

XIV. SPEECH COMMUNICATION*

Prof. M. Halle
Prof. K. N. Stevens
Prof. J. B. Dennis
Dr. A. S. House
Dr. G. Rosen

Dr. T. T. Sandel
Dr. W. S-Y. Wang
Jane B. Arnold
P. T. Brady
J. R. Carbonell

M. F. Dichter
O. Fujimura†
H. Fujisaki
M. H. L. Hecker
J. M. Heinz

A. BINAURAL HEARING: A TENTATIVE MODEL AND SOME RELATED EXPERIMENTS

In this study we have proposed a model that attempts to explain certain aspects of binaural hearing, such as sound localization and the binaural threshold shift (BTS). The BTS is defined as the change in the masked threshold of signals through binaural listening relative to monaural listening. The theoretical performance of the model is in qualitative agreement with the results of the classical, as well as the most recent psychoacoustic, experiments; furthermore, the model attempts to take into account the present status of knowledge about the ear and the nervous system at its different levels.

The model postulates an active process. It is assumed not only that the signals arriving at the two ears are processed and combined, but also that there is an adjusting mechanism that tends to optimize the result. This mechanism consists of adjustable transfer functions and switching devices controlled by means of feedback loops. In the model, sound localization is at a higher stage than the BTS, and localization judgments are made through a measure of the "effort" required to optimize the output to higher levels. It is recognized that the auditory signals do not arrive at the brain as replicas of complex audio waveforms; on the contrary, they are transformed, analyzed in the frequency domain, and finally pulse-coded when transmitted through neural paths. Furthermore, the idea that our brain behaves like a biological computer and the necessity for considering that any combination and processing of the signals arriving at the brain from both ears must deal with pulse-coded signals both lead to the conclusion that a switching logical process provides a much more satisfactory explanation than mathematical models based on crosscorrelation provide.

Some complementary and exploratory experiments were also performed. They were concerned with the influence of different factors, especially the space configuration, on the intelligibility of speech masked by a babble of four voices and with the masked threshold of lowpass-filtered bursts of noise in the presence of white noise.

*This research was supported in part by the U.S. Air Force (Air Force Cambridge Research Center, Air Research and Development Command) under Contract AF19(604)-6102; and in part by National Science Foundation.

†On leave from the Research Institute of Communication Science, University of Electro-Communications, Tokyo, Japan.

(XIV. SPEECH COMMUNICATION)

This report is an abstract of a thesis with the same title accepted by the Department of Electrical Engineering, M. I. T., September 1960, in partial fulfillment of the requirements for the degree of Master of Science.

J. R. Carbonell

B. PERCEPTION OF SOUNDS GENERATED BY TIME-VARIANT CIRCUITS

In the previous report (1) we described an experiment in which subjects were required to compare certain types of sounds that are generated by periodic pulsed excitation of a time-variant resonant circuit. The test stimulus was the output of the resonator when pulsed a fixed number of times; as the pulses were applied, the center frequency of the resonator moved linearly between two limits. Whenever the subject wished, he could hear a stimulus identical to the test stimulus, except that its resonant frequency remained stationary during the time of presentation. He could set this stationary frequency over a range of values by means of a continuously variable potentiometer, and his task was to make the comparison stimulus and test stimulus sound as nearly alike as possible. The results showed that for stimuli of 20-50 msec duration

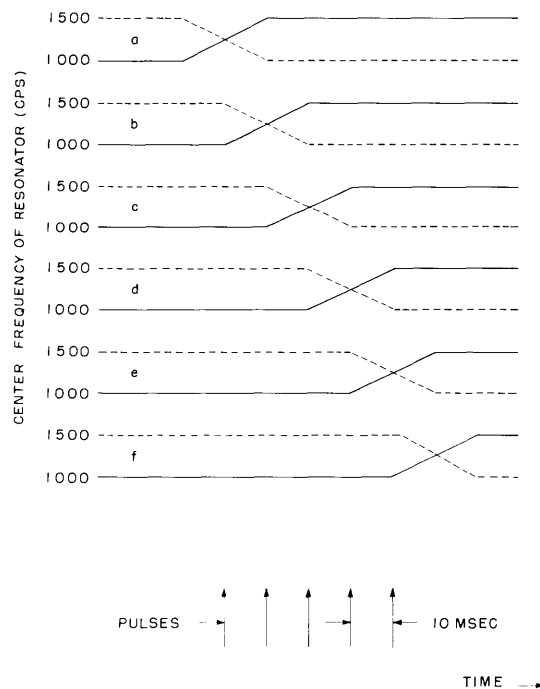


Fig. XIV-1. Stimuli used in the first test. The solid lines are paths of the resonator center frequency for rising resonances, the broken lines, for falling resonances. The five pulses used for each stimulus are shown at the base.

the matching was always made at a frequency that is very close to the final value of the frequency of the test stimulus (the last instantaneous frequency that the subject heard).

