

**NORME
INTERNATIONALE
INTERNATIONAL
STANDARD**

**CEI
IEC**

60268-16

Deuxième édition
Second edition
1998-03

Equipements pour systèmes électroacoustiques –

Partie 16:

**Evaluation objective de l'intelligibilité
de la parole au moyen de l'indice
de transmission de la parole**

Sound system equipment –

Part 16:

**Objective rating of speech intelligibility
by speech transmission index**



Numéro de référence
Reference number
CEI/IEC 60268-16:1998

Numéros des publications

Depuis le 1er janvier 1997, les publications de la CEI sont numérotées à partir de 60000.

Publications consolidées

Les versions consolidées de certaines publications de la CEI incorporant les amendements sont disponibles. Par exemple, les numéros d'édition 1.0, 1.1 et 1.2 indiquent respectivement la publication de base, la publication de base incorporant l'amendement 1, et la publication de base incorporant les amendements 1 et 2.

Validité de la présente publication

Le contenu technique des publications de la CEI est constamment revu par la CEI afin qu'il reflète l'état actuel de la technique.

Des renseignements relatifs à la date de reconfirmation de la publication sont disponibles dans le Catalogue de la CEI.

Les renseignements relatifs à ces révisions, à l'établissement des éditions révisées et aux amendements peuvent être obtenus auprès des Comités nationaux de la CEI et dans les documents ci-dessous:

- **Bulletin de la CEI**
- **Annuaire de la CEI**
Accès en ligne*
- **Catalogue des publications de la CEI**
Publié annuellement et mis à jour régulièrement (Accès en ligne)*

Terminologie, symboles graphiques et littéraux

En ce qui concerne la terminologie générale, le lecteur se reportera à la CEI 60050: *Vocabulaire Electrotechnique International* (VEI).

Pour les symboles graphiques, les symboles littéraux et les signes d'usage général approuvés par la CEI, le lecteur consultera la CEI 60027: *Symboles littéraux à utiliser en électrotechnique*, la CEI 60417: *Symboles graphiques utilisables sur le matériel. Index, relevé et compilation des feuilles individuelles*, et la CEI 60617: *Symboles graphiques pour schémas*.

Publications de la CEI établies par le même comité d'études

L'attention du lecteur est attirée sur les listes figurant à la fin de cette publication, qui énumèrent les publications de la CEI préparées par le comité d'études qui a établi la présente publication.

* Voir adresse «site web» sur la page de titre.

Numbering

As from 1 January 1997 all IEC publications are issued with a designation in the 60000 series.

Consolidated publications

Consolidated versions of some IEC publications including amendments are available. 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.

Validity of this publication

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 the date of the reconfirmation of the publication is available in the IEC catalogue.

Information on the revision work, the issue of revised editions and amendments may be obtained from IEC National Committees and from the following IEC sources:

- **IEC Bulletin**
- **IEC Yearbook**
On-line access*
- **Catalogue of IEC publications**
Published yearly with regular updates (On-line access)*

Terminology, graphical and letter symbols

For general terminology, readers are referred to IEC 60050: *International Electrotechnical Vocabulary* (IEV).

For graphical symbols, and letter symbols and signs approved by the IEC for general use, readers are referred to publications IEC 60027: *Letter symbols to be used in electrical technology*, IEC 60417: *Graphical symbols for use on equipment. Index, survey and compilation of the single sheets* and IEC 60617: *Graphical symbols for diagrams*.

IEC publications prepared by the same technical committee

The attention of readers is drawn to the end pages of this publication which list the IEC publications issued by the technical committee which has prepared the present publication.

* See web site address on title page.

**NORME
INTERNATIONALE
INTERNATIONAL
STANDARD**

**CEI
IEC**

60268-16

Deuxième édition
Second edition
1998-03

Equipements pour systèmes électroacoustiques –

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

Sound system equipment –

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

© IEC 1998 Droits de reproduction réservés — Copyright - all rights reserved

Aucune partie de cette publication ne peut être reproduite ni utilisée sous quelque forme que ce soit et par aucun procédé, électronique ou mécanique, y compris la photocopie et les microfilms, sans l'accord écrit de l'éditeur.

No part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from the publisher.

International Electrotechnical Commission
Telefax: +41 22 919 0300

3, rue de Varembe Geneva, Switzerland
e-mail: inmail@iec.ch IEC web site <http://www.iec.ch>



Commission Electrotechnique Internationale
International Electrotechnical Commission
Международная Электротехническая Комиссия

CODE PRIX
PRICE CODE

U

Pour prix, voir catalogue en vigueur
For price, see current catalogue

SOMMAIRE

	Pages
AVANT-PROPOS	4
Articles	
1 Domaine d'application	6
2 Références normatives	6
3 Définitions	6
4 Descriptions des méthodes	8
4.1 Généralités	8
4.2 La méthode STI	10
4.2.1 Généralités	10
4.2.2 Précision de la méthode STI	10
4.2.3 Limitation de la méthode STI	10
4.3 La méthode STITEL	12
4.3.1 Généralités	12
4.3.2 Précision de la méthode STITEL	12
4.3.3 Limitation de la méthode STITEL	12
4.4 La méthode RASTI	12
4.4.1 Généralités	12
4.4.2 Précision de la méthode RASTI	14
4.4.3 Limitations de la méthode RASTI	14
4.5 Méthodes de mesure	16
4.5.1 Méthode de mesure utilisant un signal d'excitation acoustique	16
4.5.2 Méthode de mesure utilisant l'injection (électrique) directe du signal d'essai	18
4.5.3 Simulation du bruit d'occupation	18
4.5.4 Essais d'intelligibilité utilisant un équipement d'analyse de séquence de type longueur maximale	18
4.5.5 Simulation du bruit d'occupation	18
4.5.6 Répétition des mesures	20
4.5.7 Analyse et interprétation des résultats	20
5 Autres méthodes pour déterminer l'intelligibilité	20
5.1 Partitions de mots phonétiquement équilibrées (liste de 256 ou 1 000)	20
5.2 Essais de rimes modifiées	20
5.3 Indice d'articulation	20
5.4 Perte d'articulation des consonnes	22
Annexe A (normative) Méthode de l'indice de transmission de la parole (STI) et méthode révisée (STI _r)	24
Annexe B (informative) La méthode STITEL	42
Annexe C (informative) La méthode RASTI	44
Annexe D (informative) Estimation de la correction dans le cas de fortes réflexions	46
Annexe E (informative) Bibliographie	50

CONTENTS

	Page
FOREWORD	5
Clause	
1 Scope	7
2 Normative references	7
3 Definitions and abbreviations	7
4 Descriptions of the methods	9
4.1 General	9
4.2 The STI method	11
4.2.1 General	11
4.2.2 Precision of the STI method	11
4.2.3 Limitations of the STI method	11
4.3 The STITEL method	13
4.3.1 General	13
4.3.2 Precision of the STITEL method	13
4.3.3 Limitations of the STITEL method	13
4.4 The RASTI method	13
4.4.1 General	13
4.4.2 Precision of the RASTI method	15
4.4.3 Limitations of the RASTI method	15
4.5 Methods of measurement	17
4.5.1 Method of measurement using an acoustic excitation signal	17
4.5.2 Method of measurement using direct (electrical) injection of the test signal	19
4.5.3 Simulation of occupancy noise	19
4.5.4 Intelligibility tests using maximum length sequence (MLS) analysis equipment	19
4.5.5 Simulation of occupancy noise	19
4.5.6 Repetition of measurements	21
4.5.7 Analysis and interpretation of the results	21
5 Other methods of determining intelligibility	21
5.1 Phonetically balanced word scores (256 or 1 000 population)	21
5.2 Modified rhyme tests	21
5.3 Articulation index	21
5.4 Articulation loss of consonants	23
Annex A (normative) Speech transmission index (STI) and revised (STI _r) methods	25
Annex B (informative) The STITEL method	43
Annex C (informative) The RASTI method	45
Annex D (informative) Estimating the correction in the case of strong reflections	47
Annex E (informative) Bibliography	51

