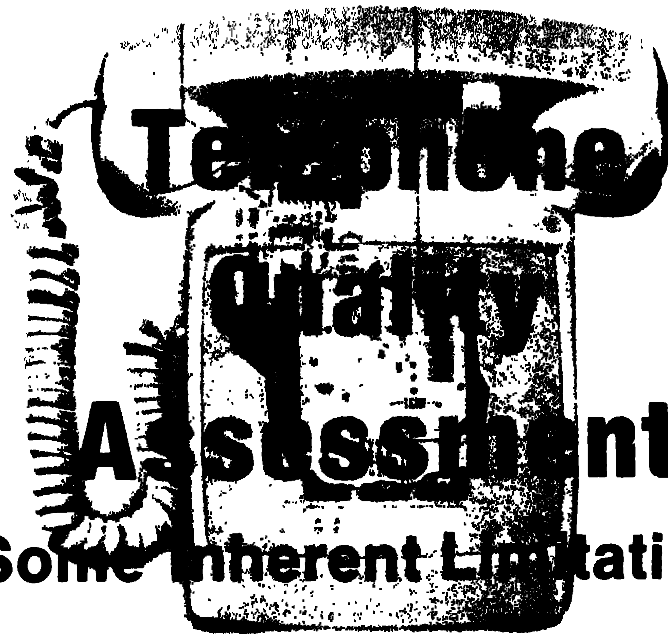


Telephone Quality Assessment: Some Inherent Limitations

B.L.K. Rao and T. Usmani

For a long time, transmission quality of telephones has been assessed on the basis of subjective comparisons between different telephones. Although this psycho-acoustic subjective method is close to real life practice, it has got various limitations which will be discussed later. The objective type electro-acoustic method of quality assessment also does not fulfil the high expectations of telephones users.

In fact, the expectations of telephone users with regard to quality have increased to a large extent today. Intelligibility and audibility of speech have now become the major factors of telephone quality. The real problem lies in the fact that they are purely subjective and totally independent of each other. Also, these parameters may vary from man to man depending upon personal liking, age group and climatic conditions etc.

In a recently published article, Mr Bernhard Rall of AEG Telefunken comments over the same problem: "Telephone user, while appreciating a good volume level, lack of distortion and a minimum of background noise, very rarely even holds the handset for optimum speech intelligibility. Apart from simple human factors, far more complicated technical factors affect speech quality characteristics such as sending, receiving, and side tone reference equivalents which are used to describe the telephone's essential properties. These are very expensive to measure and provide only an indication of mean loudness. Apart from this, simple electrical measurements like telephone input impedance show substantial differences between different types, depending on age, earphone type, manufacture and on the position of the handset when carbon microphones are used. Variation in transmission quality can also arise from the use of different telephone switching equipments, different cable lengths and dimensions, different transformer relay circuits, feeding, inductors" etc.

Test arrangements

Test arrangements may either be subjective or objective.

Subjective tests are those based on real life practices. Here, normal conversation is conducted between two parties without putting any artificial means in the system and the speech quality rated as good or bad by the persons involved. Such tests give the real picture of the system, termed as valid tests. Subjects can express their amount of satisfaction over speech quality, handset comfort, ease of dial operation, suitable loudness level etc which cannot be measured numerically.

Objective tests are tests performed in laboratory where artificial equipment arrangements are provided for electro-acoustical measurements. Here, artificial ear and artificial mouth are used alongwith various other test arrangements. These tests are simpler and quicker with high repeatability. They are best suited for regular quality evaluation as the numerical answers obtained form the easy basis of acceptance or rejection.

It is clear that subjective tests are more 'valid' tests since they are based on real life experiences. Their repeatability is low as different person's sensation of the same stimulus is quite different from each other. Hence, similar experiments with the same people may give different results at different times. On the other hand, the validity of an objective test may always be questioned as to whether the electro-acoustical test arrangements truly represent the real life practice or not.

To be more specific, recording the telephone receiver frequency response curve is invalid if the volume coupling of the receiver to the measuring microphone does not approximate the acoustic impedance of the human ear. In such a case, the receiver which was passed in the lab, may give bad results when put to actual use. Hence the tests which are devised under laboratory conditions should be firmly based on actual life experiences and the experimenter has to clear all the doubts in this regard.

