INTERNATIONAL STANDARD

IEC 60268-16

Third edition
2003-05

Sound system equipment –

Part 16:
Objective rating of speech intelligibility by speech transmission index

Equipements pour systèmes électroacoustiques –

Partie 16:
Evaluation objective de l'intelligibilité de la parole au moyen de l'indice de transmission de la parole

Reference number
IEC 60268-16:2003(E)

Publication numbering

As from 1 January 1997 all IEC publications are issued with a designation in the 60000 series. For example, IEC 34-1 is now referred to as IEC 60034-1.

Consolidated editions

The IEC is now publishing consolidated versions of its publications. For example, edition numbers 1.0, 1.1 and 1.2 refer, respectively, to the base publication, the base publication incorporating amendment 1 and the base publication incorporating amendments 1 and 2.

Further information on IEC publications

The technical content of IEC publications is kept under constant review by the IEC, thus ensuring that the content reflects current technology. Information relating to this publication, including its validity, is available in the IEC Catalogue of publications (see below) in addition to new editions, amendments and corrigenda. Information on the subjects under consideration and work in progress undertaken by the technical committee which has prepared this publication, as well as the list of publications issued, is also available from the following:

- **IEC Web Site (www.iec.ch)**

- **Catalogue of IEC publications**

  The on-line catalogue on the IEC web site (http://www.iec.ch/searchpub/cur_fut.htm) enables you to search by a variety of criteria including text searches, technical committees and date of publication. On-line information is also available on recently issued publications, withdrawn and replaced publications, as well as corrigenda.

- **IEC Just Published**

  This summary of recently issued publications (http://www.iec.ch/online_news/justpub/jp_entry.htm) is also available by email. Please contact the Customer Service Centre (see below) for further information.

- **Customer Service Centre**

  If you have any questions regarding this publication or need further assistance, please contact the Customer Service Centre:

  Email: custserv@iec.ch
  Tel: +41 22 919 02 11
  Fax: +41 22 919 03 00

INTERNATIONAL STANDARD

IEC 60268-16

Third edition
2003-05

Sound system equipment –

Part 16:
Objective rating of speech intelligibility by speech transmission index

Equipements pour systèmes électroacoustiques –

Partie 16:
Evaluation objective de l'intelligibilité de la parole au moyen de l'indice de transmission de la parole

International Electrotechnical Commission, 3, rue de Varembé, PO Box 131, CH-1211 Geneva 20, Switzerland
Telephone: +41 22 919 02 11 Telefax: +41 22 919 03 00 E-mail: inmail@iec.ch Web: www.iec.ch

Commission Electrotechnique Internationale
International Electrotechnical Commission
Международная Электротехническая Комиссия

PRICE CODE U

For price, see current catalogue

CONTENTS

INTERNATIONAL ELECTROTECHNICAL COMMISSION

____

SOUND SYSTEM EQUIPMENT –

Part 16: Objective rating of speech intelligibility by speech transmission index

FOREWORD

1) The IEC (International Electrotechnical Commission) is a worldwide organization for standardization comprising all national electrotechnical committees (IEC National Committees). The object of the IEC is to promote international co-operation on all questions concerning standardization in the electrical and electronic fields. To this end and in addition to other activities, the IEC publishes International Standards. Their preparation is entrusted to technical committees; any IEC National Committee interested in the subject dealt with may participate in this preparatory work. International, governmental and non-governmental organizations liaising with the IEC also participate in this preparation. The IEC collaborates closely with the International Organization for Standardization (ISO) in accordance with conditions determined by agreement between the two organizations.
2) The formal decisions or agreements of the IEC on technical matters express, as nearly as possible, an international consensus of opinion on the relevant subjects since each technical committee has representation from all interested National Committees.
3) The documents produced have the form of recommendations for international use and are published in the form of standards, technical specifications, technical reports or guides and they are accepted by the National Committees in that sense.
4) In order to promote international unification, IEC National Committees undertake to apply IEC International Standards transparently to the maximum extent possible in their national and regional standards. Any divergence between the IEC Standard and the corresponding national or regional standard shall be clearly indicated in the latter.
5) The IEC provides no marking procedure to indicate its approval and cannot be rendered responsible for any equipment declared to be in conformity with one of its standards.
6) Attention is drawn to the possibility that some of the elements of this International Standard may be the subject of patent rights. The IEC shall not be held responsible for identifying any or all such patent rights.

International Standard IEC 60268-16 has been prepared by IEC technical committee 100: Audio, video and multimedia systems and equipment.

This third edition cancels and replaces the second edition, published in 1998. This third edition constitutes a technical revision.

The text of this standard is based on the following documents:

FDIS	Report on voting
100/650/FDIS	100/677/RVD

Full information on the voting for the approval of this standard can be found in the report on voting indicated in the above table.

This publication has been drafted in accordance with the ISO/IEC Directives, Part 2.