COMMISSION ÉLECTROTECHNIQUE INTERNATIONALE

ÉQUIPEMENTS POUR SYSTÈMES ÉLECTROACOUSTIQUES –

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

AVANT-PROPOS

- 1) La CEI (Commission Electrotechnique Internationale) est une organisation mondiale de normalisation composée de l'ensemble des comités électrotechniques nationaux (Comités nationaux de la CEI). La CEI a pour objet de favoriser la coopération internationale pour toutes les questions de normalisation dans les domaines de l'électricité et de l'électronique. A cet effet, la CEI, entre autres activités, publie des Normes internationales. Leur élaboration est confiée à des comités d'études, aux travaux desquels tout Comité national intéressé par le sujet traité peut participer. Les organisations internationales, gouvernementales et non gouvernementales, en liaison avec la CEI, participent également aux travaux. La CEI collabore étroitement avec l'Organisation internationale de Normalisation (ISO), selon des conditions fixées par accord entre les deux organisations.
- 2) Les décisions ou accords officiels de la CEI concernant des questions techniques, représentent, dans la mesure du possible, un accord international sur les sujets étudiés, étant donné que les Comités nationaux intéressés sont représentés dans chaque comité d'études.
- 3) Les documents produits se représentent sous la forme de recommandations internationales; ils sont publiés comme normes, rapports techniques ou guides et agréés comme tels par les Comités nationaux.
- 4) Dans le but d'encourager l'unification internationale, les Comités nationaux de la CEI s'engagent à appliquer de façon transparente, dans toute la mesure possible, les Normes internationales de la CEI dans leurs normes nationales et régionales. Toute divergence entre la norme de la CEI et la norme nationale ou régionale correspondante doit être indiquée en termes clairs dans cette dernière.
- 5) La CEI n'a fixé aucune procédure concernant le marquage comme indication d'approbation et sa responsabilité n'est pas engagée quand un matériel est déclaré conforme à l'une de ses normes.
- 6) L'attention est attirée sur le fait que certains des éléments de la présente Norme internationale peuvent faire l'objet de droits de propriété intellectuelle ou de droits analogues. La CEI ne saurait être tenue pour responsable de ne pas avoir identifié de tels droits de propriété et de ne pas avoir signalé leur existence.

La Norme internationale CEI 60268-16 a été établie par le sous-comité 100C: Equipements et systèmes dans le domaine des techniques audio, vidéo et multimédia, du comité d'études 100 de la CEI: Systèmes et appareils audio, vidéo et multimédia.

Cette deuxième édition annule et remplace la première édition parue en 1988. Cette édition constitue une révision technique.

Le texte de cette norme est issu des documents suivants:

FDIS	Rapport de vote
100C/189/FDIS	100C/218/RVD

Le rapport de vote indiqué dans le tableau ci-dessus donne toute information sur le vote ayant abouti à l'approbation de cette norme.

L'annexe A fait partie intégrante de cette norme.

Les annexes B, C, D et E sont données uniquement à titre d'information.

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 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 subcommittee 100C: Audio, video and multimedia subsystems and equipment, of IEC technical committee 100: Audio, video and multimedia systems and equipment.

This second edition cancels and replaces the first edition, published in 1988. This second edition constitutes a technical revision.

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

FDIS	Report on voting
100C/189/FDIS	100C/218/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.

Annex A forms an integral part of this standard.

Annexes B, C, D and E are for information only.

ÉQUIPEMENTS POUR SYSTÈMES ÉLECTROACOUSTIQUES –

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

1 Domaine d'application

La présente partie de la CEI 60268 concerne des méthodes objectives d'évaluation de la qualité de transmission de la parole sous l'angle de l'intelligibilité. Les trois méthodes existantes, étroitement liées, sont connues sous la désignation de méthodes "STI", "STITEL" et "RASTI" (voir article 3). Ces méthodes sont destinées à l'évaluation de la transmission de la parole avec ou sans systèmes de sonorisation.

Il comporte également une étude d'autres méthodes permettant de déterminer l'intelligibilité de la parole, ainsi qu'une méthode permettant de mettre en corrélation les résultats des différentes méthodes d'évaluation.

2 Références normatives

Les documents normatifs suivants contiennent des dispositions qui, par suite de la référence qui y est faite, constituent des dispositions valables pour la présente partie de la CEI 60268. Au moment de la publication, les éditions indiquées étaient en vigueur. Tout document normatif est sujet à révision, et les parties prenantes aux accords fondés sur la présente partie de la CEI 60268 sont invitées à rechercher la possibilité d'appliquer les éditions les plus récentes des documents normatifs indiqués ci-après. Les membres de la CEI et de l'ISO possèdent le registre des Normes internationales en vigueur.

ISO/TR 4870:1991, *Acoustique - Construction et étalonnage des essais d'intelligibilité de la parole*

UIT-T Recommandation P.50:1994, *Voix artificielles*

UIT-T Recommandation P.51:1996, *Bouche artificielle*

3 Définitions

Pour les besoins de la présente partie de la CEI 60268, les définitions suivantes s'appliquent:

3.1

indice de transmission de la parole (STI)

grandeur physique qui représente la qualité de transmission de la parole sous l'angle de l'intelligibilité

3.2

indice de transmission de la parole pour les systèmes de télécommunications (STITEL)

indice obtenu par une méthode simplifiée de la méthode STI, mais encore sensible aux distorsions apparaissant dans les systèmes de communication

3.3

indice de transmission de la parole rapide, ou indice de transmission de la parole relatif à l'acoustique de la pièce (RASTI)

indice obtenu par une méthode simplifiée de la méthode STI, qui se concentre sur les perturbations et distorsions en fonction du temps (échos, réverbération)

SOUND SYSTEM EQUIPMENT –

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

1 Scope

This part of IEC 60268 concerns objective methods for rating the transmission quality of speech with respect to intelligibility. The three methods, which are closely related are referred to as the "STI," the "STITEL" 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 speech intelligibility is also included, together with a method of correlating the results of different methods of determination.

2 Normative references

The following normative documents contain provisions which, through reference in this text, constitute provisions of this part of IEC 60268. At the time of publication, the editions indicated were valid. All normative documents are subject to revision, and parties to agreements based on this part of IEC 60268 are encouraged to investigate the possibility of applying the most recent editions of the normative documents indicated below. Members of IEC and ISO maintain registers of currently valid International Standards.

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

ITU-T Recommendation P.50:1994, *Artificial voices*

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

3 Definitions and abbreviations

For the purpose of this part of IEC 60268, the following definitions apply:

3.1

speech transmission index (STI)

a 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

rapid speech transmission index; room acoustics speech transmission index (RASTI)

index obtained by a condensed version of the STI method focused on noise interference and distortions in the time domain (echoes, reverberation)

4 Descriptions des méthodes

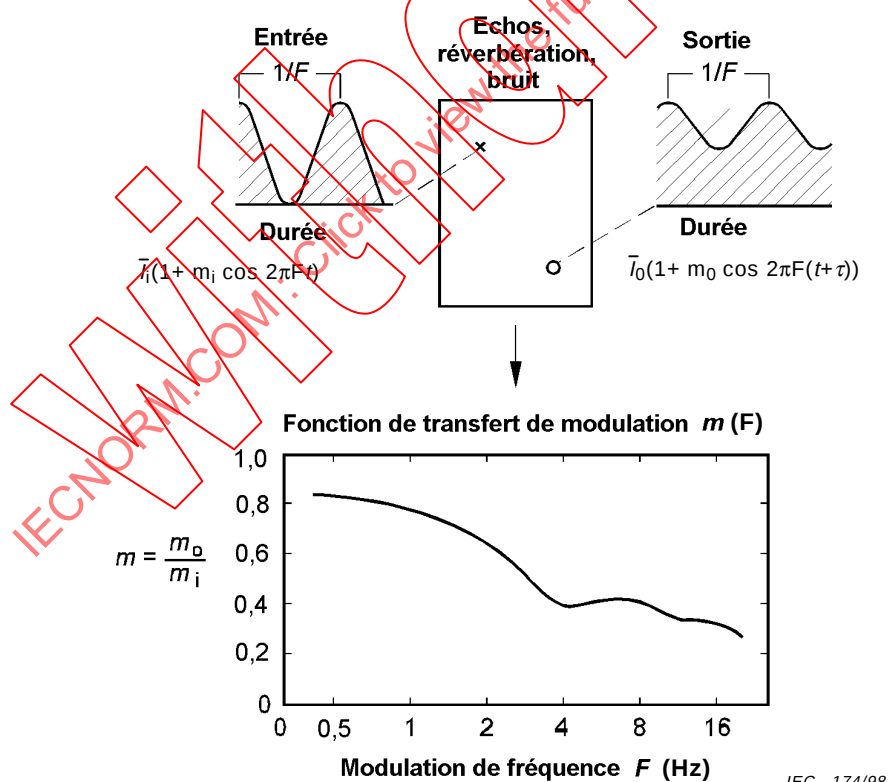
4.1 Généralités

Les méthodes peuvent être utilisées pour comparer la qualité de transmission de la parole à différents endroits et dans des conditions diverses à l'intérieur d'une même salle. On peut les utiliser en particulier pour évaluer l'effet de modifications des propriétés acoustiques, y compris les effets résultant de la présence des auditeurs ou de modifications d'un système de sonorisation [1]¹⁾. Ces méthodes peuvent également donner la valeur absolue représentative de la qualité de transmission de la parole sous l'angle de l'intelligibilité lorsque l'on compare différents espaces d'écoute soumis à des conditions semblables. L'annexe A fournit une description plus détaillée du fondement de l'indice de transmission de la parole.