The initial experiment has been extended to include test stimuli similar to those described above, but with the difference that the time-variant portion was preceded or followed by a short interval in which the resonant frequency was stationary. The objective was to determine whether or not an initial stationary segment that preceded the time-variant segment would cause the subject to make judgments that were not in the region of the final frequency. In the first test, the resonator frequency was set to move from 1000 cps (f_1) to 1500 cps (f_2), or the reverse, from 1500 to 1000 cps. The transition was accomplished in 20 msec, and the resonant frequency was stationary before and after the transition. The resonator was pulsed five times at a rate of 100 pps, and consequently some of the pulses excited the resonator while it was stationary. The six stimuli used for rising resonances, represented by the solid lines of Fig. XIV-1, were characterized by six different starting times for the transition. The pulses are shown at the base of the figure. The first pulse of stimulus a occurs at $f = 1250$ cps; the rest occur at $f = 1500$ cps. In d, the third and fourth pulses excite the resonator while it is moving. Six other stimuli, shown by the broken lines in Fig. XIV-1, were used for the "falling-resonance" case. In a single test, each of the 12 stimuli was presented to a subject, and he was instructed to match each one with the comparison (steady-resonance) sound. The frequencies of the comparison stimuli were recorded. Each of the three subjects heard the test 12 times – a total of 36 presentations of a given stimulus to all subjects. The order of stimulus presentation was changed from test to test, with the result that any given stimulus was placed before and also after every other stimulus. The results are displayed in Fig. XIV-2. The solid lines indicate the mean values for the 36 decisions for the comparison frequency. The vertical bars indicate standard deviations in cycles per second.

Stimuli with time-variant portions of several durations were studied in a second test. The 10 stimuli used for this test are shown in Fig. XIV-3. As before, the times at which the pulses occur are indicated at the base of the figure. The significant difference between this test and the first one is that now the frequency ramps are of varying duration, with the result that they all terminate at f_2 at time 10 msec after the last pulse. (See reference 1 for details of stimulus generation, especially resonator decay rate.) Each of the three subjects made 10 matchings for each of the stimuli – a total of 30 presentations of each stimulus to all subjects. The results are displayed in Fig. XIV-4 and are interpreted in precisely the same way as those of Fig. XIV-2.

On the basis of these tests, the following observations can be made.

(i) When a subject is asked to match a sound that is characterized by a time-variant resonance with a sound that has a steady resonance, he has a strong tendency to set the steady resonance at a point corresponding to the final resonant frequency of

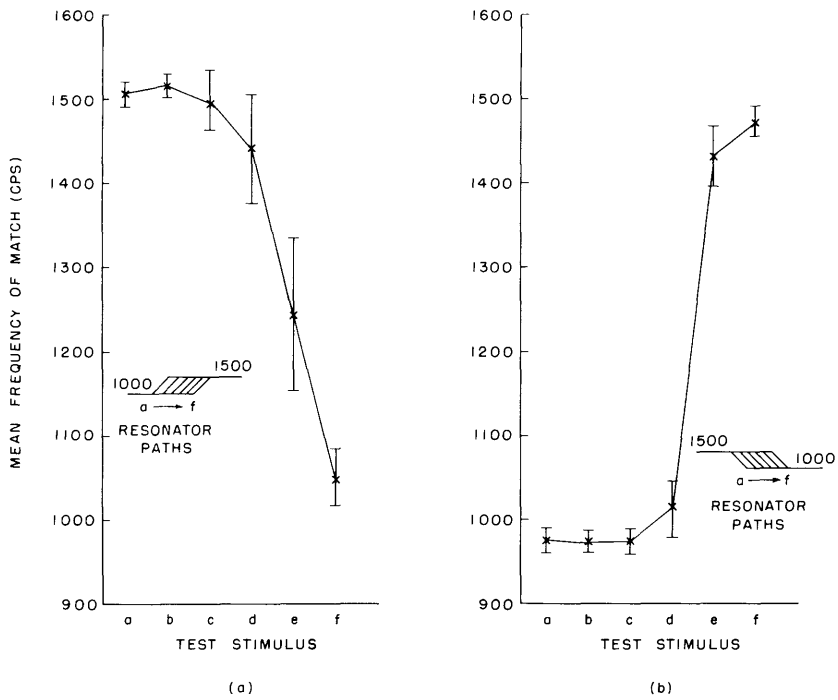


Fig. XIV-2. Results from the first test. (a) Rising resonances. (b) Falling resonances. Mean values of all the subjects' decisions, as well as the standard deviations, are plotted. Also shown is a compact display of resonator paths.