The committee has decided that the contents of this publication will remain unchanged until 2005. At this date, the publication will be

- reconfirmed;
- withdrawn;
- replaced by a revised edition, or
- amended.

A bilingual edition of this standard may be issued at a later date

SOUND SYSTEM EQUIPMENT –

Part 16: Objective rating of speech intelligibility by speech transmission index

1 Scope

This part of IEC 60268 defines objective methods for rating the transmission quality of speech with respect to intelligibility. The four methods, which are closely related, are referred to as the "STI," the "STITEL", the "STIPA" and the "RASTI" methods (see Clause 3). The methods are intended for rating speech transmission with or without sound systems.

A survey of other methods of determining or predicting speech intelligibility is also included, together with a method of correlating the results of different methods of determination.

2 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 4870:1991, *Acoustics – The construction and calibration of speech intelligibility tests*

ITU-T Recommendation P.51:1996, *Artificial mouth*

3 Definitions and abbreviations

For the purpose of this document, the following definitions apply.

3.1
speech transmission index (STI)
physical quantity representing the transmission quality of speech with respect to intelligibility

3.2
speech transmission index for telecommunication systems (STITEL)
index obtained by a condensed version of the STI method but still responsive to distortions found in communication systems

3.3
speech transmission index for public address systems (STIPA)
index obtained by a condensed version of the STI method but still responsive to distortions found in room acoustics including public address systems

3.4
room acoustics speech transmission index (RASTI)
index obtained by a condensed version of the STI method, to be used for screening purposes and focused on direct communication between persons without making use of a communication system. RASTI accounts for noise interference and distortions in the time domain (echoes, reverberation)

4 Description of the methods

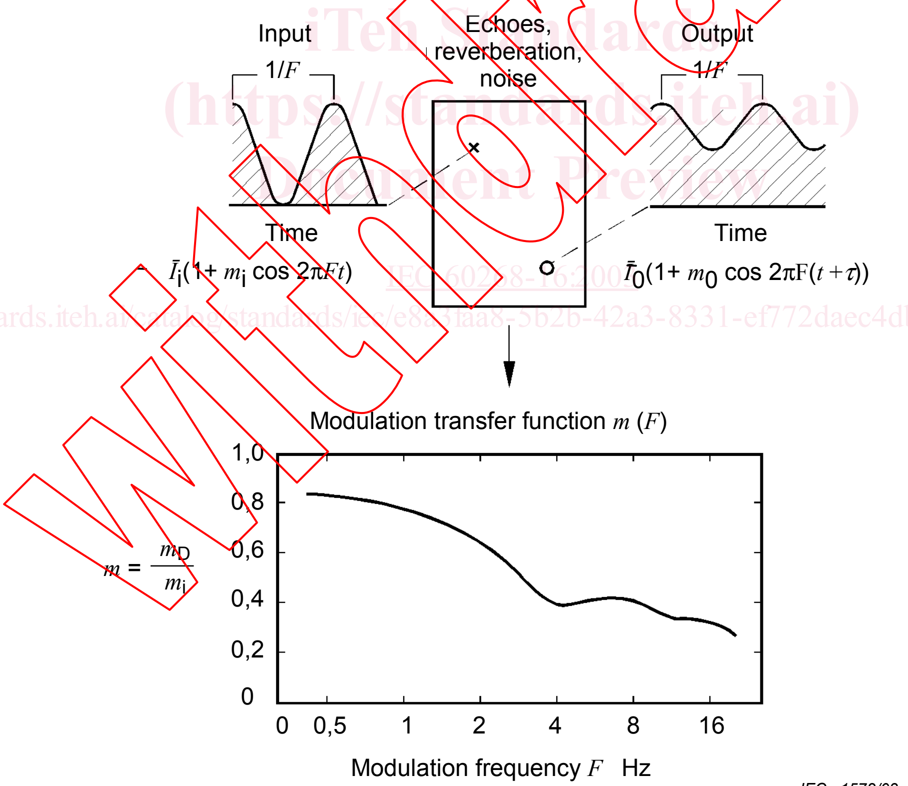
4.1 General

The methods can be used to compare speech transmission quality at various positions and for various conditions within the same listening space, in particular for assessing the effect of changes in the acoustic properties. This includes effects from the presence of an audience or of changes in any sound system [1][1]).The methods are also able to predict the absolute rating of the speech transmission quality with respect to intelligibility when comparing different listening spaces under similar conditions or assessing a speech communication channel. Annex A provides a more detailed description of the basis of the speech transmission index.

The determination of the transmission quality of speech with respect to intelligibility is based on the reduction of the modulation index m_i of a test signal, simulating the speech characteristics of a real talker, when sounded in a room or through a communication channel. The test signal is transmitted by a sound source situated at the talker's position to a microphone at any listener's position, where the modulation index is m_o.

