INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

Supplement 11

(Series P)

(03/93)

TELEPHONE TRANSMISSION QUALITY

**MEASUREMENTS RELATED
TO SPEECH LOUDNESS**

SOME EFFECT OF SIDETONE

**Supplement 11 to
ITU-T Series P Recommendations**

(Previously "CCITT Recommendations")

FOREWORD

The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of the International Telecommunication Union. The ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Conference (WTSC), which meets every four years, established the topics for study by the ITU-T Study Groups which, in their turn, produce Recommendations on these topics.

Supplement 11 to ITU-T Series P Recommendations was prepared by the ITU-T Study Group XII (1988-1993) and was approved by the WTSC (Helsinki, March 1-12, 1993).

NOTES

1 As a consequence of a reform process within the International Telecommunication Union (ITU), the CCITT ceased to exist as of 28 February 1993. In its place, the ITU Telecommunication Standardization Sector (ITU-T) was created as of 1 March 1993. Similarly, in this reform process, the CCIR and the IFRB have been replaced by the Radiocommunication Sector.

In order not to delay publication of this Recommendation, no change has been made in the text to references containing the acronyms "CCITT, CCIR or IFRB" or their associated entities such as Plenary Assembly, Secretariat, etc. Future editions of this Recommendation will contain the proper terminology related to the new ITU structure.

2 In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

CONTENTS

Supplement 11

SOME EFFECTS OF SIDETONE

(Malaga-Torremolinos, 1984; amended at Melbourne, 1988 and Helsinki, 1993)

(referred to in Series P Recommendations)

1 Introduction

Over a number of years sidetone has been studied and some important conclusions have been reached from the point of view of the subscriber in his role as both talker and listener. These conclusions relate to the effect of sidetone on a subscriber, as he hears his own voice, the way his talking level changes as a result and some effects of sidetone when the subscriber is listening in conditions of moderate to high-level room noise. These effects are summarized in Figures 1 and 3.

The relationship between talker and listener sidetone for a given telephone depends primarily on two factors: a) the geometry of its handset and b) whether or not there are any non-linear gain or loss characteristics in the sidetone path. Some guidance for telephone set designers is provided in clause 4.

Some information is also provided concerning the increasingly frequent occurence of short delay talker echo, which may be perceived as unpleasant talker sidetone.

2 Talker sidetone

Figure 1 shows that there is a preferred range for sidetone when the subscriber is talking under quiet conditions, and that the difference between the sidetone being objectionable or too quiet is of the order of 20 dB. (These results were obtained from talking-only tests and need to be confirmed by conversation tests.) The preferred range lies between 7 and 12 dB, STMR (sidetone masking rating – Recommendation P.76) [1], [5].

The acceptable range is wider and lies between an STMR of 1 dB and 17 dB, (although it must be stated that increasing STMR to a value greater than 17 dB is likely to affect only the talking level, and that only marginally). This range corresponds to the difference between the two curves at the 50% appraisals level. It is not proposed that the 17 dB figure should in any way be considered a maximum value. However, for an STMR above 20 dB, the connection sounds “dead”.

For telephone connections where the OLR is in the preferred range, the STMR values may similarly be positioned in the preferred STMR range given above. However, on high loss connections the STMR value should be close to, or even exceed 12 dB. On low loss connections the STMR value may be sometimes permitted to become less than 7 dB, but only rarely should it become as low as 1 dB, e.g. telephone sets with receive volume control. Recommendation G.121 interprets those results for transmission planning purposes.