La détermination de la qualité de transmission de la parole sous l'angle de l'intelligibilité est fondée sur la perte en taux de modulation m_i d'un signal d'essai simulant les caractéristiques de la parole d'un locuteur réel, lorsque ce signal est émis dans une salle. Le signal d'essai est transmis par une source sonore positionnée à la place du locuteur à un microphone situé à toute place d'auditeur, où l'indice de modulation est m_o .

Les caractéristiques essentielles de la source sonore sont les suivantes: dimensions physiques, directivité, position et niveau de pression acoustique.

Le signal d'essai type est constitué d'une porteuse de bruit dont le spectre dans le domaine fréquentiel recouvre celui de la parole, modulé en intensité par un signal sinusoïdal de fréquence F (voir figure 1).



NOTE – m_i et m_o sont respectivement les indices de modulation des signaux d'entrée et de sortie. T_i et T_o sont les intensités d'entrée et de sortie.

Figure 1 – Fonction de transfert de modulation: comparaison entrée/sortie

¹⁾ Les chiffres entre crochets font référence à la bibliographie donnée en annexe E.

4 Description of the methods

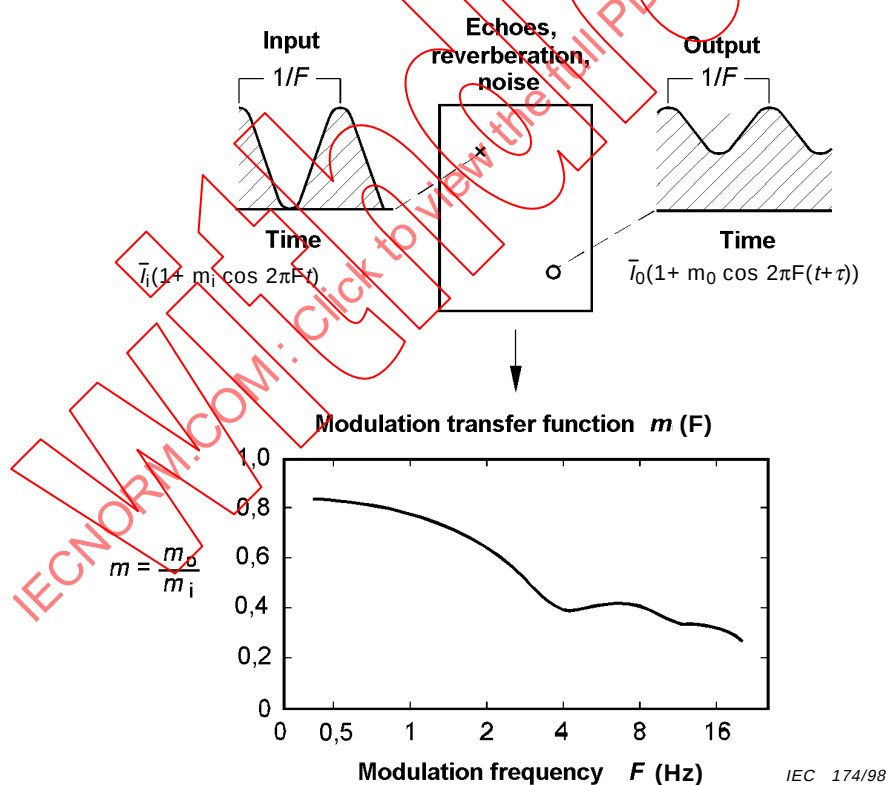
4.1 General

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 determine the absolute rating of the speech transmission quality with respect to intelligibility, when comparing different listening spaces under similar conditions. 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. 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.

Figure 1 – Modulation transfer function: input/output comparison

¹⁾ Figures in square brackets refer to the bibliography given in annex E.

La perte en taux de modulation est exprimée par la fonction de transfert de la modulation $m(F)$, qui est déterminée par:

$$m(F) = \frac{m_0}{m_i}$$

Cette perte est interprétée en termes de rapport apparent signal sur bruit (SNR), indépendamment de la cause de la réduction qui peut être la réverbération, les échos ou le bruit d'interférences déterminée par:

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

Les valeurs du rapport signal sur bruit apparent sont limitées à la gamme ± 15 dB. Les valeurs inférieures à -15 dB sont notées -15 dB et les valeurs supérieures à 15 dB sont notées 15 dB.

4.2 La méthode STI

4.2.1 Généralités

La méthode STI, décrite en annexe A, est fondée sur la détermination de la fonction de transfert de la modulation $m(F)$ pour 98 points obtenus à partir de 14 fréquences de modulation, échelonnées par tiers d'octave entre 0,63 Hz et 12,5 Hz inclus et à partir de sept bandes d'octaves dont les fréquences médianes sont échelonnées de 125 Hz à 8 kHz inclus (voir figure A.3).

4.2.2 Précision de la méthode STI

Les mesures successives ne donnent normalement pas des résultats identiques, même en présence de conditions de perturbations constantes, du fait que le signal d'essai est un bruit aléatoire ou pseudo-aléatoire de bande limitée. Les résultats sont distribués autour d'une moyenne avec un certain écart type. Ce dernier dépend, entre autres facteurs, du nombre de mesures discrètes de la fonction de transfert de la modulation (généralement 98 pour la méthode STI) et de la durée de la mesure choisie. La valeur typique de l'écart type est approximativement de 0,02 pour une durée de mesure de 10 s pour chaque $m(F)$ et avec des perturbations stationnaires. Avec un bruit aléatoire (par exemple un brouhaha de voix), des écarts types supérieurs et éventuellement une erreur systématique sont susceptibles d'être constatés. Cela peut être vérifié en effectuant une mesure en l'absence du signal d'essai. Il convient que la valeur résiduelle du STI en résultant ne soit pas supérieure à 0,20. Il convient de réaliser une estimation de l'écart type en répétant les mesures au moins pour un nombre limité de conditions.

4.2.3 Limitation de la méthode STI

Du fait de la forme des signaux d'essai et du mode d'analyse, les types de distorsions non pris en compte sont les suivants: glissements de fréquence (comme dans le cas d'appareils réducteurs d'accrochage acoustique (anti-larsen) et de transmissions radio à bande latérale unique), multiplication de fréquence (par exemple des enregistrements analogiques sur bande magnétique lus à une vitesse incorrecte) et systèmes tels que les vocodeurs, qui encodent des fragments vocaux (par exemple l'encodage linéaire prédictif qui introduit des erreurs liées à la succession de fragments vocaux/non vocaux, et à des erreurs de tonie).

Il convient de ne pas utiliser la méthode pour des canaux de transmission

- a) introduisant des glissements ou des multiplications de fréquence, ou
- b) qui comprennent des vocodeurs (c'est-à-dire vocodeurs prédictifs linéaires (LPC), codeurs prédictifs linéaires à code activé (CELP), codeurs prédictifs linéaires résiduellement activés (RELP), etc.).

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

$$m(F) = \frac{m_0}{m_i}$$

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 or interfering noise, determined by:

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

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 (e.g. 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 (e.g. analogue tape recordings played at incorrect speed) and systems such as vocoders that encode speech fragments (e.g. linear predictive coding which introduces 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 (RELPE), etc.).

4.3 La méthode STITEL

4.3.1 Généralités

Il est possible de simplifier le signal d'essai si des modulations non corrélatives (similaires à la parole), obligatoires pour l'interprétation correcte des distorsions non linéaires, sont omises. Cela permet d'envisager la modulation et le traitement parallèle de toutes les bandes de fréquences simultanément, ce qui réduit la durée de mesure. La méthode STITEL, décrite dans l'annexe B, emploie cette simplification et prend 10 s à 15 s pour une mesure.