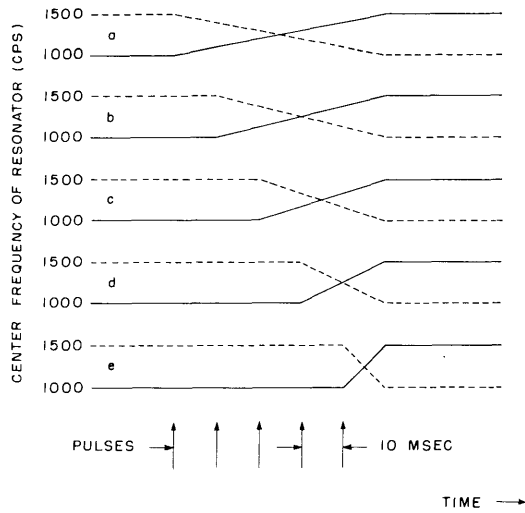


Fig. XIV-3. Stimuli used in second test. (Graphical details are the same as in Fig. XIV-1.)

(XIV. SPEECH COMMUNICATION)

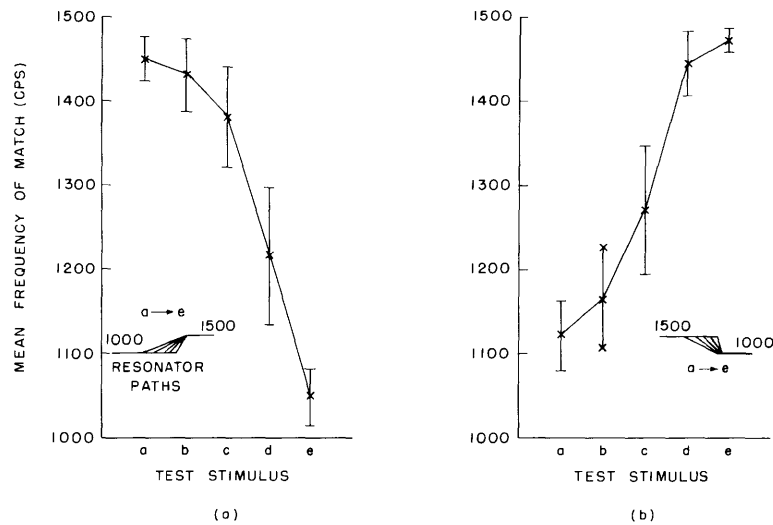


Fig. XIV-4. Results from the second test. (a) Rising resonances. (b) Falling resonances. (Graphical details are the same as in Fig. XIV-2.)

the variant-resonance sound. This effect is quite pronounced if the transition occurs near the beginning of the sound (stimuli a and b of Fig. XIV-2a and 2b), or if the transition occupies the greater part of the sound (stimuli a and b of Fig. XIV-4a and 4b). It is possible, however, with proper adjustment of the stimulus, to cause the subject to make his judgment at a frequency other than the end frequency.

(ii) The tendency to match to the end frequency is stronger for stimuli for which the resonant frequency ascends. Consider Fig. XIV-2a, stimulus f, in which the transition occurs at the very end of the sound. The subject makes a judgment at 1050 cps. With the transition occurring only 10 msec earlier, as in e, the subject has already made his judgment at a higher frequency, that is, at 1245 cps. At e in Fig. XIV-2b, the subject still has not changed his judgment to frequencies lower than the upper frequency and is still matching as if he considered the beginning portion of the sound as "most important."

(iii) The standard deviations are greater for stimuli for which matches were made at points in the intermediate frequency range between 1200 and 1300 cps. The subjects reported, on occasion, that these judgments were "harder." Therefore, we could conclude that in these cases the subjects were less sure of their answers.

(iv) The sensitivity to small (10-msec) changes in the timing of the beginnings of the ramps is quite high, especially in the test illustrated by Fig. XIV-2b. The matching frequency changes from 1015 to 4030 cps when we examine cases d and e, which differ by only a 10-msec change in the time of the onset of the ramp.

(v) A comparison of the two tests (Figs. XIV-2 and XIV-4) indicates that in the first test the subjects showed little tendency to set the matching frequency at points between

(XIV. SPEECH COMMUNICATION)

the two end frequencies (1000 and 1500 cps). The second test shows that a greater number of judgments was made in the intermediate frequency range; in fact, there is a reasonably smooth transition between the limiting frequencies as the stimulus is changed.

P. T. Brady

References

1. P. T. Brady, Perception of sounds generated by time-variant resonant circuits, Quarterly Progress Report No. 58, Research Laboratory of Electronics, M.I.T., July 15, 1960, pp. 218-220.