BRITISH ACOUSTICAL SOCIETY: SPRING MEETING at Chelsea College
London S.W.3 on Wednesday 25th April to Friday 27th April 1973

SPEECH AND HEARING: SESSION A: SPEECH PRODUCTION AND PERCEPTION

Paper No. 73SHA2

MODELS IN HEARING

M. R. Schroeder
Drittes Physikalisches Institut
Universität Göttingen

Invited Paper

In man's attempts to understand the functioning of his sense of hearing, models have played an important role. As early as the first century B.C., Lucretius postulated a model involving little grains of sand in the inner ear responding to different tones. Tartini, the noted 18th-century Italian violinist, described his "terzi tuoni" - tones that the ear itself manufactures and that are still not adequately understood. In more recent times, Helmholtz added important concepts. His "place theory" of hearing, i.e., the unique relationship between physical frequency in an auditory stimulus and the place along the basilar membrane where the frequency is "detected", is a cornerstone of our thinking about the ear.

Helmholtz's model had to give way, in time, to von Békésy's traveling wave model - not so much in a complete upheaval of prevailing concepts but, rather as a refinement of and an addition of substance to earlier ideas.

Both Helmholtz's and Békésy's models are physical models in the sense that a specific physical process is assumed to be responsible for the observed behavior. The physical process, of course, can be "summarized" by mathematical equations - in the case of Békésy's traveling waves, by the equations governing the propagation of "disturbances" on nonuniform transmission lines. The same equations, naturally, apply to other physical systems equally well. This has permitted their "visualization" in demonstration experiments, a fact exploited by Békésy in his demonstrations making use of a long row of coupled pendula simulating his traveling waves in slow motion, so that the "membrane" (pendulum) oscillations could be followed in all detail and at leisure.

The fact that such physical models can be described adequately by mathematical equations leads to what might perhaps be called the ultimate in model visualization: the simulation of the pertinent equations on a digital computer for various acoustic inputs and the microfilm "printing" of the corresponding response motions of the entire basilar membrane on successive motion-picture frames - one frame representing perhaps as little as one tenth of one millisecond so that the finished film can be viewed at a slowdown ratio of 400:1 (Lummis, 1968).

The solid foundations on which these models are built owes much to the comprehensive work of Lord Rayleigh who, characteristically, appended a critical chapter on "Facts and Theories of Audition" to his Theory of Sound.

In contrast to such physical models, where the mathematical description is merely a by-product, though often a highly welcome one, hearing theory has also benefited from models which are, from their very conception, formulated in mathematical language. A good case in point is Licklider's (1956) "triplex" theory of hearing: after the place-frequency analysis on the basilar membranes, the outputs of the two membranes corresponding to like frequencies are cross-correlated and the outputs of the (many) cross-correlators are in turn subjected to auto-correlation analysis. The cross-correlators are assumed to provide the directional information in auditory space and the auto-correlators are held responsible for such well-known - but somewhat inconvenient - phenomena as Schouten's "residue" pitch and other periodicity sensations for which no place could be found on the basilar membrane (such as the "repetition pitch" - observed even for random signals with a single echo).

While Licklider's theory built on well established terrain as far as place-frequency analysis is concerned, it broke new ground in the interaction of the postulated correlation processes. Of course, nobody, to this date, has seen the necessary multipliers and delay lines in the head. In this sense, Licklider's theory was and is a purely mathematical one. Yet there are many ways to realize correlation processes and it is not difficult to construct, from the neurophysiologically observed binaural inhibition, a mechanism at least resembling cross-correlation. Likewise, the auto-correlation function of a signal is closely related to the <u>interval histogram</u> of nerve spikes elicited by the same signal (Schroeder, 1970). And the auditory system must be sensitive to inter-spike intervals, because - apart from average firing rate and place of origin - there is little else that higher acoustic centers in the brain receive from the auditory periphery.

Thus, while large portions of Licklider's model are simply untrue if taken literally, his triplex theory continues to be an important guiding principle - to be filled with substance, modified and expanded as our knowledge of neural processes evolves.