4.3.2 Précision de la méthode STITEL

Comme avec la méthode STI (voir 4.2.2), les résultats obtenus sont des valeurs moyennes présentant un certain écart type dû à l'aspect aléatoire du bruit. L'écart type dépend du nombre de mesures discrètes de la fonction de transfert de la modulation (généralement sept pour la méthode STITEL) et de la durée de mesure choisie. Il convient que l'écart type soit évalué en effectuant des mesures répétitives, au moins pour un nombre limité de conditions.

4.3.3 Limitation de la méthode STITEL

Il convient de ne pas utiliser la méthode STITEL pour des canaux de transmission:

- a) introduisant des glissements de fréquence ou une multiplication de fréquence;
- b) comprenant des vocodeurs (c'est-à-dire LPC, CELP, RELP, etc.);
- c) pour lesquels le temps de réverbération est fortement dépendant de la fréquence;
- d) présentant des échos supérieurs à -10 dB du signal primaire;
- e) si le bruit de fond contient des sons purs perceptibles, et/ou des crêtes ou de creux marqués dans le spectre de bande d'octaves;
- f) si le bruit ambiant est de caractère impulsif et/ou que l'endroit n'est pas réellement dépourvu d'échos discrets, plus particulièrement d'échos de scintillement, dont la fréquence de répétition est un multiple entier d'une ou de plusieurs fréquences de modulation [2];
- g) si le temps de réverbération dépend très fortement de la fréquence: c'est-à-dire que, sur toute la gamme de fréquences médianes échelonnées de 125 Hz à 8 kHz, il convient que la régularité du temps d'affaiblissement précoce des bandes d'octaves ainsi que des rapports signal sur bruit reste dans les limites autorisées par la figure 3.

Si c) ou d) ou les deux s'appliquent, ou sont susceptibles de s'appliquer, il convient d'utiliser la méthode STI à la place de la méthode STITEL, ou pour vérifier les résultats obtenus par cette dernière méthode.

4.4 La méthode RASTI

4.4.1 Généralités

La diminution du nombre de bandes d'octaves peut constituer une autre simplification. C'est le cas pour la méthode RASTI, décrite dans l'annexe C, méthode selon laquelle l'analyse est limitée à deux bandes d'octaves seulement, de fréquences centrales 500 Hz et 2 kHz, et à respectivement quatre ou cinq fréquences de modulation seulement pour ces bandes. Cette méthode implique que la limitation de la bande passante et le bruit de fond présentant un spectre irrégulier ne soient pas pris en compte pour correction, pas plus d'ailleurs que l'effet de distorsion non linéaire. La méthode RASTI peut cependant être utilisée comme approche dans la plupart des cas d'applications d'acoustique de salle. Comme avec la méthode STI, on admet que certaines distorsions, en particulier celles résultant de la réverbération, soient corrigées si elles sont lisses et monotones [3].

La méthode RASTI n'est pas une mesure sûre de l'intelligibilité de la parole pour des environnements fortement perturbés par le bruit, ni de l'intelligibilité de la parole normale pour les malentendants ou les porteurs de protections d'oreille.

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 the possibility of modulating and parallel processing all frequency bands simultaneously, so 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) for which reverberation time is strongly frequency dependent;
- d) having echoes stronger than –10 dB referred to the primary signal;
- e) if the background noise has audible tones and/or marked peaks or troughs in the octave-band spectrum;
- f) 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];
- g) if the 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.

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

4.4 The RASTI method

4.4.1 General

Another simplification that can be applied is a reduction in the number of octave bands. This is the case with the RASTI method, described in annex C, in which the analysis is restricted to only two octave bands with centre frequencies 500 Hz and 2 kHz, and to only four and five modulation frequencies, respectively, in these bands. This implies that bandpass limiting and background noise with an irregular spectrum are not accounted for correctly, nor is the effect of non-linear distortion included. The RASTI method can, however, be used as a screening approach for most room acoustic applications. As with the STI method, certain distortions, particularly those from reverberation, if smooth and monotonic, are accounted for correctly [3].

The RASTI method is not a reliable measure of the intelligibility of speech uttered in high noise environments, or of normal speech to the hard of hearing or to the wearers of ear defenders.

4.4.2 Précision de la méthode RASTI

Comme pour la méthode STI (voir 4.2.2), les résultats obtenus sont des valeurs moyennes présentant un certain écart type dû à l'aspect aléatoire du bruit. L'écart type dépend, entre autres, de la durée de mesure choisie. Il convient d'évaluer l'écart type en effectuant des mesures répétées, au moins pour un nombre limité de conditions. En pratique, une durée de mesure de 10 s représente un bon compromis entre la vitesse et la précision. La figure 2 montre la précision qu'il est possible d'atteindre avec une durée de mesure de cet ordre.

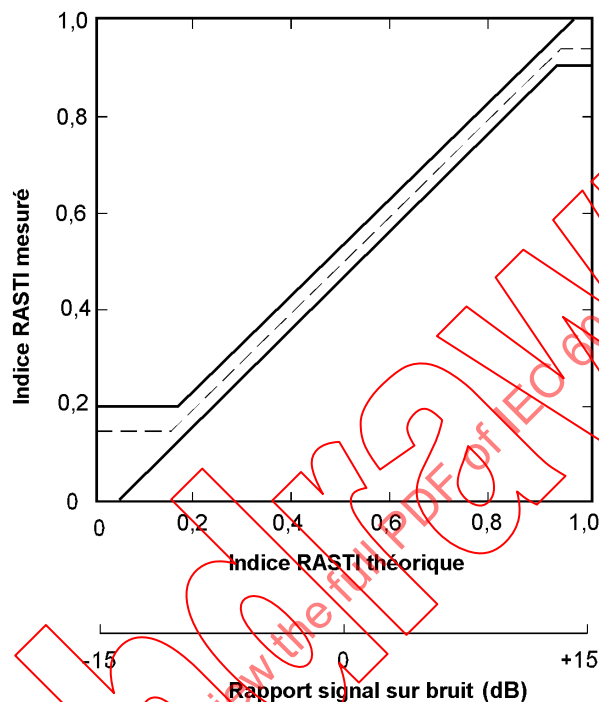


Figure 2 – Relation entre l'indice STI théorique obtenu par la méthode RASTI et l'indice STI mesuré par un équipement breveté, la durée de mesure étant approximativement de 12 s

4.4.3 Limitations de la méthode RASTI

L'application de la méthode RASTI est limitée par des facteurs liés à la transmission de la parole, au bruit ambiant et à la réverbération. Par conséquent, il convient de limiter son utilisation aux cas où les conditions suivantes sont remplies:

- absence de glissements de fréquence ou de multiplication de fréquence;
- pas d'utilisation de vocodeurs (c'est-à-dire LPC, CELP, RELP, etc.);
- transmission de parole en mode essentiellement linéaire (toute compression ou expansion de l'amplitude limitée à 1 dB) et il n'existe pas d'écrouissage d'un signal sinusoïdal donnant le même niveau de pression sonore, au même endroit de mesure, que le signal d'essai;
- la réponse en fréquence du système global entre les bandes d'octaves de fréquences centrales 125 Hz et 8 kHz est uniforme, c'est-à-dire qu'il convient que la différence du niveau de pression sonore entre deux bandes d'octaves adjacentes ne dépasse pas 5 dB;
- le bruit de fond ne contient pas de sons purs perceptibles, ni de crêtes ou de creux marqués dans le spectre des bandes d'octaves;
- le bruit de fond n'a pas de caractère impulsif et l'endroit est réellement dépourvu d'échos discrets, plus particulièrement d'échos de scintillement, dont la fréquence de répétition est un multiple entier d'une ou de plusieurs fréquences de modulation [2];
- le temps de réverbération ne doit pas dépendre trop fortement de la fréquence: c'est-à-dire que, sur toute la gamme de fréquences médianes échelonnées de 125 Hz à 8 kHz, la régularité du temps d'affaiblissement précoce des bandes d'octaves ainsi que des rapports signal sur bruit doit rester dans les limites autorisées indiquées par la figure 3;

4.4.2 Precision of the RASTI 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 upon the measuring time involved, amongst other factors. The standard deviation should be estimated by performing repeated measurements, at least for a restricted number of conditions. In practice, a measuring time of 10 s is a useful compromise between speed and accuracy. Figure 2 illustrates the accuracy obtainable with a measuring time of that order.

