere, on the contrary, is categorical and independent of connections, while the perceptual capacity of the right hemisphere is global, especially sensitive to connections, an opinion that Zaidel shares with other researchers.

After surgery that cuts the commissure of the two hemispheres, the left is completely dominant. Is the right hemisphere, in spite of its mutism, still able to release the sonorous image of a word? Two patients, after commissurotomy, were subjected to a test in which the image corresponding to two homonyms with different meanings was mixed with other images on a card with pictures, the test being to find the two images whose corresponding words "sounded alike": a problem solved by the left hemisphere of both patients but only by the right hemisphere of one of the patients. The following test, given the successful "right-hemisphere" patient, was to elicit the sonorous image of a spelled word, this time from its graphic and thus no longer visual representation. The test had a card with squares, of which one was filled not by a picture or an image but a word, the test being to "look for the rhyme" of the written word among the "hidden" words represented by pictures (the word "nail" and the image of a mail box and a letter—"mail"). The right hemisphere was unable to solve this problem.

These experiments show that, even with some difficulty, the right hemisphere is able to read entire words without having to phonetically spell each letter. They also show that A.M. Libermann's theory, mentioned above apropos the articulatory experiment as a basis for linguistic perception, is not the only possible model.

Taking into account that the right hemisphere, as opposed to the left, has some difficulty in analyzing phonetic categories, while the discrimination of vowels does not seem to present a problem for it, it seems plausible to me that the right hemisphere could be considered as a detector of *Gestalt* acoustics. Thus, the auditive field, in the temporal aspects, would be dispositive for emmelic as far as pitch is concerned, while it would be dispositive (in the same temporal aspects) for linguistics as far as a categorically oriented phonidentification is concerned.

Where do we go from there? For linguistics, it is the classic stages from the "Cry to the Phrase," from vocabulary to grammar, and so on. How far does the right hemisphere take us on this route?

Zaidel showed that with isolated words his patients, after commissurotomy, still had a very rich auditive vocabulary in the right hemisphere, corresponding to an intellectual age of up to sixteen years, and all things considered, barely inferior to the other hemisphere. There are probably differences between these two lexicons that it would be important to demonstrate by further experiments, since it seems that the vocabulary of the right hemisphere is clearly more connotative and associative, while the vocabulary of the left hemisphere is more precise and denotative or descriptive, "which seems to mean that the right hemisphere would have more difficulty managing a semantic ambiguity but would be richer in verbal associations." We are thinking here of the definition of poetry given by Paul Valéry: "... a hesitation between sound and meaning."

However, the musical vocabulary has scarcely any other value than utilitarian in a strict pedagogical context: the connotative or associative have no place in it, but the ideas of "connotation," "denotation" and "association" allow certain aspects that correspond to what we have been able to see apropos speech and emmelic, from the point of view of hemispheric specialization. "Denotative" presupposes analysis; "connotative" presupposes synthesis; and "associative" evokes a continuity of a kinetic nature.

The continuity of words creates difficulties for the right hemisphere, especially in phrases in which two adjectives are sufficient to modify the structure. Thus once more it is the linking of words that poses problems. Zeidel believes that the reason for this is to be sought in the absence of an immediate memory that would permit the hemisphere in question to decode non-redundant information. It is an interesting explanation, because the expression "non-redundant" naturally does not, in this case, imply a physiological or neurological redundance but a *semantic* redundance; in other words, precise, analytical, meaningful and referential information. Within a continuous associative context, the lack of immediate memory is perhaps compensated by a kinetic movement, close to what Karl Popper calls "continuous memory."

***

These observations reinforce the idea that the specificity of the two hemispheres (accentuated in the case of an accidental or provoked separation, in which interaction does not exist) does not so much reflect the difference of stimuli or modi as such (language-music, arithmetic-geometry, etc.) as two different behaviors, two different ways of thinking, as Sperry said, that are modally non-specific. The last few years have shown that the two types of behavior may be applied as well to visual-spatial or tactile stimuli as to auditive. We have just seen that linguistic stimuli are projected into the two hemispheres but in a different way,

https://doi.org/10.1177/039219218303112201 Published online by Cambridge University Press

according to Zaidel: ordinary or standard speech toward the left hemisphere and an extraordinary speech toward the right hemisphere. If archaic emmelic seems to reflect in an elementary way the specificity of the right hemisphere, an isorhythmic motet of the Middle Ages, a Bach fugue, a string quartet or *Vingt regards sur l'Enfant Jésus Christ* by Messiaen, would not be able to in such an elementary way, since they put into play a multitude of cognitive perceptual elements that define as many distinct categories. However, if it is true that melic contains many cognitive elements and that poetry is, as Valéry said, a "hesitation between sound and meaning," there is no doubt that when we say melic we are not speaking of language and that poetry is not music. The two kinds of basic behavior—the styles—introduce a decisive element. In music it is therefore emmelic that represents the ordinary or standard variable and melic the variable that also depends on the left hemisphere.

