HOW TO GET PARAMETERS OF THE SPEECH PERCEPTION MODEL FROM THE RESULTS OF PSYCHOACOUSTIC EXPERIMENTS?

ANATOLY VENTSOV

Pavlov Institute of Physiology
Academy of Sciences of the USSR
Leningrad, USSR, 199034

Analysis of the results of many psychoacoustical experiments has made us to conclude that we are dealing with the system adapting to the parameters of a particular set of stimuli and to the instructions, either given by the investigator or generated by the subject. In this case the concept of the continuous psychological scales seems to be invalid and some new approach to the description of the subjective mechanisms of speech perception is needed. To explain usual experimental data a special mechanism including the restricted linear scale and means for projecting the signal parameters onto the scale is proposed. Some preliminary results show the reliability of the suggested mechanism. Possible ways of further detailed investigation of the mechanism are discussed.

Any functional model of speech perception in human must be strictly formulated and efficient, i.e. it must consist of a finite set of well defined quantitative algorithms for speech signal processing and making decisions. On the whole the results of natural speech processing obtained with the model must be similar to those of human speech perception. Apparently these algorithms are to describe the processes of natural speech signal transformation into its internal subjective representation followed by subjective estimation of the results of transformation and making decision. To construct the above algorithms one may study the whole speech perception system both by neurophysiological and psychoacoustical methods. But when interpreting the results of psychoacoustical investigations one must keep in mind that the part of the system resposible for making decisions deals with parameters of subjective but not physical representation of perceived signals.

Analysing the results of many psychoacoustical investigations one has to conclude that we are dealing with the system adapting both to the parameters of a particular set of stimuli and to the instructions, either given by the investigator or generated by the subject. There are reasons to assume that in paired comparison experiments (on sound duration perception, in particular) the subjective estimates of 'longer-equal-shorter' type are more suitable than others. But when instructed to decide 'which of the two is long' (two alternative forced choice procedure), the subjects still succeed in stimulus discrimination. Thus instruction itself may cause a change in the correlation between the internal system of the subjective estimates for each stimulus and subject's responses to the stimulus. It also appears, that only by changing the instruction one can alter the 'observed' accuracy of detecting the deviations in stimulus parameters manifested in the value of the difference threshold when three categorical responses of the 'longer-equal-shorter' type are permitted /2/.

It is common observation that the classical 'point of subjective equalty' obtained in the 'constant stimuli' experiments usually falls into the center of the investigated signal parameter range, regardless of the particular mode of the studied parameter and the range value applied. Hence the obtained 'phonemic boundary' (if speech-like stimuli are involved) and the magnitude of differential limen may depend on the characteristics of a particular set of signals. Sometimes the influence of signal parameters and rigid instructions is so strongly manifested that the subject's responses do not correspond to the real physical characteristics of the signals used. Thus being presented with a steady-state vowel in the given stimulus set the subject gives under certain instruction the response - 'syllable'/1/. It follows from the above that when responding to presented signals with predetermined phonetic categories the subjects do not necessarily perceive those categories. It seems likely that the categories may be used as labels to mark the observed differences in signal parameters which are not directly connected with the given phonetic categories. Therefore in search of the numerical parameters of the speech perception system one should use signals and instructions compatible with the natural conditions of speech perception, because natural speech may be regarded as a very specific set of stimuli processed under almost unknown instructions.

To construct a reliable model of the human speech perception system it is necessary to understand among others the main principles of the system readjustment to changing external conditions whether it is the signal set parameters or the instructions that change. Some preliminary results demonstrating the possibility of these principles investigation were obtained when we studied the perception of speech-like sound duration with the paired comparison method.

The functional model of the subject's behaviour in the experiment with the constant stimuli method must include a unit for the comparison of a tested signal with a 'criterion' whether the identification or paired comparison takes place. The subject uses his internal 'criterion' in the identification task or the 'criterion' given by the investigator in the form of the standard signal in the comparison task. In fact the measurement and comparison of subjective durations take place in neural network, so one must take into account possible signal transformation errors ('noise') and the threshold character of neuron reactions. Therefore at least three independent quantitative parameters of the model should be taken into account: mean square root error (on the supposition of normal distribution) and difference thresholds for positive and negative increments of signal durations. It has been shown in /4/ that the only possibility to get the above parameters is to apply in psychoacoustic experiments 'longer-equal-shorter' or 'same-different' responses. The classical two alternative forced choice procedure cannot give the necessary model parameters.