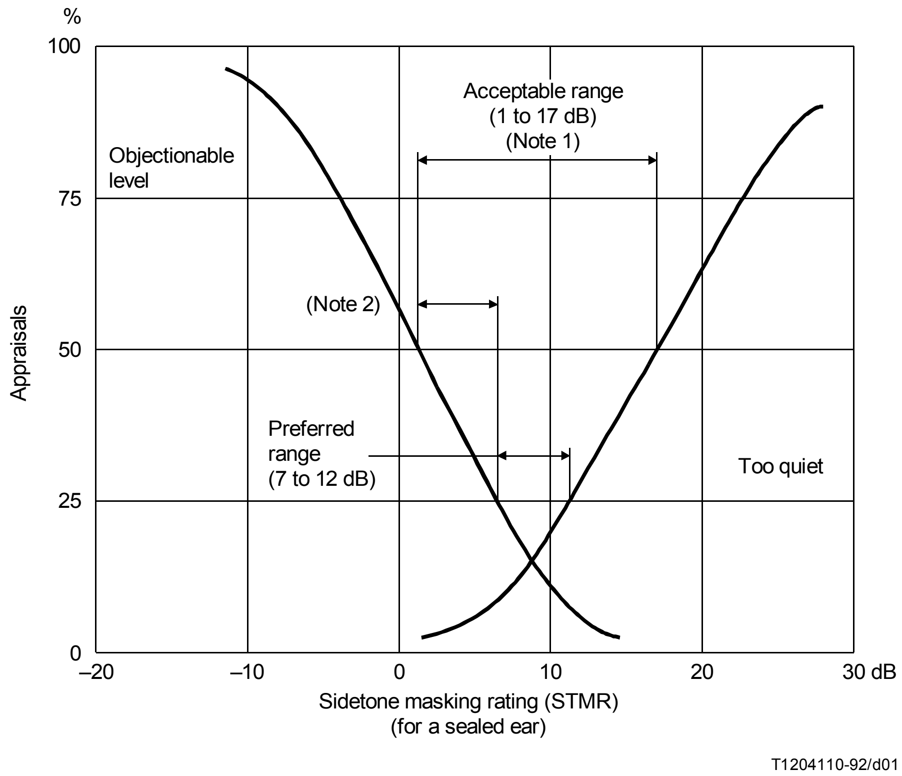
Figure 2 shows the way in which the talking level changes with sidetone level [1], [2], [3], [4]. These results were obtained by means of conversation tests [6], for a connection close to the preferred overall loss. The speech voltage will also be a function of room noise for the same connection conditions.

3 Listener sidetone

High room noise in the subscriber's environment disturbs the received speech in two ways:

i) noise being picked up by the handset microphone and transmitted to the handset receiver via the electric sidetone path;

ii) noise leaking past the earcap at the handset receiver.



NOTES

1 Conversational conditions will determine what part of this range is acceptable for a given connection.

2 This part of the acceptable range (1 to 7 dB) should only be entered with caution, e.g. on low loss connections, (see Recommendation G.121) or where there is a receive volume control.

FIGURE 1

Curves showing sidetone levels that are objectionable and too quiet, together with the preferred range, for the subscriber as a talker

Studies have shown that at low frequencies the earcap leakage path dominates over the electric sidetone path in much the same way as the human sidetone signal does in talker sidetone. The weightings applied in the STMR loudness calculation are therefore applicable and the listener sidetone rating (LSTR, Recommendation P.76) has been developed, which makes use of the room noise sidetone sensitivity (see 9/P.64) in the STMR rating method (see Recommendation P.79).

Results of subjective tests from two Administrations [7], [8] (using in this case a mean opinion scale of 0-10) are given in Figure 3. In each case the LSTR was derived by making use of Δ_{Sm} [see Recommendations P.10, P.64, P.79 and the *Handbook on Telephonometry*, 3.3.17c)] to convert the sidetone sensitivities S_{meST} to S_{RNST} before calculating LSTR (Australian results) or applied as a weighted correction to STMR (Swedish results) as described in A.4.3.3/G.111. Room noise levels were comparable at 55-59 dBA.

Based upon these results Recommendation G.121 recommends that a value of 13 dB LSTR should be striven for.

The value 13 dB is based on a 10 dB LSTR (which may be considered a minimum value), where no further improvement in mean opinion score was possible by increasing LSTR (see Figure 3), plus an allowance of 3 dB reflecting the fact that room noise in some office locations can exceed the values used in these experiments. Other tests (Sweden) have also suggested that a higher figure might be more appropriate.

The value that is satisfactory in a given telephone connection will depend on such factors as the level of room noise, the OLR of the connection, the talking levels used, etc. This is still under study.