As I have just said, many of Zaidel's conclusions agree with the result of observations of the linguistic function in a normal person. The center of interest in this case is nothing less than categorical perception. Libermann's hypothesis served as a basis for Zaidel's research. It admits a fundamental distinction between two types of perception: discrimination on the one hand and identification on the other, or, if preferred, discernment and designation, or labeling. Roughly, we may say that the stimuli are disposed in a linear continuum, on the scale of physical perceptions: identification consists in distinguishing some of them in this continuous series, then in giving them a name. The ear is able to discriminate many more sounds than the subject can identify in absolute terms, but identification is a condition *sine qua non* of categorical perception. In this case, being the identification of words and syllables, that means, *primo*, an acoustic lead must be found that will progressively distinguish two consonants from each other in a continuum, in other words, a category must be defined; *secundo*, it must be possible to compare, two by two, the pairs of stimuli belonging to the different categories thus defined. Generally, everything indicates that the perception of linguistic differences *within* a category is more difficult than the perception of differences *between* the categories, which also results from Zaidel's observations.

Opposed to the categorical perceptual mode is found the continuous mode, represented in speech by vowels. If categorical perception is difficult, the perception of a series of vowels in speech is still more so and takes more time and effort for the control of articulation. As far as explosive consonants are concerned, the essential lead is found in the variations in the spectrum of attack, among others the change of formants. Now, this attack is short (30-50 ms.), while the key of vowels is found especially in the frequency of the first three formants and their duration, which is relatively long. Short vowels (40-50 ms.) are perceived more categorically than long ones allowing stable formants. My opinion concerning the importance of the time factor in the perceptual difference between consonants and vowels is confirmed by the most recent results of psycho-physical research in linguistics (Pisoni, 1979). This temporal aspect already appears in the neuronal responses to vowels at the level of the cochlear nuclei of the lower brain stem (Wallin, 1982).

However, categorical perception also exists outside of speech, for example, in music. In a 1974 study, J.E. Cutting and B.S. Rosner showed that instrumental musical sounds—thus without a direct link to language as in singing—are easily identified by listeners as "plucked" or "bowed," which roughly corresponds to the difference between consonants and vowels. Music can "reproduce" these neuronal processes. A good example is the use Berlioz made in the fourth movement of his *Symphonie Fantastique* (measures 82-90) within the same motif, of the *series* of *pizzicato-arco* for the strings and *staccato* for the woodwinds, intended to create the effect of a surprise. Because though the *pizzicato* shares the same quality with the emission of explosive consonants, which is to *require an effort* to be clearly perceptible ("the sound escapes us and we perceive the event, almost immediately, as phonetic": Studdert-Kennedy, after Cutting-Rosner), neither the *pizzicato* nor the *arco* are received as other than musical sounds.

This is expressed very well by our authors:

> "This fact [namely, that the decoding of the sound of a plucked or bowed string cannot be phonetic], coupled with the result of the first experiment, which demonstrated that rise time can cue perceptual categories in both speech and music,

https://doi.org/10.1177/039219218303112201 Published online by Cambridge University Press

suggests that certain aspects of phonetic coding may be intimately related to the coding of naturally occurring nonlinguistic sounds."

This proposition renovates my fundamental thesis, which is that the perception of music is integrated into a global auditive structural field of an unstable nature. Again I quote:

> "The fabric of speech perception and the mechanisms behind it could not have been woven wholly out of new cloth. Remnants of underlying auditory, nonlinguistic processes should and do show through. The categorical perception of musical sounds varying in rise time is apparently one of these threads."

Among the infra-linguistic processes envisaged by Cutting and Rosner in the passage quoted above is the perception of sounds that I have qualified as non-organic, to distinguish them from the perception in the higher mammals of the specific sounds of their species and our speech. The only method we have to study the "substrate" of the parallel evolution of language and music is to study the vocal communication in the higher mammals. This is, in part, the objective of modern ecological research, with ramifications in the fields of neurology, psychology, sociology and linguistics, research that has been multiplied in recent years with very interesting results, especially with regard to the specific vocalization of primates.