While objective tests are artificial and results may be in terms of somewhat irrelevant quantities, subjective tests are imprecise and many important factors are largely unknown as well as uncontrollable. Thus, to adopt a balanced measurement procedure, the physical and other conditions in the field are first surveyed. Laboratory tests are then conducted under conditions shown to be relevant by the field survey. The final assessment is then made in the field using, as far as possible, the same assessment criteria as used in the laboratory. Both subjective and objective types of measurements are then made to arrive at a particular acceptable passing range for a particular characteristic. In fact both types of systems have got their own limitations, where a practical correction factor is to be evaluated for necessary correlation between each other.

The wide variety of characteristics for telephone connection to achieve fundamental information requires systematic planning of telephone network. It must always take into account factors including loss, circuit and room noise, various distortions, side tone, echo, propagation time, effects of various voice switching devices and finally the overall effects of all practical combinations.

THE TEST PIECE

Telephone instrument, the subscriber apparatus, has got its own various limiting factors for a perfect design. Handset dimensions, microphone and earphone characteristics, side tone balancing etc are the vital factors of interest in this regard.

Handset dimensions

The distance between the centre of the opening of ear canal which is taken as reference point of listening and the centre of the lips which defines the reference point of mouth is widely studied internationally over various human heads for the development of an ideal handset. One more problem is that the radiation pattern of a real mouth does not allow all the speech sounds to enter the mouthpiece of the handset while speaking. True reproduction of speech into electrical and back to acoustical form is also considerably difficult. Another serious complication arises due to the fact that different sounds are generated from different positions in the mouth which refers to variation in source position.

According to an expert, "the variability between the physical dimensions among the human heads including the fact that the female heads are not only smaller than male ones, but also yield less reproducible results, causes complication for the design of telephone handsets... An angle is also required to define the direction in which speech is emitted from the mouth into the mouthpiece of the microphone. This is very difficult to define precisely."

Carbon microphone

Carbon microphones which are used as acoustical to electrical energy converters have their own severe drawbacks. Perhaps it is the weakest link in the telephone network where a lot of precautions have to be taken for proper output response. Proper conditioning of the microphones before actual measurements is also very essential and has been described by the International Telegraph and Telephone Consultative Committee (CCITT) in the following manner:

"The CCITT considers that since the characteristics of carbon microphones are strongly dependent on conditioning techniques, it is necessary to follow a consistent procedure prior to measuring sensitivity/frequency characteristics

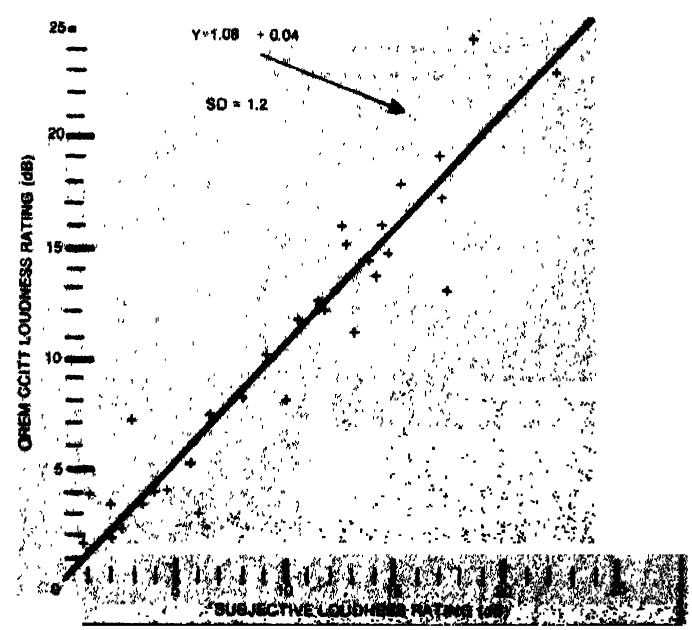


Fig. 1: Comparison of results using subjective and instrumental methods for determination of loudness ratings

in order to obtain reproducible results. The following steps are specified for the standard conditioning method:

1. Place the handset in a holding fixture with the handset clamped in a position corresponding to that in which the microphone is going to be measured.

2. Connect the microphone or telephone set terminals as required to the DC feed circuit and appropriate terminating load.

3. Turn the feed current on. After 5 seconds, condition the microphone by rotating it through an arc slowly and smoothly. The microphone face should reach a vertical plane during the initial part of the rotation. In this vertical plane, a reference vector should be visualised which passes through the centre of the microphone and points straight up. Rotation should be continued until this reference vector points straight down (i.e. movements of the reference vector through 180°). The direction of rotation should then be reversed and the microphone returned to its starting position.