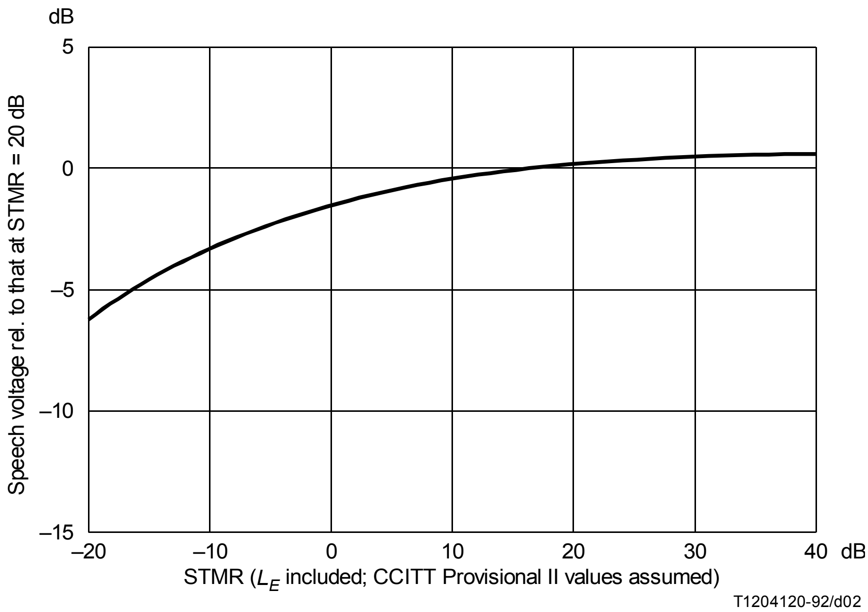


FIGURE 2

Speech voltage as a function of STMR

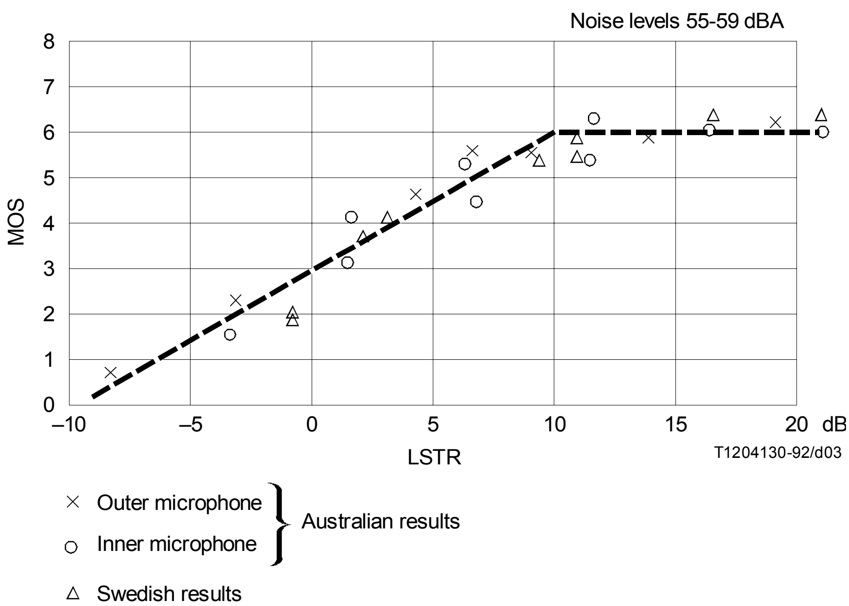


FIGURE 3

MOS as a function of LSTR calculated from Australian and Swedish test results

4 Relationship between talker and listener sidetone

4.1 Telephones having linear sidetone characteristics

For telephones having linear gain or loss characteristics in the sidetone path, the relationship between talker and listener sidetone levels is controlled by the geometry of the handset. There are two aspects of the geometry that seem most important: the distance from the mouth to the transmitter port and the size of the obstacle created by the transmitter end.

For speech inputs, a handset having a large transmitter end positioned close to the mouth experiences a greater sound pressure at its transmitter port than a handset having the transmitter end positioned farther from the mouth (distance effect) or one having a small transmitter end (obstacle effect). However, for diffuse field room noise inputs, the sound pressure at the transmitter port is independent of the size and shape of the handset. Thus, if the STMR level is the same for the two handsets, the one with the large transmitter end close to the mouth will have less electrical gain in its sidetone path, which will result in a greater LSTR value.