The sound gesture of primates consists of vocal and non-vocal signals, the latter produced with the hands and feet, lips, tongue and teeth. The animal strikes his chest with his fists or stomps, smacks his lips or tongue, grits his teeth or drums with a branch on a stone or another branch. All these gestures have their equivalent correspondents in human behavior, in connection with daily emotional situations, or as elements of a magic ritual, but since they "sound" but do not "speak" they are often grouped under the heading of ethno-musicology and included in a musicological context, which is a regrettable source of methodological confusion.

It is probable that no primate can reproduce all the linguistic sounds used by the human voice, since the anatomical formation

https://doi.org/10.1177/039219218303112201 Published online by Cambridge University Press

of their organs of phonation is somewhat different from that of man, where the roof of the tongue is more developed and more mobile, and where the larynx is found in a depressed position with regard to the tongue. The vocal repertory of the different species of primates that we have been able to examine up until now is located on a continuous line on which are found traces of all the variants of animal cries. Some of these registers contain morphologically stereotyped and acoustically distinct—or "discrete"—signals. Others are so varied that it is difficult, even impossible, to fix their limits: they are "gradual." In the one case, internal variations are slight, while the distance between the groups may be very great. In the other, the internal variations are quite evident, while the distances are blurred.

There have been attempts to look, in the behavior of such or such a species, for the lowest common denominator that would allow the prediction as to whether the repertory of each would be discrete or gradual. First the habitat has been investigated (according to whether the animal lives on the ground or in the trees). Often direct visual contact is lacking with the tree-dwelling monkeys, which would favor the development of a "discrete" register for communicating over a distance, on a sonorous background that is very rich in noises. It is the same for the small vocalizing monkeys running the risk of being tracked down. It has been shown, however, that this approach is not sufficient if it is not completed by studies on the social relationships between individuals within groups.

If vocal communication occurs at a short distance, or even at sight, the ambiguity of the cries may be greater, since the gesture and the visual contact may clarify them. The existence of discrete models within a variable repertory may be attributed to a change in situation, entailing, for example, the need to give a cry of alarm, which implies a distinction in the register having the value of a signal.

With monkeys living on the ground, like the macaques, both hypotheses are valid. In this case, we are dealing with groups that are generally numerous, eminently sociable, already disposing of different types of signals, not just acoustic. All this is favorable to a "gradual" vocalization. Apparently, there is a selective predilection in favor of gradual signals that could be due to the fact

https://doi.org/10.1177/039219218303112201 Published online by Cambridge University Press

that the animals that develop this repertory by this can refine their information, which in turn allows them to transmit a more subtle and more complex degree of social contact.

J.G. Gautier (1974) classifies the modifications of the gradual repertory under three headings: *transitional, ontogenic* and *internal*. The first may be illustrated by the aggressive cries of the macaque, that go from muttering to howling. The second corresponds to the phases in the evolution of the young during their growth, from maturation to apprenticeship. The third covers the entire scope of the variations of a determined cry of an individual or within the group.

S. Green (Zoloth, 1979), who has especially studied the Japanese species (*Macasa fuscata*), analyzes what he calls their "cooing" and observes a certain correlation between the morphological characteristics of this cry and certain types of behavior. The animal would have thus developed a system of vocal signals among which he would choose such or such variations as a precise communication. Since each type of "coo" also permits a certain degree of internal variation, the new member is put through a severe test. There is thus good reason to assume that evolved primates also have access to a "perceptual constance," in other words, to the fact that an expressive sound made by the same subject can change from one situation to another but nevertheless remains perceived as "similar" or "identical."

M. Beecher *et al.* (1979) also speculated as to whether these curious Japanese animals were not also endowed with a selective attention and had no preference for one or the other ear, in other words, if their perception of signals was lateralized or not.

To answer the first question, they began by distinguishing two kinds of "coo" that Green identified as "smooth-early-high-coo" (SE) and "smooth-late-high-coo" (SL). The first are the cries of contact of isolated animals, separated from the flock, alone or in groups, or the cries of young animals, temporarily separated from their parents or their playmates within the group. These cries often receive an antiphonal response. The SL cries are also cries of contact but of a different nature, denoting either a sexual invitation or a deference of subordinate to superior. While the animal emitting an SE shows a mentally-balanced state, the cries of the SL type rather express a more emotional reaction. Each of the two types of cries contains a shrill phase of maximal frequency ("peak") that appears for the one in the first two-thirds and for the other in the final third of the emission. During two different experiments, it was a matter of having the subjects distinguish the pitch of the sound, excluding the place of the peak in the curve, or of recognizing the location of the peak in excluding the pitch. The results showed that the Japanese macaque can learn to discern the two on the basis of a natural classification proper to his communication system, but like a small child he does it with the aid of a selective attention that, according to Patricia K. Kuhl (1977), is more sensitive to the category of vowels than to the melodic contour of pitch.

