Adaptive measurement of speech intelligibility with the use of the maximum likelihood procedure

Jan Zera
Central Institute of Labour Protection – National Research Institute
Czerniakowska 16, 00-701 Warsaw, Poland
jazer@ciop.pl

and
Chopin Academy of Music
Okolnik 2, 00-368 Warszawa, Poland

Abstract
An adaptive maximum-likelihood procedure (MLP), typically used for threshold measurements in psychoacoustics, was adopted for the measurement of speech intelligibility, with the use of the modified rhyme test (MRT) [1]. The advantage of the MLP over the fixed step up-down adaptive procedures usually employed in adaptive speech intelligibility tests is that the MLP imposes no restrictions on the selection of the target percent-correct word score. An adaptive test based on the MLP permits quick estimation of the S/N ratio at any word score, pre-selected as an input parameter. Owing to its high efficiency, the MLP is a useful tool for rapid assessment of the S/N ratio of communication systems. As an example of such an application of the MLP–MRT adaptive test presented are preliminary measurements of speech intelligibility conducted on listeners wearing hearing protectors.

1. Introduction
The measurement of speech intelligibility with the use of adaptive procedures was described in a number of papers [2, 3, 4, 5, 6, 7]. When such procedures are used, the S/N ratio is determined in each presentation of a test item by the subject’s responses to preceding test items. The published studies were carried out with the use of staircase up-down procedures with a fixed step size [8]. These procedures make it possible to estimate only certain values of percent-correct responses, such as 50%, 70.7%, or 79.4%. The present paper describes a speech intelligibility test based on an adaptive maximum-likelihood procedure (MLP) originally developed for the measurement of sensory thresholds [9, 10]. The MLP is considered the most efficient adaptive procedure [11]. Its effectiveness results in a small number of test items required to complete an adaptive run. Moreover, the MLP allows to conduct the measurement for any, freely selected percent-correct word score.

This paper presents certain properties of an adaptive speech intelligibility test based on the MLP. The investigations reported here included an assessment of variability and accuracy of level estimates predicted with the use of Monte Carlo simulations as well as measured for real listeners, in a psychoacoustic experiment. Finally, presented is an example application of the adaptive MRT for the assessment of the influence of a hearing protector on speech intelligibility.

2. The adaptive test
The adaptive MLP is based on the assumption that the shape of the subject’s response curve associated with a given task is known prior to conducting the measurement and is invariant when expressed as a function of stimulus intensity. In each stimulus presentation, the stimulus level is adjusted so as to correspond to a given percent-correct value on the most recently determined response curve. The S/N ratio is changed in every presentation of a speech sample, according to the target S/N ratio corresponding to certain percent-correct word score on the response function selected with the maximum likelihood.

The adaptive test described here was based on the Modified Rhyme Test (MRT) published by House et al. [1]. The MRT is a closed response set test with a vocabulary of 300 monosyllabic CVC words split into six 50-word lists. In each test item, the subject selects a single word from six possible responses. For the purpose of the adaptive MLP-MRT it was assumed that the response curve used in the MRT is cumulative Gaussian on a logarithmic S/N ratio scale (Fig. 1). Since the MRT response is a six-alternative forced-choice task, the response curve ranges from 16.7 to 100% correct word score. The slope of the response curve corresponding to a standard deviation of $\sigma = 6.8$
dB was determined by fitting the cumulative normal curve to the data reported by House *et al.* [1].

![Figure 1](image1.png)

*Figure 1.* The response curve used by the adaptive MRT (solid line). The standard deviation of estimates of the S/N ratio calculated from the variance of the binominal distribution is shown by a dashed line. The filled circle indicates the minimum variance of $\sigma = 9.9 \cdot \sqrt{n}$ dB (the sweetpoint) at S/N = 1.4 dB. Additional points on the response curve show the levels at which the measurements were conducted. Figure taken from ref. [13].

### 2.1. Numerical simulations

Although the MLP makes it possible to set the testing level at any target S/N ratio and therefore at any target percent-correct word score, the minimum variance of level estimates are obtained when the measurements are conducted at a specific point on the response curve called the sweetpoint [12, 10]. Estimating the S/N ratio away from the sweetpoint results in an increase in the variability of estimates as shown in Fig. 1.

Monte Carlo computer simulations (Fig. 2) of the adaptive MLP-MRT, and runs comprising number of trials from 5 to 200, were conducted at target correct word scores selected as shown by circles in Fig. 1 to investigate the variability and bias in S/N estimates [13]. These simulations revealed (Fig. 2) that in most conditions standard deviation of level estimates agreed well with theoretical relation $1/\sqrt{n}$. Moreover, only a small amount of bias was observed in the data. For 25 test items which were also used in experiments with subjects, the bias was about -0.3 dB for range of levels corresponding to the middle region of the response curve. Bias was larger (-1.68 dB) for levels close to the chance level of 26.7% on the response curve.

![Figure 2](image2.png)

*Figure 2.* Monte Carlo simulations of adaptive MLP–MRT. The upper panel shows the procedure bias (0 dB corresponds to the speech-to-noise ratio as in Fig. 1). Lower panel shows the standard deviation of the level estimates. The number of test items in an adaptive run ranges from 5 to 200 (run length). The results are obtained for the target correct word score of 26.7% (cross), 41.7% (circle), 58.3% (diamond), 65.1% (square), 75% (triangle), and 90% (star) corresponding to circles in Fig. 1. Dashed lines show the theoretical limit for the standard deviation resulting from standard deviation calculated in Fig. 1. Figure taken from ref. [13].