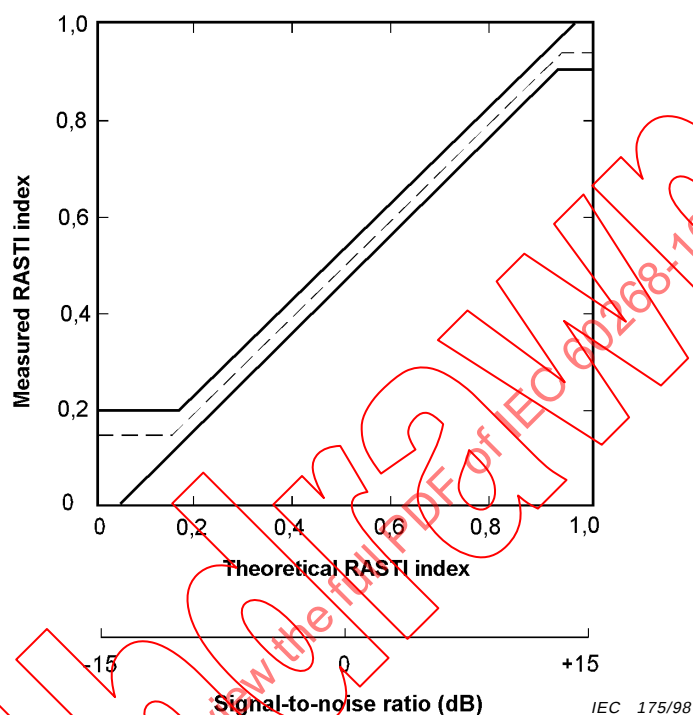
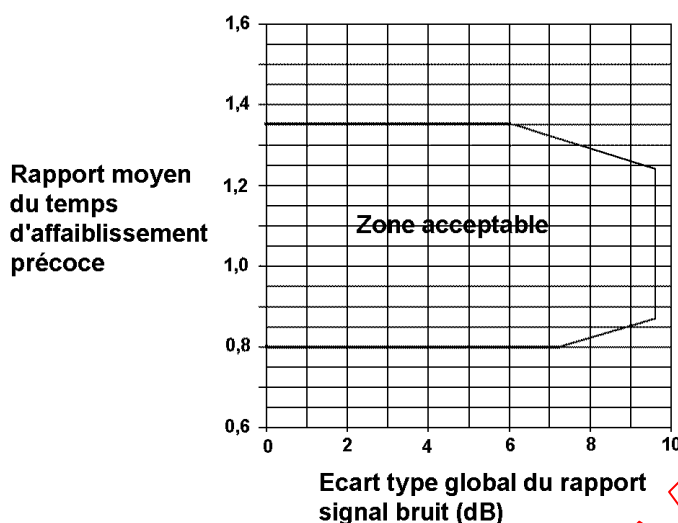


Figure 2 – Relationship between the theoretical STI by the RASTI method and the STI measured by a proprietary equipment with a measurement time of 12 s approximately

4.4.3 Limitations of the RASTI method

The application of the RASTI method is limited by factors concerned with speech transmission, background noise and reverberation. Therefore its use should be restricted to cases where the following conditions are met:

- no frequency shifts or frequency multiplication;
- no use of vocoders (i.e. LPC, CELP, RELP, etc.);
- essentially linear speech transmission (any amplitude compression or expansion limited to 1 dB) and no peak clipping of a sinusoidal signal giving the same sound pressure level at the measuring position as the test signal;
- overall system frequency response between the octave bands centred on 125 Hz and 8 kHz is uniform, i.e. the difference in sound pressure level between any two adjacent octave bands should not exceed 5 dB;
- background noise is free of audible tones and of marked peaks or troughs in the octave-band spectrum;
- background noise is not impulsive and the space is substantially free of discrete echoes, particularly flutter echoes whose repetition frequency is an integral multiple of one or more of the modulation frequencies [2];
- reverberation time is not 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;



NOTE 1 – L'écart global du rapport signal sur bruit est la somme algébrique des différences entre les rapports signal sur bruit des bandes d'octaves pour les cinq bandes d'octaves de fréquences médianes 125 Hz, 250 Hz, 1 kHz, 4 kHz et 8 kHz et la moyenne arithmétique des rapports signal sur bruit des bandes d'octaves 500 Hz et 2 kHz. Les rapports signal sur bruit dépassant ± 15 dB sont respectivement fixés à $+15$ dB ou -15 dB.

NOTE 2 – Le rapport moyen du temps d'affaiblissement précoce représente la moyenne des bandes d'octaves de fréquences médianes 125 Hz, 250 Hz, 1 kHz, 4 kHz et 8 kHz, divisée par le temps moyen d'affaiblissement précoce des bandes d'octaves 500 Hz et 2 kHz.

Figure 3 – Conditions dans lesquelles les résultats RASTI et STI ne diffèrent pas de plus de 0,05

h) le bruit de fond ne varie pas de façon significative avec le temps.

Il est envisageable d'utiliser les méthodes STI et STITEL lorsque les conditions c) ou d) ne sont pas remplies. Si e), f) ou g) ne sont pas remplies, il convient d'utiliser la méthode STI.

Dans les cas où les conditions a), b) ou h) ne sont pas remplies, il est nécessaire d'utiliser une autre méthode telle que les tests de mots phonétiquement équilibrés ou de rimes modifiées, etc.

4.5 Méthodes de mesure

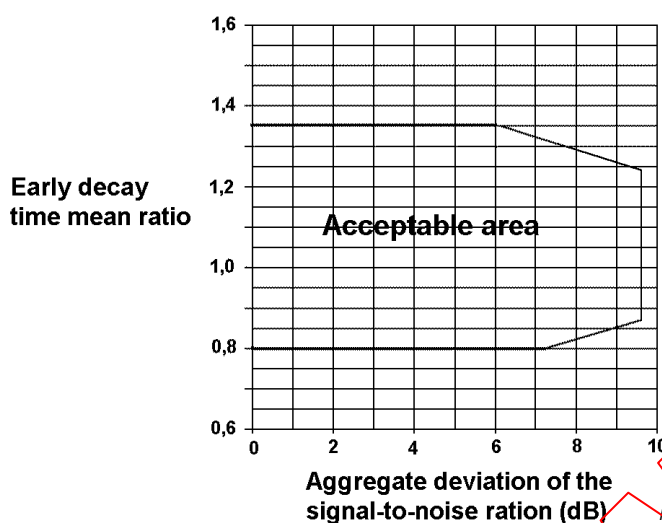
Il convient d'éviter toute compression, non-linéarité d'amplitude, fréquence non stationnaire ou traitement temporel avant d'effectuer les mesures RASTI, mais il est essentiel de s'assurer que les effets sur les niveaux de pression sonore produits par le système en essai soient compensés.

NOTE – Cela n'est pas nécessaire pour les mesures STI ou STITEL.

Généralement, dans un espace d'écoute, l'intelligibilité de la parole dépend de la directivité de la source, par conséquent, il convient d'utiliser une bouche artificielle présentant des caractéristiques de directivité similaires à celles de la tête/bouche humaine (voir UIT-T Recommandation P.51) pour assurer une meilleure précision lors de l'évaluation de l'intelligibilité de locuteurs sans amplification. Quand la parole est transmise par l'intermédiaire d'un système électroacoustique, il ne faut normalement pas de simulateur à moins qu'il n'existe un microphone de proximité ou un microphone antibruit.

4.5.1 Méthode de mesure utilisant un signal d'excitation acoustique

- Placer la source (bouche artificielle ou haut-parleur d'essai approprié) sur l'axe approprié du microphone, à la distance normale de locution (mesurée à partir du cercle formé par les lèvres, pour la bouche artificielle) et l'orienter dans la direction normale de locution.
- Régler le niveau du signal d'essai reçu par le microphone à un niveau égal à celui de la parole dans des conditions normales de fonctionnement. Il convient d'ajuster le niveau de pression sonore (SPL) en utilisant une pondération A et d'avoir un niveau de 68 dB.



NOTE 1 – The aggregate deviation of the signal-to-noise ratio is the algebraic sum of the differences of the octave-band signal-to-noise ratios in the five octave bands centred on 125 Hz, 250 Hz, 1 kHz, 4 kHz and 8 kHz from the arithmetic mean of the signal-to-noise ratios of the 500 Hz and 2 kHz octave bands. Signal-to-noise ratios which exceed ± 15 dB are set to +15 dB or –15 dB, respectively.