Without interrupting the DC current or jarring the microphones, repeat this process two more times. The speed of rotation is not critical but should be slow enough to ensure that the effect of centrifugal force on the carbon granules is negligible. Finally, return the handset to the measuring position."

The real voice sensitivity measurement of carbon microphone is very difficult. A voice test team at Standard Telecommunication Lab (STL) in England was set up to determine the sensitivity frequency characteristics of a wide range of telephone subsets. They noted that, "although it is just over one hundred years since the carbon microphone was invented, it is evident from a number of factors that we do not as yet know how to measure its performance adequately. For instance, if we compare subjective test results of loudness ratings with loudness rating calculations made from objective measurements of the sensitivity frequency characteristics of carbon microphones, large discrepancies exist which are not wholly accounted for by expected subjective test variability. Documentary evidence of these facts is quite clear in the results of the CCITT programme for the development of a new method for loudness rating of telephone sets."

Fig. 1 shows the results of some sound loudness rating determinations made subjectively and objectively in the telephone set evaluation laboratory at STL. Here it is plain that a few telephone sets appear to need special treatment, or perhaps different treatment, from others with respect to the determination of their sensitivity frequency characteristics for loudness rating calculations.

This team took speech repetition tests using various methods including real voice excitation tests. Sets of male voice as well as female voice were used. The final loudness rating calculations showed that about 2dB discrepancy exists in the determination of carbon microphone sensitivity comparison.

The sensitivity of carbon microphone is defined as follows:

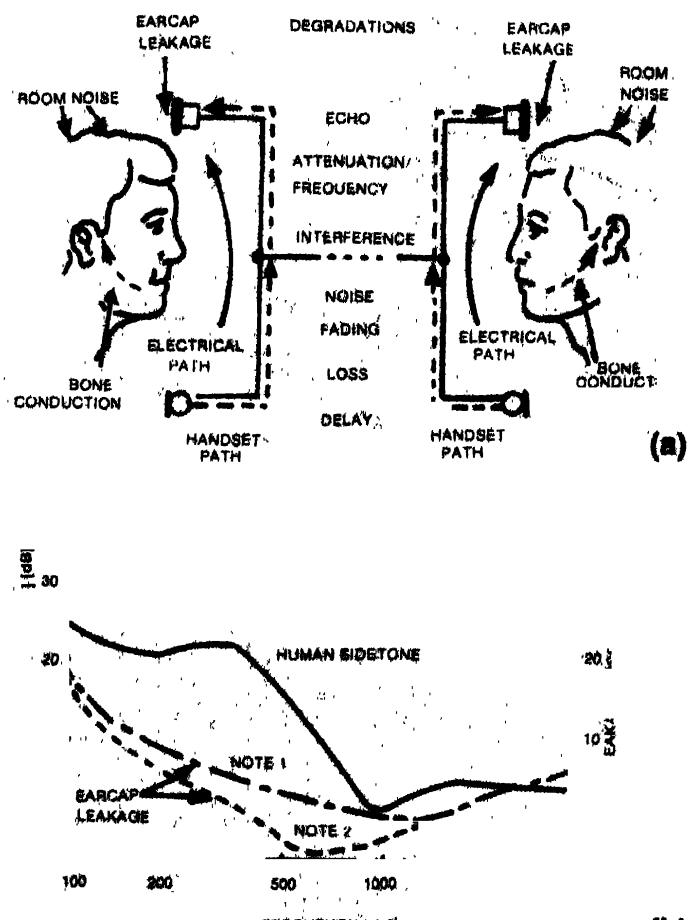
$$S_M = 20 \log_{10} \frac{\text{Open Circuit Output EMF}}{\text{Free Field Sound Pressure}}$$

at mouth reference point, limits being relative to 1V per N/m^2.

By expressing the sensitivity in this manner, the effects of differences in the lengths of handset will be appropriately reflected in the results under practical condition of service, considerable departure from so-called 'model' position will naturally occur and these will be accompanied by variations in vocal level.

Earphone characteristics

The sensitivity of the earphone represents the output acoustical energy with reference to the given electrical power. It is measured at the ear reference point by coupling a



NOTE 1: EARCAP LEAKAGE USED IN CALCULATION OF SIDETONE LOUDNESS RATING (STMR)
NOTE 2: EARCAP LEAKAGE UNDER STUDY BY THE CCITT AS BEING REPRESENTATIVE OF REAL TELEPHONE USE.