Occasionally, a "new" model is not really new but simply the addition of more specific assumptions to an existing one. An example is the author's (1972) "Integrable Model for the Basilar Membrane". For a basilar membrane having, in addition to the well-known logarithmic relationship between place and characteristic frequency, a phase velocity which is likewise an exponential function of place (and a loss-factor (1/Q) independent of place), the equations of Zwislocki (1950) and Peterson and Bogert (1950) can be integrated in closed form yielding surprisingly simple relations between such seemingly unrelated observables as the low-frequency phase-slope (in Radians per Hz) and the high-frequency amplitude-slope (in dB/octave). (The latter equals, approximately, the former times 6 times the characteristic radian-frequency of the place of observation.) The agreement of the model with the Mössbauer-technique measurements of Johnstone, Taylor and Boyle (1970) and Rhode (1971) is good enough to make it a useful and convenient tool in further studies of the basilar membrane especially its nonlinearities (Hall, 1973).

A Mathematical Model of the Hair Cell Function

Spike data recorded from primary acoustic fibers (see, for example, Rose *et al.*, 1967), show some of the more puzzling properties encountered in audition. Thus, "period histograms" (i.e., the distribution of nerve firing probabilities during one period of the acoustic stimulus) for signals above a level of about 50 dB (*re* 0.0002 dyn/cm^2) are remarkably insensitive to stimulus amplitude - as if a strongly amplitude-limiting device had intervened somewhere between acoustic stimulus and neural spikes. Yet, for one polarity ("positive", say) of the acoustic signal, the period histogram is a close replica of the stimulating waveform - or the waveform thought to stimulate the hair cell (Brugge *et al.*, 1969). (For "negative" polarities the period histogram is suppressed below the spontaneous firing level.) In the language of electrical engineering: there must be a half-wave rectifier and a waveform-preserving automatic gain control at (or near) the point of transduction from mechanical motion to neural activity. Needless to say, neither has been found. Existing models, like Weiss' (1966) neuron model, though useful in some respects, show grossly different behavior.

While pondering these problems, it recently occurred to the author that a very simple mathematical process can account for most (if not all) of the existing primary acoustic spike data. Furthermore, *it seems that this mathematical model can be given physico-chemical substance*. The model is characterized by the following "rules":

1. Quanta of a (chemical) agent are generated in the hair cell (or the hair itself) at a fixed average rate.
2. Some quanta disappear at a rate proportional to their number without noticeable effects in the observed nerve fiber.
3. The remaining quanta disappear *and* cause firing in an attached fiber at a rate proportional to their number and a membrane "permeability function" which is linearly related to the mechanical pressure on the hair over a given range of "positive" pressure. For "negative" pressures,

the membrane permeability is reduced to values near zero.

The observed rectifier action then resides in the asymmetric dependence of membrane permeability on pressure for which there is ample physico-chemical evidence, at least for artificial lipid membranes (H. Träuble, 1971, H. Träuble and D.H. Haynes, 1971).

The observed "gain control" property of the transduction process is a consequence of the assumed proportionality of firing probability with the number of available quanta: with increasing amplitude of the mechanical stimulation, the average membrane permeability increases proportionally while concurrently, because of the extra "drain" on the quanta, their expected number decreases. For large mechanical amplitudes, these two effects (increase in membrane permeability and decrease in available quanta) cancel each other out leading to the observed independence of nerve firing probability on stimulus <u>amplitude</u> while preserving stimulus <u>waveform</u> (for "positive" signals).

Other properties of this model for the mechanical-to-neural transduction process are being studied in collaboration with J. L. Hall of Bell Laboratories, who also contributed to the formulation of the model. Logan and Shepp (1973) and Mallows (1973) analyzed several mathematical aspects of the discrete Markov process underlying the model. Logan and Shepp first showed that the intervals between successive neuron firings are geometrically ("exponentially") distributed - in agreement with neurophysiological data beyond the influence of refractory effects.