NOTE 2 – The average early decay time ratio is the average over the octave bands centred on 125 Hz, 250 Hz, 1 kHz, 4 kHz and 8 kHz, divided by the average early decay time of the 500 Hz and 2 kHz octave bands.

Figure 3 – Conditions under which RASTI and STI results do not differ by more than 0,05

h) background noise does not vary substantially with time.

The STI and the STITEL methods may be considered where conditions c) or d) are not met. Where e), f) or g) are not met, the STI method should be used.

In cases where conditions a), b) or h) are not met, it is necessary to use another method such as phonetically balanced word or modified rhyme tests.

4.5 Methods of measurement

Any compression or non-linear amplitude or non-stationary frequency or temporal processing should be bypassed before carrying out RASTI measurements, but it is essential to ensure that any consequent effects on the sound pressure levels produced by the system under test are compensated.

NOTE – This is not necessary for STI or STITEL measurements.

Generally, in a listening space, speech intelligibility depends upon the directivity of the source; therefore a mouth simulator having similar directivity characteristics to those of the human head/mouth (see ITU-T Recommendation P.51) should be used for the highest accuracy when assessing the intelligibility of unamplified talkers. When speech is relayed through a sound system, a simulator is not normally required unless a close talking or noise cancelling microphone is involved.

4.5.1 Method of measurement using an acoustic excitation signal

- Set the source (artificial mouth or suitable test loudspeaker) on the axis of the appropriate microphone at the normal speaking distance (measured from the lip-circle for the artificial mouth) and direct it in the normal speaking direction.
- Set the test signal level at the microphone to equal that of speech under normal operating conditions. The sound pressure level (SPL) should be set using A-weighting and the level should be 68 dB.

- c) Vérifier que le spectre du signal d'essai est correct et ne présente pas d'écart supérieur à ± 1 dB par rapport à la gamme de fréquence comprise entre 88 Hz et 11,3 kHz (les limites des bandes d'octaves de 125 Hz et 8 kHz). Régler l'égalisation (le cas échéant) de la bouche artificielle ou du haut-parleur d'essai, selon le cas, pour satisfaire à cette exigence.
- d) Lancer la séquence d'essai STI, STITEL ou RASTI. Normalement, et quand elle est disponible, il conviendra de choisir l'option «avec bruit».

4.5.2 Méthode de mesure utilisant l'injection (électrique) directe du signal d'essai

- a) Suivre la procédure décrite en 4.5.1, en remplaçant l'étape b) par l'étape b) ci-dessous, et en sélectionnant l'endroit d'injection du signal de façon à ce qu'il soit situé le plus près possible de l'entrée normale du microphone, afin d'englober au maximum le système dans l'essai.
- b) Régler le niveau de signal d'essai de façon à obtenir un niveau de pression acoustique pondéré A, égal au niveau $L_{Aeq,10}$ du système pour une locution continue.

4.5.3 Simulation du bruit d'occupation

L'effet du bruit d'occupation peut être déterminé soit

- a) en entrant manuellement des données relatives au bruit dans le tableau des données de bruit utilisé par l'équipement de mesure, ou
- b) en mélangeant un signal de bruit artificiel ou enregistré de contenu et de niveau spectral adéquat avec l'entrée du signal dans l'analyseur.

4.5.4 Essais d'intelligibilité utilisant un équipement d'analyse de séquence de type longueur maximale (MLS)

Les limitations indiquées en 4.2.3 (STI) ou 4.4.3 (RASTI) s'appliquent également à cet équipement.

NOTE – Si le signal d'essai MLS traverse d'abord un filtre linéaire présentant un temps d'intégration suffisant pour engendrer une distribution d'amplitude approximativement gaussienne (par exemple le filtre vocal utilisé pour compenser le bruit, voir 4.5.4b), les distorsions, telles que celles résultant des limitations de la vitesse de balayage de l'amplificateur, seront correctement traitées comme un bruit plutôt que comme un faux écho. Voir également [4].

- a) Il est permis de réaliser les mesures directement ou en effectuant des enregistrements numériques des signaux d'essai et de la réponse du système (à partir du microphone d'essai) pour le post-traitement. Si l'entrée du système passe par un microphone de proximité, il convient d'appliquer le signal d'essai à partir d'une bouche artificielle (voir UIT-T Recommandation P.51). Si la conception du microphone rend cette méthode physiquement impraticable, il convient d'utiliser un transducteur approprié, tel qu'une petite enceinte mono-source équipée d'un haut-parleur unique de haute qualité (diamètre du cône ne dépassant pas 100 mm), et de décrire les résultats.
- b) Il convient de régler l'équipement d'essai de façon à produire une séquence d'une durée égale ou supérieure à 1 s, et d'utiliser le filtre vocal. Pour un essai STI, il convient de vérifier que les bandes passantes à 1 dB correspondant à la fois au signal d'essai (avant le filtre vocal) et à la section réceptrice s'étendent au moins de 88 Hz à 11,3 kHz.
- c) Il n'est pas exclu que des types particuliers d'équipements de mesure n'imposent pas l'emploi d'autres conditions de mesure spéciales.

NOTE – Les techniques de moyennage et de perte de l'ondulation ne sont pas appropriées à cette mesure.

4.5.5 Simulation du bruit d'occupation

Il est permis de l'effectuer de la façon suivante.

- a) Si l'on utilise une restitution directe (câblée) de la réponse du système, il convient de mélanger avec le signal restitué un bruit de même distribution spectrale que le bruit d'occupation attendu, au niveau approprié, afin d'obtenir le rapport signal ambiant sur bruit escompté (voir note).

- c) Check that the test signal spectrum is correct to within ± 1 dB over the range 88 Hz to 11,3 kHz (the limits of the 125 Hz and 8 kHz octave bands). Adjust the equalization (if any) of the artificial mouth or test loudspeaker, as necessary, to satisfy this requirement.
- d) Run the STI, STITEL or RASTI test sequence. Normally, and where available, the “with noise” option should be selected.

4.5.2 Method of measurement using direct (electrical) injection of the test signal

- a) Follow the above procedure, replacing step b) by step b) below, and selecting the injection point for the signal to be as close as possible to the normal microphone input, so as to include as much of the system as possible in the test.
- b) Set the test signal level so that an A-weighted sound pressure level, equal to the $L_{Aeq,10}$ of the system for continuous speech, is produced.

4.5.3 Simulation of occupancy noise

The effect of occupancy noise can be determined either

- a) by manually entering noise data into the noise data table used by the measuring equipment; or
- b) by mixing an artificial or recorded noise signal of the correct spectral content and level with the signal input to the analyzer.

4.5.4 Intelligibility tests using maximum length sequence (MLS) analysis equipment

The limitations given in 4.2.3 (STI) or 4.4.3 (RASTI) also apply with this equipment.

NOTE – If the MLS test signal is first passed through a linear filter of sufficient impulse response duration to result in an approximately Gaussian amplitude distribution (e.g. the speech filter that is used when compensating for noise, see 4.5.4b)), distortions, such as those due to slewing rate limitations in the amplifier, are correctly treated as noise, rather than as a false echo. See also [4].

- a) Measurements may be made directly, or by making digital recordings of the test signals and the system response (from the test microphone) for post-processing. If the system input is via a close-talking microphone, the test signal should be applied from an artificial mouth (see ITU-T Recommendation P.51). If the design of the microphone makes this physically impracticable, a suitable transducer, such as a small, single-source, high quality loudspeaker (cone diameter not exceeding 100 mm), should be used and described with the results.
- b) The test equipment should be set up to provide a sample length of at least 1 s, and the speech-shaping filter should be used. For an STI test, the 1 dB bandwidths of both the test signal (prior to the speech-shaped filter) and of the receiving section should be at least 88 Hz to 11,3 kHz.
- c) Specific types of measuring equipment may require the imposition of other special measuring conditions.

NOTE – Averaging and ripple reduction techniques are not appropriate for this measurement.

4.5.5 Simulation of occupancy noise

This may be accomplished as follows.

- a) If direct (cabled) recovery of the system response is used, noise of the same spectral distribution as the expected occupancy noise should be mixed with the recovered signal at the correct level to give the expected ambient signal-to-noise ratio (see note).