Fig. 2: (a) Sidetone paths. (b) Human sidetone loss and earcap leakage/frequency characteristics.

suitable sound pressure measuring device designed to represent the acoustical load of human ear which is termed as artificial ear. Mathematically, sensitivity is expressed as

$$S_E = 20 \log_{10} \frac{\text{Output Sound Pressure in N/m}^2}{\text{Voltage Across Input Terminals}}$$

The unit being dB relatively to 1 N/m² per volt. Here also the background noise as well as the artificial coupling of the ear piece with the human ear play an important role. A good sealing between earphone and ear, and proper shape of handset etc are also taken into account.

Side tone balancing

Side tone is the reproduction in a telephone receiver, of sounds picked up by the microphone of the same telephone sub-set. Side tone balance is satisfactory if the environmental or room noise picked up by the microphone, possibly including the speaker's own voice, are only heard at an attenuated level in the receiver. Inadequate side tone balance adversely affects conversation due to: (i) room noise at listener's end, and (ii) speaker induced into lowering his own voice. As there is increasing traffic over telephone sub-sets located in rooms with high ambient noise level, side tone balancing has become very important.

The CCITT-recommended STRE of about 9 dB using NOSFER, is not met in many cases. This is due to mismatching between the hybrid balancing network and the line impedance seen at the terminals of a telephone sub-set. Effective measures to improve the side tone balance can only be taken if the influence of the parameters of transmission path and telephone instrument is known. To this end, it is necessary to analyse the circuit characteristics of telephone sub-set, the type of cable used, and the impedance transparency of the trunk circuits of switching system, taking into consideration all devices connected to the transmission path.

In fact, the transmission performance of a telephone connection varies as far as the customer is concerned, not only with the overall loss in the circuit but also with the amount of side tone present in the circuit. There are various types of side tones, controllable and uncontrollable, both of them having their own effects.

During a telephone conversation, the speech enters a talker's ear by three paths, electrical path through the telephone set, a bone conduction path and an earcap leakage path. Also, during the conversation, room noise enters the listener's ear via two paths, the electrical path and the earcap leakage path. These effects are known as talker's side tone (speech side tone) and listener's side tone (room noise side tone) respectively. In addition, there can be an acoustic path through the handset itself, resulting from imperfect sealing of transducers (Fig. 2).

TRANSMISSION QUALITY PARAMETERS

Perception of speech transmission quality is two dimensional and may be divided into two independent factors—intelligibility and naturalness. Intelligibility depends upon the loudness of speech signals while naturalness represents the individual characteristic of the speaker (identity etc).

Till today, no instrument is available for measurement of overall transmission quality. Different telephone administrations have adopted different methods, which basically