For the sound source, the important characteristics are the physical size, the directivity, the position and the sound pressure level.

The typical test signal consists of a noise carrier with a speech-shaped frequency spectrum and a sinusoidal intensity modulation with modulation frequency F (see Figure 1).



NOTE m_i and m_o are the modulation indices of the input and the output signals, respectively. $\bar{I}_i$ and $\bar{I}_o$ are the input and output intensities, the intensities being equal to the square of the sound pressure levels (p^2).

Figure 1 – Modulation transfer function: input/output comparison

1) Figures in square brackets refer to the bibliography.

The reduction in the modulation index is quantified by the modulation transfer function *m(F)* which is determined by

$$m(F)=\frac{m_o}{m_i}$$

and is interpreted in terms of an apparent signal-to-noise ratio (*SNR*), irrespective of the cause of the reduction which can be reverberation, echoes, non-linear distortion components or interfering noise, determined by

$$SNR_{\text{App}} = 10\lg\left(\frac{m(F)}{1-m(F)}\right)$$

The values of the apparent signal-to-noise ratio are limited to the range ±15 dB. Values less than –15 dB are given the value of –15 dB and values greater than 15 dB are given the value of 15 dB.

4.2 The STI method

4.2.1 General

The STI method, described in Annex A, is based on the determination of the modulation transfer function *m(F)* for 98 data points, obtained for 14 modulation frequencies at one-third octave intervals ranging from 0,63 Hz up to and including 12,5 Hz and for seven octave bands with centre frequencies ranging from 125 Hz up to and including 8 kHz (see Figure A.3).

4.2.2 Precision of the STI method

Because the test signal is band-limited random or pseudo-random noise, repetition of measurement does not normally produce identical results, even under conditions of steady interference. The results centre on a mean with a certain standard deviation. This depends, amongst other factors, on the number of discrete measurements of the modulation transfer function (usually 98 for the STI method) and the measuring time involved. Typically, the value of the standard deviation is about 0,02 for a measuring time of 10 s for each *m(F)* and with stationary noise interference. With fluctuating noise (for example, a babble of voices), higher standard deviations may be found possibly with a systematic error. This can be checked by carrying out a measurement in the absence of the test signal. This should result in a residual STI value less than 0,20. An estimate of the standard deviation should be made by repeating measurements for at least a restricted set of conditions.

4.2.3 Limitations of the STI method

Due to the form of the test signals and the analysis, the types of distortion not accounted for are frequency shifts (such as those found with devices for preventing acoustic feedback and with single sideband radio transmissions), frequency multiplication (for example, analogue tape recordings played at incorrect speed) and systems such as vocoders that encode speech fragments (for example, linear predictive coding which might use code-book related synthesis or the introduction of errors related to voiced/unvoiced speech fragments and pitch errors).

The method should not be used for transmission channels

a) which introduce frequency shifts or frequency multiplication, or
b) which include vocoders (i.e. linear predictive speech coder (LPC), code-excited linear predictive coder (CELP), residually excited linear predictive coder (RELP), etc.).

Without specific corrections, the STI method is not a reliable prediction measure of the intelligibility of speech for hearing-impaired listeners [17] or to the wearers of ear defenders.

4.3 The STITEL method

4.3.1 General

A simplification can be applied to the test signal if the uncorrelated (speech-like) modulations required for the correct interpretation of non-linear distortions, are omitted. This opens up the possibility of modulating and parallel processing all seven frequency bands simultaneously, thus reducing measuring time. The STITEL method, described in Annex B, employs this simplification and takes 10 s to 15 s for a measurement.

4.3.2 Precision of the STITEL method

As with the STI method (see 4.2.2), results are mean values with a certain standard deviation, due to the randomness of noise. The standard deviation depends on the number of discrete measurements of the modulation transfer function (typically seven for the STITEL method) and the measuring time involved. The standard deviation should be estimated by performing repeated measurements, at least for a restricted number of conditions.

4.3.3 Limitations of the STITEL method

The STITEL method should not be used for transmission channels

a) which introduce frequency shifts or frequency multiplication;

b) which include vocoders (i.e. LPC, CELP, RELP, etc.);

c) which introduce strong non-linear distortion components;

d) for which reverberation time is strongly frequency-dependent. Over the range of centre frequencies 125 Hz to 8 kHz, the uniformities of the octave-band early decay times and signal-to-noise ratios should fall within the permitted area shown in Figure 3;

e) having echoes stronger than –10 dB referred to the primary signal;

f) if the background noise has audible tones and/or marked peaks or troughs in the octave-band spectrum;

g) if the background noise is impulsive and/or the space is not substantially free of discrete echoes, particularly flutter echoes whose repetition frequency is an integral multiple of one or more of the modulation frequencies [2].

If c), d), or e) or all three apply, or possibly apply, the STI method should be used instead, or used to verify the results obtained by the STITEL method.