As far as the identification of the peak is concerned, all the monkeys showed a marked predilection for the right ear (left hemisphere), while the left ear (right hemisphere) seemed distinctly more able to recognize the pitch of the sounds.

We have thus been able to see that the phenomena of perception that were believed to be limited to the linguistic sphere are much more general in scope. Categorical perception and perceptual constance, as well as the phenomena of auditive laterality in certain animals, seem to be close to the analogous phenomena observed in the new-born infant. There is also a certain morphological-anatomical similarity between the human infant and certain monkeys. The register of the supra-laryngeal voice of the infant, according to Ph. Lieberman, is not noticeably different from that of the chimpanzee, or, as he says, "the effects of the pharyngeal cavity on the quality of speech would be analogous in the newborn child and the chimpanzee." Observations of this kind are naturally instructive for us, and I would like, in conclusion, to mention some of the questions raised on this point by specialists.

Patricia K. Kuhl (1979) has no doubt that

> "certain predispositions for the perception of speech-sound categories exist in nonhuman mammals, and lead naturally to the suggestion that the repertoire of speech sounds was originally selected, in the evolutionary scheme of things, precisely because they were so ideally suited to the auditory system."

This may be summarized as follows: the aptitude to perceive a certain acoustical continuum categorically is by nature psycho-physical; certain acoustical phenomena engender natural categories in which acoustical changes are perceived only beyond a certain threshold. The sounds that may thus be opposed are useful for communication, in the sense that the system tolerates a rather wide margin of uncertainty without endangering the perception of the passage to the perception of a higher category.

We oppose an objection to this thesis, one that arises from Libermann's observation on "articulatory experience" (see above). K.N. Stevens (1972) demonstrated that some combinations of the frequencies of formants resist very well even a considerable articulatory activity. Speech can thus "discern" certain zones of acoustic attributes, offering such a resistance to the "brutality" of articulatory treatment, and from that decipher the code of a linguisic communication. The choice of these privileged regions would be pinpointed both by the stability of the acoustical dimension, for example, between the first two formants, as well as by their psycho-acoustical attributes as they are presented once the auditive respective the articulatory mechanisms have accomplished their task. I think we may summarize the problem in this way: the relationship of articulation to what Kuhl and Stevens call "acoustic" and the relationship of acoustics to information have evolved in such a way as to permit a certain tolerance with regard to "faults." If I have understood them correctly, we are not far from the descriptive analysis of evolution that I have outlined as a hypothesis, based on the idea of an auditive system considered as an instable open system.

III. Expanding the Hypothesis

The model of the expressive sound world of man that I submitted at the beginning of this article may now be completed. I will begin with the preceding hypothesis to make the point of the discussion. To conclude, I will indicate what possible rapport I see between the dispositive system conforming to the model and some new ideas concerning unstable open systems that through a multi-oscillatory activity may be transformed, giving rise to other structures with a higher and more specific degree of orientation.

As we have seen, it has been proved that non-human primates have a categorical capacity for perception and a perceptional constance; that they show within a complex social context a preference for non-discrete, gradual signals; and that their auditive functions are accompanied by a selective attention that seems tied to a functional lateralization.

In human speech function, the attack and the ability to produce explosive consonants play a predominant role. Ph. Lieberman has shown that the hominid monkeys have all the anatomical conditions for this. The throat of these animals even lends itself to what we call *falsetto*, the fundamental tone of which is higher relative to the register of a normal voice, and its energy is relatively feeble in the upper harmonics. Generally, however, spectrograms of monkeys reveal more energy in these upper harmonics than in what is supposed to represent normal speech in man.

These anatomical qualities can end in a linguistic phonation only if they are accompanied by an adequate muscular control. Such a control is only possible if it supposes, as preliminary condition, a modelization of corresponding neuronal regions that do not exist to a high enough degree in non-human primates.

Like Patricia K. Kuhl, Ph. Lieberman marks the resemblance between the perceptual models of the primates and the perceptual model that guides linguistic activity in the human infant. The region of the brain of these monkeys toward which converge the data of sensorial information in all its forms is homologous with the region of the human brain that, in the left hemisphere, presents a manifest hypertrophy. This hypertrophy exists in the infant, but he does not use it yet and seems to discriminate bilaterally.