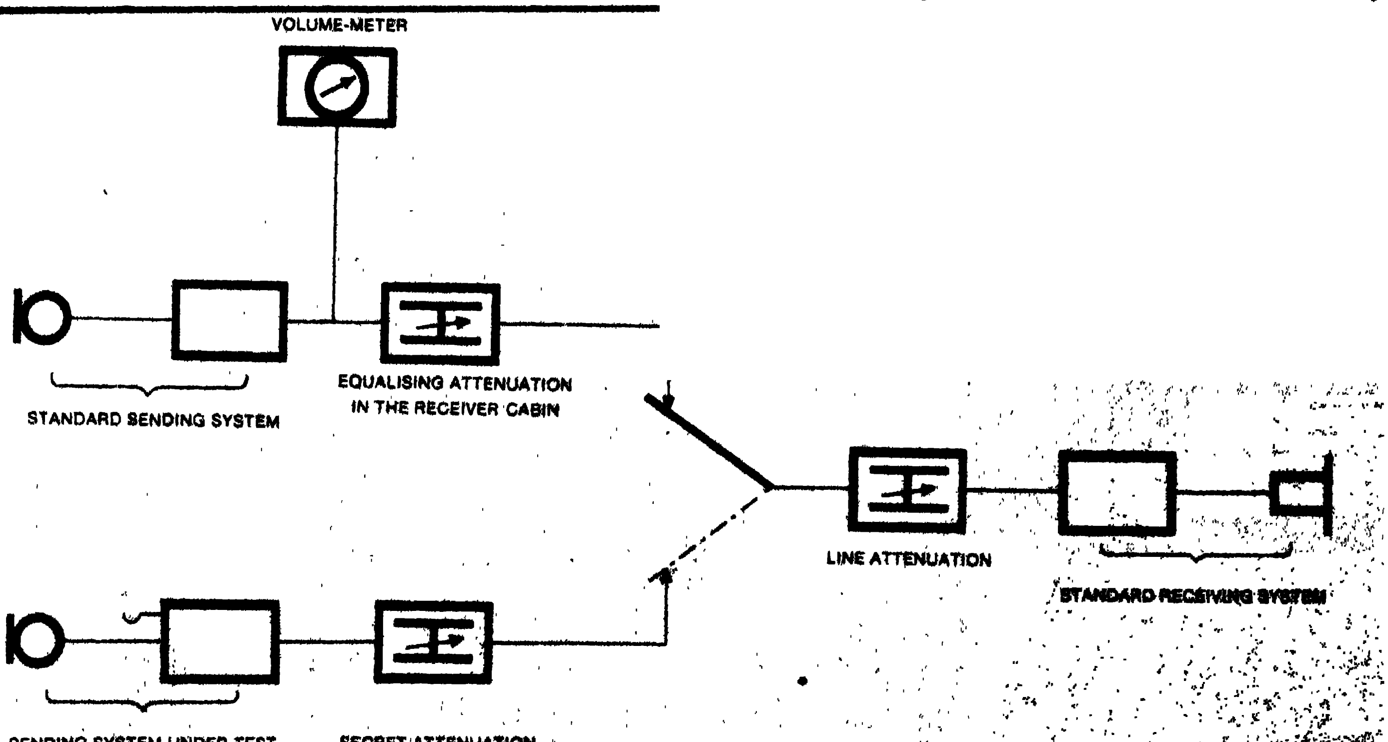


Fig. 3: Measurement of sending reference equivalent.

depend upon subjective ratings. CCITT has also recommended a few of these methods and has tried to standardise them as much as possible. Some of these methods are described below.

Reference equivalent

This is the characteristic for the loudness of a telephone system, expressed in dB, indicating the extent by which the system under test is louder or less louder than the zero level determined in the reference system of new fundamental system for determination of reference equivalent (NOSFER) approved by CCITT.

Thus the reference equivalent is a relative loudness value measured by means of voice and ears with respect to NOSFER system. NOSFER is a set of equipments available in CCITT laboratories which comprises of the following:

1. The transmission path, sub-divisible into sending end, junction and receiving end.
2. A centralised apparatus for the supply of room noise and interconnection facilities.
3. A calibration equipment.

The comparison is direct if the NOSFER is used and indirect if a secondary system compared previously with NOSFER is used as standard.

The measuring set up for sending reference equivalent is shown in Fig. 3. One of the persons participating in the measurement reads the test text into the microphone of the system under test and into that of the standard system (NOSFER). Another person at the receiving end listens to the speaker and adjusts both systems to identical loudness by means of an attenuator box. Normally, a trained group of three persons each participate at both ends. The attenuation value chosen, shows how much worse/better the test system is, compared to the reference system which is in turn the send reference equivalent of the test system. Reference equivalent is negative if the test system is better than the reference system. Otherwise it is positive.

The block schematic used for the measurement of receive reference equivalent is shown in Fig. 4. Here the listeners have to listen in the receiver under test and compare it with the receiver of the standard system (NOSFER). The identical loudness is achieved with the help of attenuator box and the attenuation value indicates the receive reference equivalent of the system under test.

