A MECHANISM OF ATTENTION IN HEARING: COMPUTER SIMULATION AND EXPERIMENTS
M. Tarasova, N. Dubrovskiy

To cite this version:

HAL Id: jpa-00230662
https://hal.archives-ouvertes.fr/jpa-00230662
Submitted on 1 Jan 1990

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
A MECHANISM OF ATTENTION IN HEARING : COMPUTER SIMULATION AND EXPERIMENTS

M.V. TARASOVA and N.A. DUBROVSKIY

N.N. Andreyev Acoustical Institute, Academy of Sciences of the USSR, 117036 Moscow, Shvernik St. 4, U.S.S.R.

Abstract - A model of "attention" mechanism in binaural hearing is presented together with experimental data gained to verify for predictions of a model.

1 - INTRODUCTION

A widely known manifestation of the attention mechanism in hearing is the so called "cocktail party effect", that is, the ability of a person to listen to one talker in the presence of others separated in azimuth though. In this paper we discuss our model of the attention mechanism tuned to spatial parameters of signals /1/ and present the results of our experiments which support predictions of the model.

2 - THE MODEL OF "ATTENTION" MECHANISM

The model /1/ was based on the statistical theory of optimal signal detection in noise. The incoming signal passed through space-frequency filters of right and left ears and was added to the input-related internal noise \( \xi(t) \) which represents a spontaneous neural activity in the auditory system. Then the sum of signal and noise from every ear comes on the processing block. Actual values of the human interaural base (17.5 cm) and directivity pattern of each ear were taken into account.

We considered a case of signal reception with a known azimuth \( \theta_0 \) and of its processing by the model via an optimal algorithm. This algorithm works in the following way: by using a discrete temporal representation we can write down a received message as \( Z = \lambda S + \xi \). Making a decision whether signal \( S \) is present or not is equivalent to solving a problem of verifying a simple statistical hypothesis \( H_0 : \lambda = 1 \) versus an alternative one \( H_1 : \lambda = 0 \). The corresponding procedure consists in constructing a statistics \( T(Z) \) and comparing it to a threshold \( K_\alpha \) (the hypothesis \( H_0 \) is accepted when \( T(Z) > K_\alpha \) and \( H_1 \) is accepted when \( T(Z) < K_\alpha \)). The threshold value \( K_\alpha \) is calculated from the a priori settled significance level \( \alpha \) ("false alarm" probability). The algorithm of making a decision is considered as good one if the probability \( \gamma \) of \( T(Z) \) being more than \( K_\alpha \) (probability of correct detection) is high enough when the signal is present (\( \lambda = 1 \)). So, optimization of the decision rule must consist in searching such a statistics \( T^*(Z) \) that provides a maximum value of \( \gamma \), the value \( \alpha \) being fixed (Neuman-Pierson criterion). The Neuman-Pierson lemma claims that the optimal statistics has the form of likelihood ratio and for gaussian noise it is useful to choose: \( T^*(Z) = \ln (f_s(Z)/f_{\xi}(Z)) \), that leads to: \( T^*(Z) = Z^T R^{-1} S \), where \( R \) is the noise correlation matrix. The probability characteristics of the reception process can be estimated by
using the signal-to-noise ratio of the statistics $T^*(Z): \rho_{X} = (3/\pi)^{1/2}$. The value $\rho_{X}$ defines the "statistical discriminability" of $H_0$ and $H_1$ hypotheses for a given algorithm of signal processing. If the expected signal $S_0$ on which the "attention" of the optimal detector is tuned, and the received signal $S$ differ in some parameter then the probability characteristics of the reception can be estimated using the value $\rho(S, S_0) = (3/\pi)^{1/2}$. Notice that $\rho(S, S_0) < \rho(S, S) = \rho_{X}$ and the equality takes place only if $S = S_0$. In order to describe quantitatively the decreasing of the detection quality when $S = S_0$, we have introduced a "loss function": $\mathcal{L}(S, S_0) = \rho(S, S_0)/\rho(S, S)$. We studied the decrease of detection efficiency as a function of "mismatch" between the expected source azimuth $\theta_0$ (on which the "attention" of the model is tuned up) and the actual one $\theta$, i.e. we examined the behavior of the loss function $\mathcal{L}(\theta, \theta_0) = \rho(\theta, \theta_0)/\rho(\theta_0, \theta_0)$. 

3 - EXPERIMENTAL VERIFICATION OF THE MODEL

Psychoacoustical experiments were carried out on detection of signals, the listener attention being tuned to different directions of signal arrival. Three adult males aged 23-33 years took part in experiments. A listener was placed sitting on a chair standing on a metallic net in the center of an anechoic chamber 12 x 12 x 9 m$^3$. Seven loudspeakers were located uniformly in front of the listener on a semicircumference 6 m in diameter. An angle in azimuth plane between adjacent loudspeakers was 30 degrees. Broadband stimuli were also used in the experiment.