While studying sound duration discrimination with the stimulus sets containing different standards and the 'same-different' response procedure the familiar results have been obtained: the difference threshold has appeared to grow up with increasing standard duration, the relative difference threshold being almost constant /4/. This effect could be realized in the algorithm of speech signal analysis providing that the continuous nonlinear subjective scale for durations was introduced into the algorithm. But when the subjects were presented with a large set of signals, consisted of several subsets with its own standard each, they developed the idea of an undivided set, for which the single difference threshold (positive and

negative) was established /5/. Evidently the rule of relative difference threshold constancy does not work in the case, which is totally inconsistent with the concept of the nonlinear continuous scale for subjective durations.

Suppose however that there exist a restricted linear subjective scale and special mechanism of projecting parameters of the particular set of signals onto it, which provides effective use of the scale, with difference threshold being constant in the units of the scale. Then to obtain an efficient algorithm we must know what is projected onto the scale and how it is done.

Speaking of what is projected let us consider two possibilities. 1) Subjective parameters of a particular set of signals are transformed so that their minimum and maximum values fall onto the initial and final points of the scale respectively. 2) The maximum values of negative and positive differences in the parameter between two compared signals fall onto the above mentioned points of the scale. Accordingly the difference thresholds measured in experiments are not unlikely to turn out a constant if related to the whole experimental range either of signal durations or of differences of compared signal durations.

The sign of duration difference in the signal pair may be changed in two different ways: 1) when signals in a pair are presented in the permanent order (standard-test) throughout the experiment the test is made longer or shorter than the standard; 2) when the test signal always exceeds the standard in duration their order in the pair varies (standard-test or test-standard). When the stimulus set is organized according to the first way the range of signal durations is equal to that of differences in signal durations. When the set is organized according to the second way the duration range is a half of that of duration differences. Now comparing the relative thresholds obtained in both cases we can find out what is projected onto the suggested subjective scale.

The independent values of the difference thresholds were obtained in the experiments with tonal pulses /4/ and steady-state vowels /5/. Then the difference between the positive and negative thresholds was calculated and used as a measure of the subjective differential sensitivity. Let us call it the insensitivity zone.

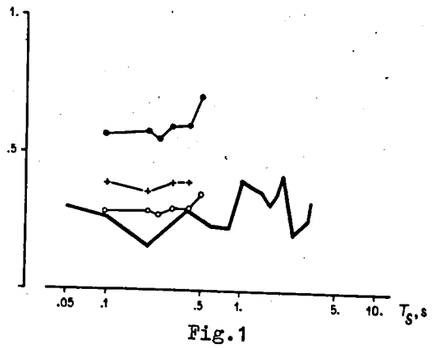

Fig.1

The normalized widths of the suggested zone are plotted on fig.1 against the standard stimulus durations. The normalization was achieved by dividing corresponding values of the zone size either by the stimulus duration range or by the range of the stimulus duration differences. The crosses mark the results obtained when the stimuli have been organized according to the first way mentioned above. The circles mark the results obtained with the stimuli organized in the second way, the filled ones representing the normalization by the range of the stimulus durations and the unfilled ones - by the range of the stimulus duration differences. For the comparison the thick line represents the normalized doubled differential limens for the duration of the gap between two acoustical clicks obtained in /3/, where the experimental stimulus set was organized according to the first way.

As far as one can judge from the picture the normalization by the range of the duration differences brings into a good agreement the results obtained in a quite different experimental conditions. So it may be concluded that it is the difference of the compared signal parameters that is projected onto the suggested subjective scale.

Though the collected data are insufficient to make any final conclusion, they nevertheless show a possible way of studying the processes of creating the internal psychological representation of natural speech signals. It may also be supposed that in search of the mechanisms of projecting the signal parameter onto the proposed subjective scale, one must investigate the dynamics of the subject's responses to one and the same stimulus throughout the experiment.

REFERENCES

/1/ Dechovitz D., Mandler R., "Effects of transition length on identification and discrimination along a place continuum", Hask. Lab. Stat. Rep. on Speech Research, SR-51/52, pp. 119-129. 1977.

/2/ Fernberger S., "Instructions and the psychological limen", Amer. Journ. Psychol., v.43. 1931.

/3/ Getty D.J., "Discrimination of short temporal intervals: A comparison of two models", Percep. & Psychophysics, v.18, pp.1-8, 1975.

/4/ Ventsov A.V., Malinnikova T.G., "Modelling the subjective mechanism of duration comparison", in Issledovanije modelej recheobrazovanija i rechevosprijatija, Leningrad, Nauka, 1981, pp.19-36.

/5/ Ventsov A.V., "What the differential subjective sensitivity on duration depends on?", in Doklady X Vsesojuznoj Akusticheskoj konferencii, Moscow, 1983, sec.X, pp.18-21.