The subject's task was to select the correct answer from the set of words displayed on the screen. The MRT words were presented in the background of continuous pink noise at 75 dB SPL. The speech level was an adaptive variable. A PC-compatible computer, equipped with the System II from Tucker-Davis Technologies (TDT) controlled stimulus presentation, sampled the responses and executed the adaptive procedure. The speech samples recorded by a male speaker were digitized from a commercially available...
MRT recording, and stored on a computer hard disk. The speech samples and noise were presented binaurally through Beyer Dynamic DT 911 headphones.

Figure 3. Signal-to-noise ratio determined using the adaptive MLP–MRT. The S/N ratio adjusted by the adaptive procedure is displayed on the abscissa. Measured correct word score for estimated levels is displayed on the ordinate. Mean result of 15 adaptive runs per point measured at target correct word score of 26.7%, 41.7%, 58.3%, 65.1%, 75%, and 90% (solid symbols). Open symbols represent the subject’s response curve derived from all data collected across the adaptive runs. Dashed line shows the cumulative normal function as in Fig. 1. One subject (non-native speaker). Figure adopted from ref. [13].

At each selected target percent correct word score 15 adaptive runs consisting of 25 test items were used to estimate the corresponding level. These were followed by another 25 test items were presented at a fixed S/N ratio to estimate the percent correct word score at the level determined by the adaptive test.

The S/N ratio estimated in adaptive runs is shown on the abscissa in Fig. 3. The percent correct word score is shown on the ordinate. Coordinates of each solid point represent the level estimated by the MLP–MRT procedure, and the mean percent-correct word score determined at levels obtained from the adaptive runs. At the high end of the percent correct score, the responses do not reach 100% correct, even for high S/N values. This finding may be attributed to the fact that a non-native speaker of English participated in the experiment. The curve with open symbols is the subject’s response curve derived from all data collected across the adaptive runs. As such, this curve has not been influenced by any possible bias related to the adaptive procedure therefore it may be used as a sort of a reference curve. Results of the adaptive test are located close to the reference in the middle range of correct word scores. The observed differences in level do not exceed 1 dB, and differences in percent correct scores are less than 2%. In general, it may be concluded that the bias of data obtained by the MLP–MRT procedure used in this experiment is smaller than the bias level predicted from simulations.

2.3. Example of the use of an adaptive MRT for assessment of the effect of wearing hearing protectors on speech intelligibility

Three listeners, aged from 23 to 46 participated in this experiment. The listeners were native speakers of Polish and had normal hearing. Listeners were tested individually in a sound-insulated room. Speech samples were reproduced from a loudspeaker located in front of the listener, at distance of 1.5 meter. Four loudspeakers located in the corners of 3 x 3 meter square were used for reproduction of noise. The listener was seated in the center of the square and used a computer screen to select the response from the word table presented on the monitor screen.

In the first condition, pink noise at 86 dB SPL was used as a masker and no hearing protector was used. In second condition, the listeners were using Peltor type H7A hearing protector. An adaptive run comprised 15 test items. Each listener performed ten measurements (adaptive runs) per condition. Since the listeners were non-native speakers of English, they were given training sessions to become familiar with the test material. During the training sessions no noise was presented, and the speech items were played through the loudspeaker at a comfortable listening level.

Figure 4. Change in S/N ratio required to maintain target percent word score when Peltor H7A hearing protector is worn, as estimated by the adaptive MRT test. Results of listeners L1, L2 and L3 are shown by triangles circles and squares, respectively. Filled symbols and open symbols refer to 75% and 58% correct word score, respectively. Average standard error of level estimates equals 1.1 dB.
The effect of the use of hearing protector on the level estimated by the adaptive maximum likelihood MRT test is shown in Fig. 4. Results of the three listeners are indicated by different symbols. The S/N ratio determined by the adaptive MRT when hearing protectors were donned or doffed is shown on the abscissa in Fig. 4, for target correct word score of 58% and 75% (open and filled symbols, respectively).

An adaptive test is a convenient means to express the influence of wearing the hearing protector by a change of effective speech-to-noise level required to preserve similar percent correct word score. This change is shown by arrows in Fig. 4. For listeners L1 and L2, lower S/N ratio by about 2-3 dB is needed to obtain a correct word score of 75% when the hearing protector is used. For listener L3, the improvement is even larger, amounting to about 4 dB. At 58% correct word score, the change in effective level is less that 1 dB for listeners L1 and L2, and maintains at about 4 dB for listener L3.

3. Summary

The present experiments have demonstrated that speech intelligibility can be effectively measured by the adaptive MLP-MRT procedure, with the use of only 15 or 25 test items in a run. The advantage of the maximum likelihood procedure is that the S/N ratio quickly stabilizes at the level of interest, and most of the words can be presented close to such a level. The ability of the MLP to converge in a small number of trials is an advantage over other adaptive methods. In the present experiment, the adaptive run practically converged in the listener's initial five responses. It was thus possible to measure the percent intelligibility using the listener's responses taken from the final part of the adaptive run.

The MLP-MRT procedure allows to conduct the test at a level corresponding to some preselected target intelligibility soon after the start of the adaptive run. The procedure may be a useful tool for quick assessment of communication systems at various values of target intelligibility, with practically no influence of the starting level in the test. The adaptive method can be easily extended to other intelligibility tests employing different number of test items in the response scheme.

4. Acknowledgements

This work was supported by a grant from KBN PCZ 16-21-2.3, and KBN 06-21.

5. References