Prior to the test session each subject's detection thresholds were determined for bursts of noise having a bandwidth of 0.1-12.0 kHz arriving from every loudspeaker. The duration of bursts was 200 ms, the rise and fall times were 20 ms. Two-alternative-forced-choice (2AFC) technique was used: signals appeared at random in the first or in the second observation intervals, which were marked by simultaneous flashes of light. The listener task was to indicate the interval a signal was presented in. After each wrong answer the signal level was raised by 2 dB; after two correct answers in succession the signal level was lowered by 2 dB. Estimates of detection thresholds turned out to be essentially the same for all loudspeakers varying for different listeners from 4 dB to 6 dB relatively to $2 \cdot 10^{-5}$ Pa.

During the test session signals arriving at random from any of seven loudspeakers were presented in series containing 200 trials with the same sound level. For each series estimates of correct detection probability (CDP) and false alarm probability were calculated allowing also to determine the detectability index $d'$. The $d'$ value defines statistical discriminability by a listener of the absence and the presence of a test signal in a series of trials. 2AFC technique allowed to obtain reliable data concerning the detection ability of the auditory system.

A listener's head being fixed, his attention was directed to a definite loudspeaker by three different ways: visually, statistically and verbally (with the aid of instructions). A visual tuning was achieved by flashing a lamp attached to a given loudspeaker. A statistical way of tuning consisted in presenting signals from different loudspeakers following a special distribution: 40% of presentations came from the loudspeaker which was intended to be tuned to and the remaining 60% were distributed uniformly over six others. Actually, a listener had no information on this distribution. In the third mode a listener got instructions prior to the experiment to expect signals only from a definite loudspeaker although they came actually from all loudspeakers but predominantly from the selected one (40% of presentations). CDP for signals arriving from the direction of tuning is considerably higher than for those arriving from other directions. As a rule, the maximal effect of attention was achieved when all three modes of listener's concentration were applied in combination. The case is presented in Fig. 1, which demonstrates the relation of $d'$ as well as CDP to an angle of signal arrival $\theta$, the attention being paid to $\theta_0 = -30^\circ$ (Fig. 1a), to $\theta_0 = 0^\circ$ (Fig. 1b) and to $\theta_0 = +60^\circ$ (Fig. 1c). The horizontal dashed line represents experimental data obtained without concentration of the listener's attention, and the solid curve 1 represents experimental data with attention concentration. The effect of increasing CDP for any direction of tuning is reliable with a significance level better than 0.05. Qualitative agreement between the experimental and the model data for $\mathcal{L}$ values is observed for all directions of tuning. Comparison of calculated $\mathcal{L}$ curves for different stimuli bandwidths with the ex-
Fig. 1 - Comparison of experimental and calculated data. The CDP and d' values vs azimuth $0$. The horizontal dashed line represents the experimental data obtained without subject concentration; the solid line 1 represents data with subject tuning. Curves 2 through 5 show the calculated values for $(0.5-2.0)$ kHz; $(0.3-3.0)$ kHz; $(0.1-5.0)$ kHz and $(0.02-20.0)$ kHz.
Experimental ones leads to the conclusion, that the listeners used only a part of the whole signal bandwidth due to the difference in absolute threshold values at different frequencies. So, the effective frequency band is close to 0.5-2.0 kHz and the experimental data exhibit the best agreement with calculated values for that band (curve 2 on Fig.1).

Our experimental data reveal that the spatial selectivity of auditory signal detection (at 3 dB loss of efficiency) varies from $\Delta = 52^\circ$ for $\theta_0 = -30^\circ$ to $\Delta = 74^\circ$ for $\theta_0 = +60^\circ$. It implies that for getting panoramic auditory perception the number of independent spatial channels $(360^\circ/\Delta)$ should be $5 + 7$. For 1 dB loss in $\lambda$ value that number can be $18 + 20$.

4 - CONCLUSION

There are two fundamental properties of auditory perception: panoramness and selectivity. The former enables a listener to perceive signals from all possible directions, while the latter allows him to increase the efficiency of perception for signals arriving from a direction where the listener's attention is tuned to.

In our model the influence of attention tuning on the efficiency of detection of exactly known signal is estimated. An optimum algorithm can be realized provided that signals pass through a bank of frequency filters resembling the critical band filters in hearing. The results of our experiments are shown to be in qualitative agreement with the model predictions, thereby proving the attention mechanism to play actually an important role in the detection of acoustic signals.

REFERENCES