With the inclusion of refractory properties and other refinements, mathematical analysis is being replaced by computer simulation. Among the conditions tested to date are period histograms (for sinewaves of widely different frequencies and complex signals); interval histograms (for periodic signals and noise) and post-stimulatory histograms (for tone and noise bursts). The agreement with corresponding neurophysiological data is remarkable.

Future Outlook

As we probe deeper and deeper into the mysteries of human perception, auditory and otherwise, our reliance on models, particularly of the mathematical kind, is bound to increase markedly. True, in due course, these models may gain neurophysiological and physico-chemical substance but, as we progress to ever higher centers of awareness, there will probably remain a "mathematical residue" not amenable to further interpretation. The situation may evolve in a manner similar to Quantum Field Theory in physics where questions concerning the "meaning" of an abstract formalism are meaningless and where even the formulation of self-consistent theories and models whose predictions are in agreement with observed facts are anything but trivial.

References

von BÉKÉSY, G. (1960). Experiments in Hearing, E. G. Wever, Ed. (McGraw-Hill, New York).

BRUGGE, J. F., ANDERSON, D. J., HIND, J. E., and ROSE, J. E. (1969). "Time Structures of Discharges in Single Auditory Nerve Fibers of the Squirrel Monkey in Response to Complex Period Sounds," J. Neurophysiol. 32, 386-401.

HALL, J. (1973). "Simulation of a Nonlinear Model of the Basilar Membrane on a Digital Computer," J. Acoust. Soc. Amer. 53, 324(A).

JOHNSTONE, B. M., TAYLOR, K. T., and BOYLE, A. J. (1970). "Mechanics of the Guinea Pig Cochlea," J. Acoust. Soc. Amer. 47, 504-509.

LICKLIDER, J.C.R. (1956). "Auditory Frequency Analysis," in Colin Cherry (Ed.) Information Theory (Academic Press, New York). (See also Licklider's chapter, "Three Auditory Theories," in S. Koch (Ed.) Psychology: A Study of a Science (McGraw Hill Book Co., New York 1959).

LOGAN, B. F., and SHEPP, L. A. (1973). "Birth and Death Model of Neuron Firing," J. Appl. Probability (to be published).

LUMMIS, R. C. (1968). "The Secret Code of Hearing," Bell Labs Record, 46, No. 9, 261-266.

MALLOWS, C. L. (1973). (Private communication).

PETERSON, L. C., and BOGERT, B. P. (1950). "A Dynamical Theory of the Cochlea," J. Acoust. Soc. Amer. 22, 369-381.

RHODE, W. S. (1971). "Observations of the Vibrations of the Basilar Membrane in Squirrel Monkeys using the Mössbauer Technique," J. Acoust. Soc. Amer. 49, 1218-1231.

ROSE, J. E., BRUGGE, J. F., ANDERSON, D. J., and HIND, J. E. (1967). "Phase-locked Response to Low-Frequency Tones in Single Auditory Nerve Fibers of the Squirrel Monkey," J. Neurophysiol. 30, 769-793.

SCHROEDER, M. R. (1970). (Unpublished notes)

SCHROEDER, M. R. (1972). "An Integrable Model for the Basilar Membrane," Symposium on Hearing Theory, Eindhoven To be published in J. Acoust. Soc. Amer.

TRÄUBLE, H. (1971). "The Movement of Molecules Across Lipid Membranes: A Molecular Theory," J. Membrane Biol. 4, 193-208.

TRÄUBLE, H. and HAYNES, D. H. (1971). "The Volume Change in Lipid Bilayer Lamellae at the Crystalline-liquid Crystalline Phase Transition," Chem. Phys. Lipids 7, 324-335.

WEISS, Th. F. (1966). "A Model of the Peripheral Auditory System," Kybernetik 3, 153-175.

ZWISLOCKI, J. (1950). "Theory of the Acoustical Action of the Cochlea," J. Acoust. Soc. Amer. 22, 778-784. Based on J. Zwislocki, Acta Otolaryng. Suppl. 72 (1948).