- b) Si l'on utilise l'enregistrement numérique de la réponse du système, il convient de mélanger avec le signal enregistré avant post-traitement un bruit de distribution spectrale identique au bruit d'occupation attendu, au niveau approprié, afin d'obtenir le rapport signal ambiant sur bruit escompté (voir note).

NOTE – En fonction d'artefacts particuliers de certains équipements/logiciels, il peut être nécessaire d'augmenter le niveau de bruit de 3 dB par rapport au niveau demandé simplement pour obtenir le rapport signal ambiant sur bruit attendu.

4.5.6 Répétition des mesures

Il convient d'effectuer des mesures répétées à plusieurs reprises et d'en établir une valeur moyenne. Il convient également d'en tirer une estimation de l'écart type et de l'inclure parmi les résultats.

4.5.7 Analyse et interprétation des résultats

Il est important d'examiner la matrice réduite de la modulation pour déterminer la fiabilité des résultats.

Par principe, il convient que les valeurs de chaque colonne de bandes d'octaves diminuent quand la fréquence de modulation augmente. Des valeurs constantes ou légèrement décroissantes dans une colonne indiquent la présence de bruit. Des diminutions importantes indiquent que la réverbération constitue l'effet principal. Des valeurs qui diminuent au début, puis augmentent avec la fréquence de modulation indiquent la présence de réflexions périodiques ou fortes, susceptibles d'occasionner une conclusion trop optimiste. Il convient, si cet effet est détecté, d'en faire part avec les résultats et d'appliquer une correction estimative.

L'entraînement des spécialistes est nécessaire pour déterminer la correction estimée nécessaire. L'annexe D fournit un aperçu du mode opératoire. Dans le cas où le temps de réverbération dépasse 4 s, il convient de prendre des précautions afin d'éviter les effets de distorsion de repliement.

5 Autres méthodes pour déterminer l'intelligibilité

NOTE 1 – Dans les essais parlés, l'insertion de mots dans des phrases porteuses produit un effet équivalent à de la réverbération/des échos au cours de la présentation des mots d'essai.

NOTE 2 – Voir également [5].

5.1 Partitions de mots phonétiquement équilibrées (liste de 256 ou 1 000)

Les limitations sont indiquées dans l'ISO/TR 4870. Il convient de noter qu'en raison du fondement de la méthode sur la reconnaissance de mots par les auditeurs, il n'existe aucune contrainte concernant les caractéristiques du système électroacoustique ou celles de l'environnement.

5.2 Essais de rimes modifiées

Les limitations sont similaires à celles indiquées dans l'ISO/TR 4870. Il convient de noter qu'en raison du fondement de la méthode sur la reconnaissance de mots par les auditeurs, il n'existe aucune contrainte concernant les caractéristiques du système électroacoustique ou celles de l'environnement.

5.3 Indice d'articulation

NOTE – L'indice d'articulation se réfère également à l'indice d'intelligibilité de la parole (SII).

Les limites sont données par le document cité en référence [6].

- b) If digital recording of the system response is used, noise of the same spectral distribution as the expected occupancy noise should be mixed with the recorded signal, prior to post-processing and at the correct level to give the expected ambient signal-to-noise ratio (see note).

NOTE – Depending on particular artifacts of certain equipment / software, it may be necessary to set the noise level 3 dB higher than that needed simply to give the expected signal-to-noise ratio.

4.5.6 Repetition of measurements

The measurements should be repeated several times and the results averaged. Also, an estimate of the standard deviation should be derived from them and included with the results.

4.5.7 Analysis and interpretation of the results

It is important to examine the modulation reduction matrix to determine the reliability of the results.

As a rule, the values in each octave-band column should decrease with increasing modulation frequency. Constant or slightly reducing values in a column indicate the presence of noise. Large reductions indicate that reverberation is the main effect. Values that first reduce and then increase with modulation frequency indicate the presence of periodic or strong reflections, which may produce an over-optimistic conclusion. It is recommended that if this effect is detected, it should be reported with the results and an estimated correction applied.

Specialist training is required in order to determine the necessary estimated correction. Annex D provides an outline of the procedure. For reverberation times exceeding 4 s, care should be taken that aliasing effects are avoided.

5 Other methods of determining intelligibility

NOTE 1 – In spoken tests, embedding the words in carrier phrases results in representative reverberation/echoes during the presentation of the test words.

NOTE 2 – See also [5].

5.1 Phonetically balanced word scores (256 or 1 000 population)

The limitations are given in ISO/TR 4870. It should be noted that, because the method is based on the perception of words by listeners, there are no limitations in respect of the characteristics of the sound system or those of the environment.

5.2 Modified rhyme tests

The limitations are similar to those given in ISO/TR 4870. It should be noted that, because the method is based on the perception of words by listeners, there are no limitations in respect of the characteristics of the sound system or those of the environment.

5.3 Articulation index

NOTE – The articulation index is also being referred to as the speech intelligibility index (SII).

The limitations are given in [6].

5.4 Perte d'articulation des consonnes

Les limitations sont similaires à celles indiquées dans l'ISO/TR 4870. Il convient de noter qu'en raison du fondement de la méthode sur la reconnaissance de mots par les auditeurs, il n'existe aucune contrainte concernant les caractéristiques du système électroacoustique ou celles de l'environnement. Cependant, si l'on utilise une autre méthode de mesure, il n'est pas exclu que des contraintes existent dans ce cas.

IECNORM.COM: Click to view the full PDF of IEC 60268-16:1998

Withd2Wn

5.4 Articulation loss of consonants

The limitations are similar to those given in ISO/TR 4870. It should be noted that if the measurement procedure is based on the reception of words by listeners, there are no limitations in respect of the characteristics of the sound system or those of the environment. If, however, another method of measurement is used, there may be limitations in those respects.

IECNORM.COM: Click to view the full PDF of IEC 60268-16:1998

Withd 2020

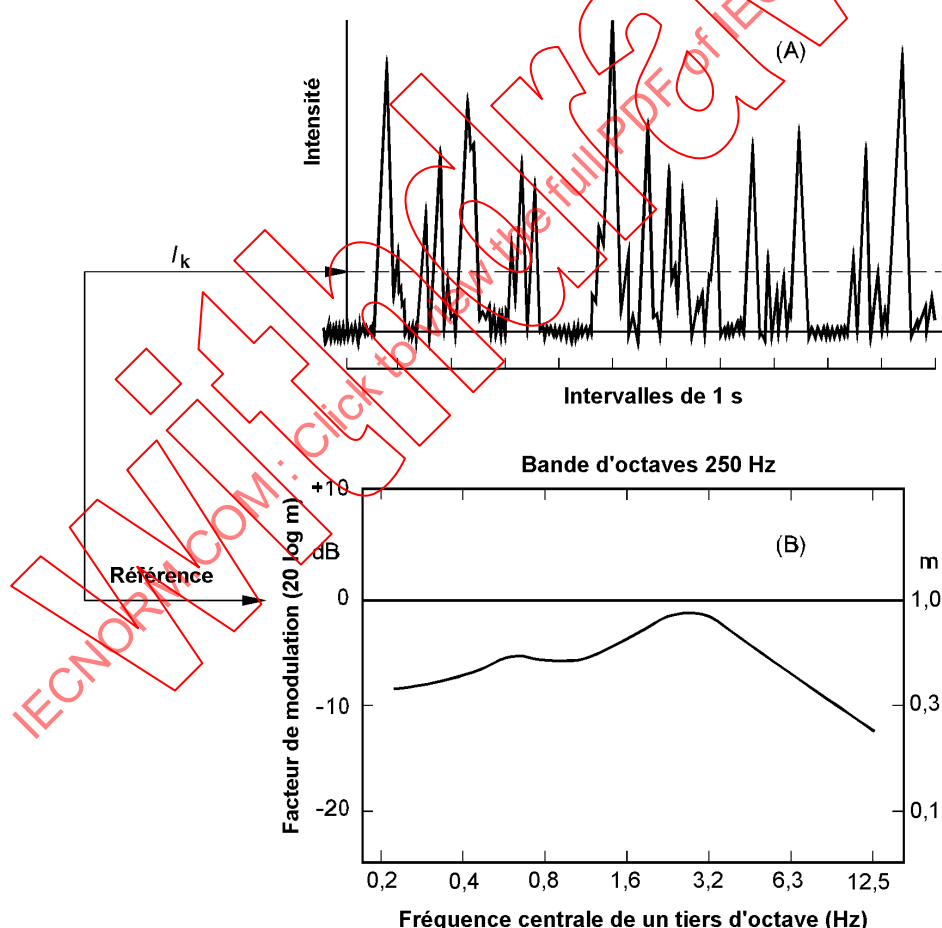
Annexe A (normative)

Méthode de l'indice de transmission de la parole (STI) et méthode révisée (STI_r)

A.1 Contexte

A.1.1 Fonction enveloppe et spectre enveloppe

Un discours continu peut être considéré comme une séquence de fragments élémentaires de parole, appelés phonèmes. Chaque phonème est caractérisé par un spectre de fréquence spécifique. Pour assurer la clarté, il faut que les différences spectrales des phonèmes soient préservées. Ces différences spectrales peuvent être caractérisées par la fonction enveloppe à l'intérieur d'un certain nombre de bandes de fréquences. Une distorsion du signal vocal, du fait du bruit ou de la réverbération par exemple, engendre une perte des différences spectrales entre phonèmes, ce qui se reflète par une perte des fluctuations de la fonction enveloppe.