According to Kuhl, the new-born child pays more attention to the category of vowels than to the line of contours of pitch. We may deduce, therefore, that his capacity to identify low-frequency and acoustically stable sounds, such as vowels, corresponds to a perceptive capacity that is much more discriminatory than identificatory. Such a discriminative power represents the neural analysis of temporal qualities, in which I have believed to recognize the characteristic of emmelic perception. (The sequences of neuronal impulses usually adopt a phasic oscillatory coordination). Some researchers, relying upon clinical observations of phenomena of recession that accompany aphasias, have presented the hypothesis that music represents a period that is phylogenetically anterior to that of speech.

G.H. Monrad-Krohn, in a series of articles extending from 1937 to 1963, studied the pathological deformations of linguistic prosody, and in the last of these articles he especially retains what he calls "grunts," that is, "those inarticulated noises in which modifications in tone, accent and rhythm are the only bearers of the intellectual and/or emotional content of the message," in other words, those exclamations that Sperry and others localize in the right hemisphere and which represent for them a conscious experience, having a corresponding social significance.

Monrad-Krohn gives some examples of analogous sound gestures that he takes from his clinical observations. One cannot help being struck by their richness in vowels and voiced consonants at the expense of the hard attacks represented by explosive consonants. This is also true for the repertory catalogued by Sperry. In this case, we again encounter the preference of the right hemisphere for an acoustical stability. Monrad-Krohn shows, with the aid of spectrograms, that emotionally "positive" grunts present a wider set of formants than do gestures having only an interrogative meaning. He interprets this, taking changes in frequency into account, as a proof of the semantic value of sound gestures. The distribution of the "peaks" would thus have a certain relationship with behavior, as we have seen in the variations of "coo."

The notion of "breath-group" enters into Ph. Lieberman's research on the origin of language:

> "The phonetic feature that speakers make use of to segment the train of words is the *breath-group*, which enables the listener to group words into meaningful sentences. It probably is one of the most central, basic aspects of language, and it, or some equivalent feature, must have been present in the earliest forms of hominid language... In the production of normal speech the acoustic cues that characterize the normal breath-group are a consequence of minimal deviation from the respiratory activity that is necessary to sustain life."

In my initial remarks on the tonotopic system, I called attention to the importance of the feedback that covers at the same time the auditive system and others, among them neural mechanisms commanding the respiratory cycle. In an earlier study, quoted above (Wallin, 1982), I showed how, in this feedback system, the interval of time varies from the beginning of the change in electric activity in the internal muscles of the larynx, to the commencement of phonation, in rapport with different linguistic sounds. The interval is considerably shorter for surd consonants than it is for vowels or voiced consonants. This interval is partially determined by impulses coming from the respiratory mechanisms, governed by motor units located in the *ambiguus* nucleus of the lower brain stem, which in turn receives its orders through conduits coming from the frontal parts of the neocortex.

There is a certain relationship of coordination between the tonic activity of the laryngeal muscles (in the absence of phonation) and the spontaneous respiratory cycle. The latter provokes impulses in the muscles that, so to speak, are superimposed on the tonic activity proper to the laryngeal muscles. These impulses are phasically coupled with the oscillatory curve of respiration.

As Ph. Lieberman has shown, respiratory rhythm changes with linguistic activity. Spontaneous expiration lasts around two seconds but may vary during linguistic production from three-tenths of a second to forty seconds, which supposes a radical change in the harmonic relationship between the respiratory cycle and the muscular activity of the larynx. Since the time of preparation for the beginning of phonation is different for vowels or voiced consonants, on the one hand, and surd consonants on the other, the lag between the phases must be less for the voiced pair, and especially for emmelic music. Hence the optimalization of the sensory-motor activity that is, apparently, expressed by integral multiple oscillations, would be greater for emmelic than for speech. In support of this hypothesis is the fact that music not only presents a larger acoustical stability among its constituent elements but that it starts from these elements to develop, through periodic additions or associations, constructions resulting in composed structures or *Gestalts*, that become more and more important.