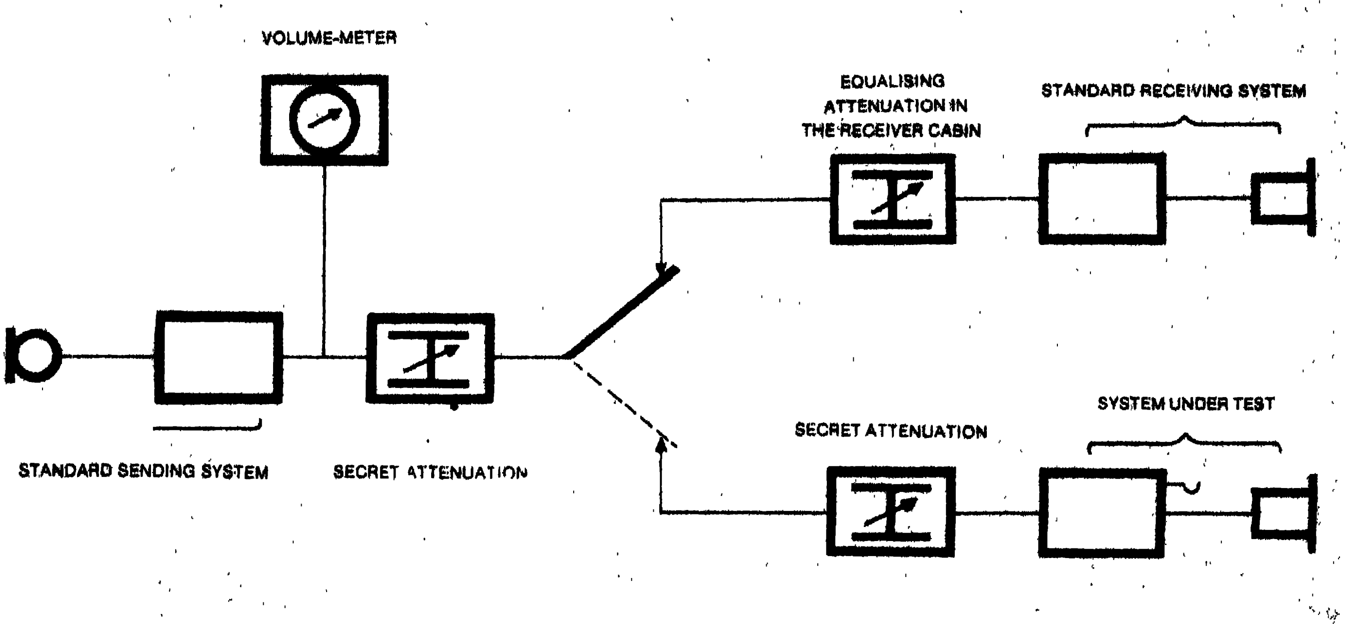


Fig. 4: Measurement of receiving reference equivalent.

The reliability of the measurements of the reference equivalent may be affected by many factors, as for instance the tone and the loudness of the speaker, the variations of the same during measurements, the frequency response of the standard and the system under test, the relationship between the frequency characteristics, the hearing abilities, and the fatigue of the listener etc. The temperaments and the objectivity of the people conducting the experiments also make differences in these measurements. CCITT records these limitations in the following manner:

"A desire to depart from use of reference equivalents as defined by recommendation P72 arises for the following reasons:

1. Reference equivalents cannot be added algebrically, discrepancies of at least ±3 dB are found.
2. Replication accuracy of reference equivalent is not good. Changes in crew can cause changes of as much as 5 dB.
3. Increments of real (distortionless) transmission loss

are not reflected by equal increments of reference equivalents. 10dB increase in loss results in an increase in reference equivalent of only about 8 dB."

Articulation tests

It is the test of intelligibility of the speech sound, which is taken as quality measure. Here one has to distinguish the sound syllables, words and sentences. The ratio between the correctly understood sound elements and the total number of transmitted elements becomes the measure of articulation score. While a well conducted articulation test can provide an enormous amount of information about the speech quality which will be very close to real life practice, they are quite time consuming, requiring a highly trained test team to conduct the test for a fairly long time, sometimes several weeks. Analysis of the test results is also a delicate job which needs expert advice at all stages.

Opinion tests

In these tests, conversation between two teams takes place over telephone connection under test. The participants are then asked to rate the speech quality. Rating is done in five steps starting from bad to excellent. The procedure is repeated many times so that different test variables are taken into account. The opinion score is then combined giving the mean opinion score which is combined in curves.

As is evident, these tests are purely subjective, and reliability is very low since noise level and attenuation play an important role in final results.

Conclusion

Over the last few years, the expectation of telephone users regarding quality of telephone has increased tremendously as quality has become a way of life today. Various telephone organisations in different parts of the world are busy to determine systems for absolute measurements of overall quality of telephones. New parameters for telephone quality are being identified in different telephone laboratories and being standardised by CCITT.

Indian Telephone Industries Ltd (ITI) has been making serious efforts to simplify and standardise the quality assessment methods for telephones. Over the years, several new test arrangements have been developed and many instruments have been added in quality testing labs. In Naini, ITI has imported B & K electro-acoustic telephone transmission measuring system type 3354 for final assessment of telephone quality. A Canadian microprocessor based automatic telephone test equipment is also under procurement. ITI has also established a life test laboratory as well as a climatic test laboratory. Efforts are being made to establish an anechoic chamber for measurement of quality parameters of a telephone under absolute conditions. □