One Administration [9] has shown that the difference in LSTR and STMR levels for a sample of 26 linear telephone sets is highly correlated to the logarithm of the distance between the transmitter port (centre of the external opening for the microphone on the surface of the handset) and the centre of the lip ring of the artificial mouth when the handset is placed in the LRGP test position (see Recommendation P.64). It has the following empirical relationship:

$$\mathrm{LSTR} - \mathrm{STMR} = 33 - 20 \log(d)$$

where the distance d from the transmitter port to the centre of the lip ring is measured in millimetres. There may be small perturbations on the order of +/–1 dB about this relationship depending on the obstacle size presented by the transmitter end of the handset.

NOTE – This relationship is based on measurements of telephones having more or less conventional handsets. It may not be applicable for handsets with extreme geometries or for operator headsets that have their transmitter ports located behind the lip plane.

4.2 Telephones having non-linear sidetone characteristics

Non-linear gain or loss characteristics may be used in the electrical sidetone path to increase the LSTR-STMR difference. Carbon transmitters, for example, frequently are less sensitive to the lower input levels of room noise than they are to the higher input levels of speech. Such a characteristic may be introduced into telephones having linear microphones through the use of various non-linear gain circuits.

If the same non-linear gain function is used in both the send and sidetone paths of the telephone, then the LSTR-STMR difference may be approximated by measuring the difference in send sensitivities due to speech and room noise inputs, DELSM, as described in Recommendation P.64. An STMR difference may then be calculated according to the method given in Appendix I/G.111. However, if the send and sidetone paths do not have the same non-linear gain characteristics (e.g. automatic gain control circuitry in the receive path that affects sidetone), then the DELSM method will give erroneous results. In this case, the LSTR and STMR values must be measured directly.

5 Short delay talker echo perceived as sidetone

Talker echo can have a detrimental effect on transmission quality at delay times of a few milliseconds, even though the delay is not long enough for it to be perceived as an echo signal separate from the sidetone. Such echoes can occur, for example, due to reflections from the analogue trunk port of a digital PBX or on local analogue calls through a digital exchange. Unless the hybrid that converts the 4-wire digital PBX or exchange back to a 2-wire analogue circuit is well matched, some reflections will occur. Because of the digital processing times involved, these talker echo signals have a few milliseconds of delay. Sidetone provides a beneficial masking of low-level short delay talker echo, but as the talker echo level increases, it interacts with the sidetone in an unpleasant manner (hollow sounding sidetone, rain-barrel effect, etc.).

The objectively measurable effect of short delay talker echo is that it produces ripples in the sidetone frequency response. The reflected talker echo signal is added to the direct sidetone signal with a phase relationship that increases the signal at some frequencies and decreases it at others. The spacing between the ripples is equal to the reciprocal of the delay. When the reflected talker echo signal is small relative to the direct sidetone, the ripples are small. As the talker echo signal increases in magnitude, the ripples increase in size until the peaks are 6 dB above the in-phase signal and the troughs are very deep due to almost exact out-of-phase cancellation. At even higher talker echo levels (or lower sidetone levels) the amount of ripple again decreases, but the predominant signal is then the delayed talker echo.

People perceive short delay talker echo combined with sidetone differently than an equivalent level of pure sidetone, even though they may not be able to detect that a separate echo signal is present. Thus, a simple sidetone measure such as STMR is not adequate to describe the effect of the combined signal. Talker echo, even with very short delay times, must be treated as a separate impairment to transmission quality. Recommendation P.11 and Supplement 3 to Series P Recommendations provide some guidance as to how both sidetone and talker echo may be taken into account in predicting the quality of a telephone connection, but this subject remains under study.

References

[1] CCITT Contribution COM XII-No. 50, Study Period 1977-1980 (ITT).

[2] CCITT Contribution COM XII-No. 171, Study Period 1977-1980.

[3] CCITT Contribution COM XII-No. 199, Study Period 1977-1980 (Australia).

[4] CCITT Contribution COM XII-No. 116, Study Period 1977-1980 (Hungary).

[5] CCITT Contribution COM XII-No. 152, Study Period 1981-1984 (NTT).

[6] Results of conversation tests sent directly to Special Rapporteur for Question 9/XII, British Telecom, 1978.

[7] CCITT Contribution COM XII-No.151, Study Period 1981-1984 (Australia).

[8] CCITT Contribution COM XII-No.70, Study Period 1985-1988 (Sweden).

[9] CCITT Contribution COM XII-106, Study Period 1989-1992 (Sweden).

Printed in Switzerland
Geneva, 1994