A.M. Libermann (1962, see above) sustained that we perceive

https://doi.org/10.1177/039219218303112201 Published online by Cambridge University Press

speech partly on the basis of a kinetic memory or an articulatory experience. It is well known that musical listening is accompanied by a change in rhythm in cardiac pulsations and in muscular movement, for example in the legs. It is thus probable that kinetic experience ("the sense of effort") plays an important role in musical performance as well as in listening to music. If the optimalization of effort is greater in musical activity than in linguistic activity, the combination of all these factors perhaps explains why certain aphasic subjects are able to sing words that they cannot emit by speaking. The oscillatory stability of the system of musical feedback is greater than it is in linguistic production, where it is constantly changing.

The overall image may be interpreted so that the mode of auditive discrimination that favors acoustically stable stimuli has priority. The bilateral functional image that in monkeys seems to be multi-modal and is likewise so in the new-born child would be the phylogenetic point of departure of a distinction between *Homo sapiens* and hominids. Hominids were already a highly specialized species with regard to other higher mammals. *Homo sapiens* added certain innovations of capital importance for the functioning of his central nervous system. One of them is the coordinating role of the frontal brain; another, the lateralization permitting linguistic functions.

In spite of this lateralization, the human brain retains an integral unity, especially in the phylogenetically older emotive and attentive aspects of its activity, so that the new specialization of the left hemisphere also profits the right hemisphere. The primitive faculty of responding to acoustically stable stimuli develops autonomously because of their strong ties with emotive functions, the heritage of the sound gesture that precedes speech. The new topographical division that I have just mentioned ends in a division of tasks between the substrate and the sonorous gesture, one analyzing the phonetic aspects, the other recreating an acoustic totality. The two functions are autonomous, in spite of interhemispheric coordination, but their autonomy operates differently in each part, in that the emmelic line—under the ever-increasing influence of the left hemisphere and the faculty it has to conceptualize—is more and more specialized in a perception that is non-semantic and *per se*. From the coordination with the regular periodicity that is established in the exogene stimuli found in the models of ammelic impulses comes a stability of internal or endogene temporal structures that are progressively affirmed: music as a pace-maker. As the object of emmelic perception is detached from the variations of synergic relationships and is oriented more and more toward an autonomous synchronization of the different oscillatory activities, internal and external, emmelic becomes melic. Not all individuals cross this threshold, whose crossing demands a concurrence of different qualities. The same is true of societies and civilizations: throughout history, music, side by side with its autonomous functions, is constantly associated with a synergic context like rites, social ceremonies, dance, theatre, etc. It is possible that the first emmelic was born of the union of music with dance.

* * *

Thus, the expressive sonorous world in man is divided into three functional zones. We no longer say *in principio erat verbum,* but in the beginning functional reflex signs existed (Kaila) that were especially manifested by a sound gesture having on the one hand an autonomous existence, with its easily-discernible distinctive characteristics, and on the other the faculty of being linked to phylogenetically more recent regions of language and emmelic. Between the regions is a complex relational structure, rigorously coordinated. According to recent studies, some species of higher mammals show a disposition to cross the boundary of the younger regions. Man is comfortable in the three regions: signs of atavistic reaction have stronger emotive qualities but weaker designative and propositional qualities.

Speech takes part in all of it with a practically unlimited possibility of variations but always with a direct or indirect ponderance in favor of the designative. This is also true for poetry. In emmelic or melic, values of contrast, similarity, permanence or transience have meaning only by virtue of their combinatory power and not because of their rapport with a real object or whatever event: perception refers here to a *sui generis* process.

Neither is the primary function of music to establish a meaning-

https://doi.org/10.1177/039219218303112201 Published online by Cambridge University Press

ful rapport with an external object, nor to be in correlation with neural phenomena representing a stimulus corresponding to this object. This is one reason, among others, that led me to question the communicative value of music. We could ask ourselves instead if it is not rather an indication of the state of mind of the creator, as Guido d'Arezzo expressed (after William Blake) when speaking of sensations as so many windows.

Secondly, music enters into a complex semiotic situation or into what I call "a field of expressive synergy," in which acoustical parameters such as intensity or frequency and the graduated variations of these parameters in relation to each other end by creating experiences (*near* or *far*, *right* or *left*, *large* or *small*) that under certain conditions approach analogous experiences or reenforce the analogy between impressions received via other modal conducts, pictures, or dance, for example.

In the case of music, the underlying "economic" principle probably goes back to a subtle sensitivity for the integral multiples and thus the internal morphological affinities characterizing the temporal aspects of the auditive processes of the brain, starting in the fibers of the auditory nerve, continuing in the nuclei of the brain stem and ending in the neocortex. The same principle applies to the coordination of the oscillatory auditive system with the respiratory cycle and the muscular activity of the larynx (regardless of whether or not there is phonation).

The teleology of speech is based on our faculty to identify a phonetical flow through a categorical perception. Musical perception, on the other hand, is based on a continuous and appositional acoustic discrimination. If this last, freed from the necessity for a rapid semantic interpretation of external stimuli that endlessly change, is in fact slower, more periodic and regular, we must probably see in that the consequence of the fact that it uses specific temporal approaches at the level of cellular activity, and particular transfer processes.

However, the difference is relative, since speech and music occupy neighboring zones that intersect. Speech offers variants (the linguistic variant of Zaidel, for example) that show aspects imprinted with continuity (sonic languages and poetry, for example), while in the emmelic zone we see a tendency toward a categorical and cognitatively more charged aspect: melic. This relativity reflects the constant collaboration and interdependence of the two hemispheres, that ends in an integration of the two modes of thought.

* * *

It goes without saying that this dispositive system has nothing stable or closed about it. There are continual exchanges of energy and information between the central auditive processes and other internal systems, across the different structures of feedback, that are never in perfect equilibrium nor closed. Energy and information are equally exchanged between motor functions centrally directed and the surrounding world, which lends credit to the notion of a "socio-cultural dispositive field." Over the course of time, the internal oscillatory activity fits itself in among "pacemaker" type of stimuli, such as—most distinctly—music.

The functional adaptation of the auditive field to specific dispositives of a higher degree of complexity is made, as we have seen, by leaps and not linearly. Inside this system, there are distances of an extreme sensitivity between stable states, with regard to the neural modes of excitation. The instability of the auditive field, according to this hypothesis, is one of the inherent or immanent qualities of the system. It follows that the emmelic fluctuations of music are not epiphenomenal "deviations," with regard to a system of rigid rules, but qualities, that through the oscillatory mechanisms of feedback can be reintegrated at a higher and more specific level of activity. This is precisely what is illustrated by emmelic-melic, in all its innumerable synchronic and diachronic variations and what resolves the tonotopic paradox.

What I have said here about the auditive-dispositive field is, in sum, rather close to the descriptive analysis of structures called "dissipative," unstable, or open. In my opinion, it is within the framework of the theory of dissipative structures that research should be directed in the future to form its "matrix."

> "It is through the instability of dissipative structures in open systems that certain disturbances of the homogeneity of the system, within the framework of tolerable fluctuations, may ... be amplified, modeled and presented macroscopically as a stationarily developed structure of the system (a spatio-

https://doi.org/10.1177/039219218303112201 Published online by Cambridge University Press

periodic dissipative structure) or like an oscillation (most often in the form of a temporal-spatial-periodic dissipative structure)" (Sinz, 1980).

There is no better way to describe the relationship in the auditive field between, on the one hand, general sound reception and on the other, the specific perception of music.

Nils L. Wallin
(*Royal Swedish Academy of Music*)

BIBLIOGRAPHY

BEECHER M.D., PETERSEN M.R., ZOLOTH S.R., MODDY D.B., STEBBINS T.C., "Perception of Conspecific Vocalization by Japanese Macaques. Evidence for Selective Attention and Neural Lateralization," *Brain, Behavior and Evolution*, 1979, vol. 16, nos. 5-6, p. 443.

BOGEN J.E., GORDON H.W., "Musical Tests for Functional Lateralization with Intracarotid Amobarbital," *Nature*, 1971, vol. 230, p. 524.

CRICK F.H.C., "Thinking about the Brain," *Sci. American*, 1979, vol. 241, no 3, p. 131.

CUTTING J.E., ROSNER B.J., "Categories and Boundaries in Speech and Music," *Percept. Psychophys.*, 1974, vol. 16, p. 564.

DOBZHANSKY T., *The Biology of Ultimate Concern*, New York, New American Library, 1967.

VON ECONOMO C., "Über Progressive Cerebration," *Kl. Wochenschrift*, Vienna, 1951, no. 44, p. 597.

FLANAGAN J.L., "A Difference Limen for Vowel Formant Frequency," *J. Acoust. Soc. Amer.*, 1955, no. 27, p. 613.

GAUTIER J.P., "Field and Laboratory Studies of the Vocalizations of Talopoin Monkeys, Miopithecus talopoin," *Behavior*, 1974, no. 41, p. 209.

GESCHWIND N., "The Anatomical Basis of Hemispheric Differentiation," *Experimental Studies of Hemisphere Function in the Human Brain*, ed. Dimond-Baumont, 1973.

GESCHWIND N., "Anatomical Asymmetry as the Basis for Cerebral Dominance," *Fed. Proc.*, 1978, 37 (9), p. 2263.

IMBERTY M., (a) "Acculturation Tonale and structuration perceptive du temps musical chez l'enfant"; (b) "Peut-on parler de compétence musicale à propos de la structuration du temps musical chez l'enfant?," *Basic Musical Functions and Musical Ability*, papers given at a seminar arranged by the Royal Swedish Academy of Music, ed. Hans Astrand, 1981, publications issued by the Royal Swedish Academy of Music, no. 32, p. 131.

KIMURA D., "Functional Asymmetry of the Brain in Dichotic Listening," *Cortex*, 1967, 3, p. 163.

KUHL P.K., "Speech Perception in Early Infancy: Perceptual Constancy for the Vowel Categories /a/ and /o/," *J. Accoust. Soc. Amer.*, 1977, 61, suppl. 1, S39 (A), p. 374.

KUHL P.K., "Models and Mechanisms in Speech Perception. Species Comparisons Provide Further Contributions," *Brain, Behavior and Evolution*, 1979, vol. 16, no. 5-6.

LEVY J., NEBES R.D., SPERRY R.W., "Expressive Language in the Surgically Separated Minor Hemisphere," *Cortex*, 1971, vol. 7, p. 48.

LIBERMANN A.M., COOPER F.S., HARRIS K.S., MACNEILAGE P., "A Motor Theory of Speech Perception," *Proc. Speech Communication Seminar*, Stockholm, 1962.

LIEBERMAN PH., *On the Origin of Language. An Introduction to the Evolution of Human Speech*, London-New York, MacMillan, 1975.

MONRAD-KROHN G.H., "The Third Element of Speech; Prosody and its Disorders," *Problems of Dynamic Neurol.*, Jerusalem, 1963, p. 101.

NICHOLIS G., PRIGOGINE J., "Fluctuation in Non-Equilibrium Systems," *Proc. Nat. Acad. Sci.*, 1971, 68, p. 2102.

NKETIA J.H.K., *The Music of Africa*, New York, W.W. Norton and Co., 1974.

PISONI D.B., "On the Perception of Speech Sounds as Biologically Significant Signals," *Brain, Behavior and Evolution*, 1979, vol. 16, no. 5-6, p. 330.

PRIGOGINE J., "Order Through Fluctuations: Self-Organizations and Social System," *Evolution and Consciousness*, ed. E. Jantsch, C.H. Waddington, London-Amsterdam, Addison-Wesley Publ. Co., 1976, p. 93.

SINZ R., *Chronopsychophysiologie, Chronobiologie und Chronomedizin*, Berlin, Akademie-Verlag, 1980.

SPERRY R.W., "Forebrain Commissurotomy and Conscious Awareness," *J. Med. and Phil.*, 1977, vol. 2, no. 2, p. 101.

STEVENS K.N., "The Quantal Nature of Speech: Evidence from Articulatory-Acoustic Data," *Human Communication: A Unified View*, ed. David Denes Jr., New York, McGraw-Hill, 1972.

USTVEDT H.J., "Über die Untersuchung der musikalischen Funktionen bei Patienten mit Gehirnleiden, besonders bei Patienten mit Aphasie," *Acta Med. Scan.*, Suppl. LXXXVI, Stockholm-Helsinki.

WALLIN N.L., "Some Notes on the Neurobiological and Genetic Aspects and Musical Ability," papers given at a seminar arranged by the Royal Swedish Academy of Music, ed. Hans Astrand, publications issued by the Royal Swedish Academy of Music, 1981, 32, p. 179.

WALLIN N.L., "Den musikaliska hjärnan. En kritisk essä om musik och perception i biologisk belysning," with an English summary, KUNGL, *Musikaliska Akademiens Skriftserie*, 3," Skrifter frän musikvetenskapliga institutionen, Göteborg, 6, Stockholm, 1982.

ZAIDEL E., "Language, Dichotic Listening and the Disconnected Hemispheres," *Brain Inform. Serv. Conference Report*, 42, Los Angeles, 1976, p. 103.

ZAIDEL E., "Lexical Organization in the Right Hemisphere," *Cerebral Correlates of Conscious Experience*, ed. P.A. Buser, A. Rougel-Buser, 1977, Inserm. Symposium, 6, 1978.

ZOLOTH S., GREEN S., "Monkey Vocalization and Human Speech: Parallels in Perception?," *Brain, Behavior and Evolution*, 1979, vol. 16, no. 5-6, p. 430.