NOTE – Le spectre est normalisé par rapport à l'intensité moyenne I_k .

IEC 177/98

Figure A.1 – Fonction enveloppe (schéma A) d'un signal vocal de 10 s pour une bande d'octaves de 250 Hz et le spectre enveloppe correspondant (schéma B)

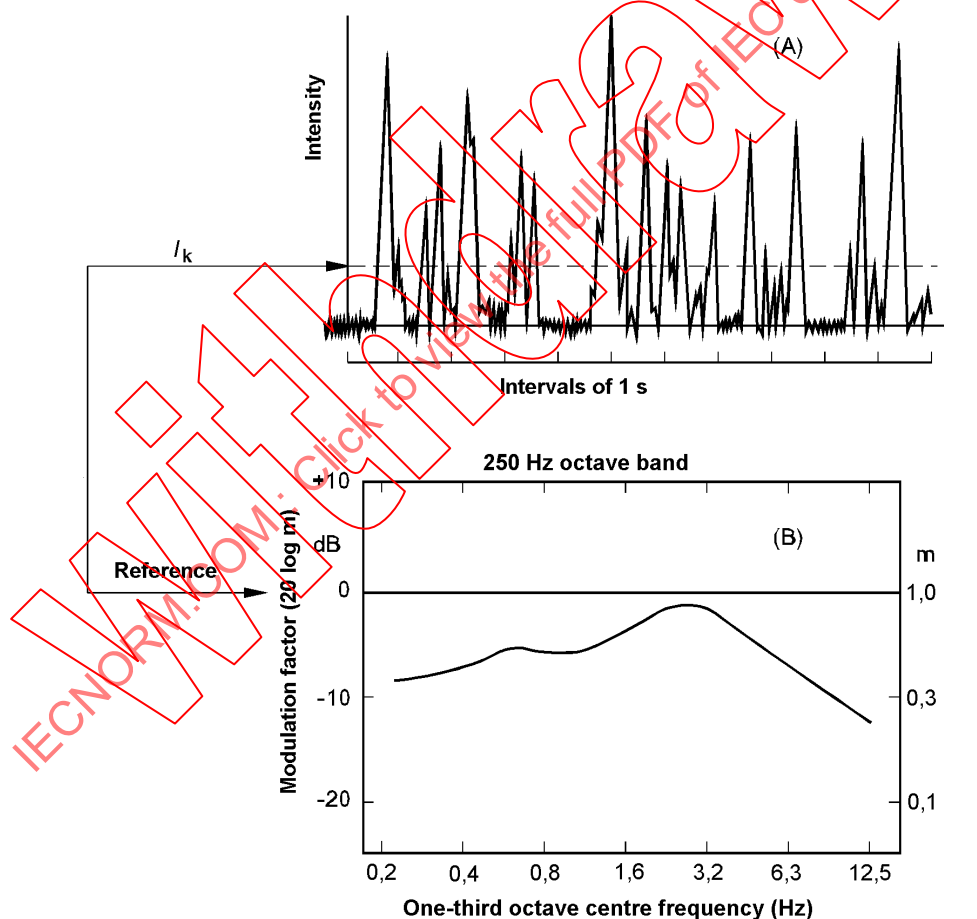
Annex A (normative)

Speech transmission index (STI) and revised (STI_r) methods

A.1 Background

A.1.1 Envelope function and envelope spectrum

Connected discourse can be considered as a sequence of the smallest speech fragments, called phonemes. Each phoneme is characterised by a specific frequency spectrum. Clarity requires that the spectral differences of the phonemes are preserved. These spectral differences can be typified by the envelope function within a number of frequency bands. Distortion of the speech signal, such as by noise or reverberation, results in a reduction of the spectral differences between phonemes and this is reflected in a reduction in the fluctuations in the envelope function.



NOTE – The spectrum is normalized with respect to the mean intensity I_k .

Figure A.1 – Envelope function (panel A) of a 10 s speech signal for the 250 Hz octave band and corresponding envelope spectrum (panel B)

La figure A.1 (A) montre la fonction enveloppe pour une bande de fréquences d'octaves de 250 Hz. La forme de la fonction enveloppe est unique pour une séquence spécifique de phonèmes. Le spectre enveloppe fournit une description plus générale des fluctuations de la fonction enveloppe et résulte d'une analyse par bande de tiers d'octave de la fonction enveloppe. Habituellement, un extrait de discours d'une minute est analysé afin de refléter la distribution spectrale des fluctuations de l'enveloppe autour de l'intensité moyenne, donnant ainsi l'indice de modulation en fonction de la fréquence de modulation (voir figure A.1 (B)).

Une comparaison des spectres enveloppe mesurés directement au niveau du locuteur avec des spectres correspondants obtenus via une voie de transmission indique la perte des fluctuations dues à la voie en question, et conduit à la fonction de transfert de modulation ou MTF, qui représente le facteur de perte de l'indice de modulation comme étant une fonction de la fréquence de modulation.

A.1.2 La fonction de transfert de modulation (MTF)

La raison sous-jacente à l'application du concept de MTF aux études relatives à l'acoustique des salles a été décrite par ailleurs [3,7,8]. La MTF quantifie l'étendue des pertes de modulations du signal original comme une fonction de la fréquence de modulation. Les modulations sont définies par l'enveloppe de l'intensité du signal comme cela est le cas dans le domaine de l'intensité où le bruit ou la réverbération affecterait uniquement la profondeur de modulation d'une modulation sinusoïdale sans modifier sa forme. La figure A.2 en fournit une illustration pour la bande d'octaves de fréquence médiane 250 Hz dans le cas de deux systèmes de transmission simples, l'un avec réverbération seulement (cas A; $T = 2,5$ s) et l'autre avec interférence de bruit seulement (cas B; rapport signal sur bruit $S/N = 0$ dB).

Dans le cas de la réverbération, la MTF a la forme d'un filtre passe-bas: les fluctuations les plus rapides étant relativement plus affectées. Dans le cas théorique d'une réverbération purement exponentielle, la MTF peut être mathématiquement dérivée (voir figure A.2 cas A) et le produit FT (F = fréquence de modulation, T = temps de réverbération) détermine la coupure. Dans le cas du bruit, la MTF est définie par le rapport signal sur bruit et elle est indépendante de la fréquence de modulation: le bruit en augmentant l'intensité moyenne, réduit l'indice de modulation pour toutes les fréquences de modulation.

A.1.3 Détermination de la fonction MTF

La MTF d'une voie de transmission acoustique peut être déterminée de multiples façons, le principe consistant à dériver le facteur de perte de modulation à partir de la comparaison des modulations d'intensité à l'entrée et à la sortie de la voie. Ainsi, est-il possible d'utiliser des signaux vocaux a), la réponse impulsionnelle b) ou des signaux d'essai spéciaux c).

Figure A.1 (A) shows the envelope function for the 250 Hz octave frequency band. The shape of the envelope function is unique for a specific sequence of phonemes. The envelope spectrum gives a more general description of the fluctuations of the envelope function and results from a one-third octave-band analysis of the envelope function. Typically a 1 min speech excerpt is analyzed to reflect the spectral distribution of the envelope fluctuations about the mean intensity so giving the modulation index as a function of modulation frequency (see figure A.1 (B)).

Comparison of the envelope spectra obtained directly from the speaker with corresponding spectra obtained via a transmission path gives the reduction in fluctuations due to the path and leads to the modulation transfer function or MTF, which represents the reduction factor of the modulation index as a function of modulation frequency.

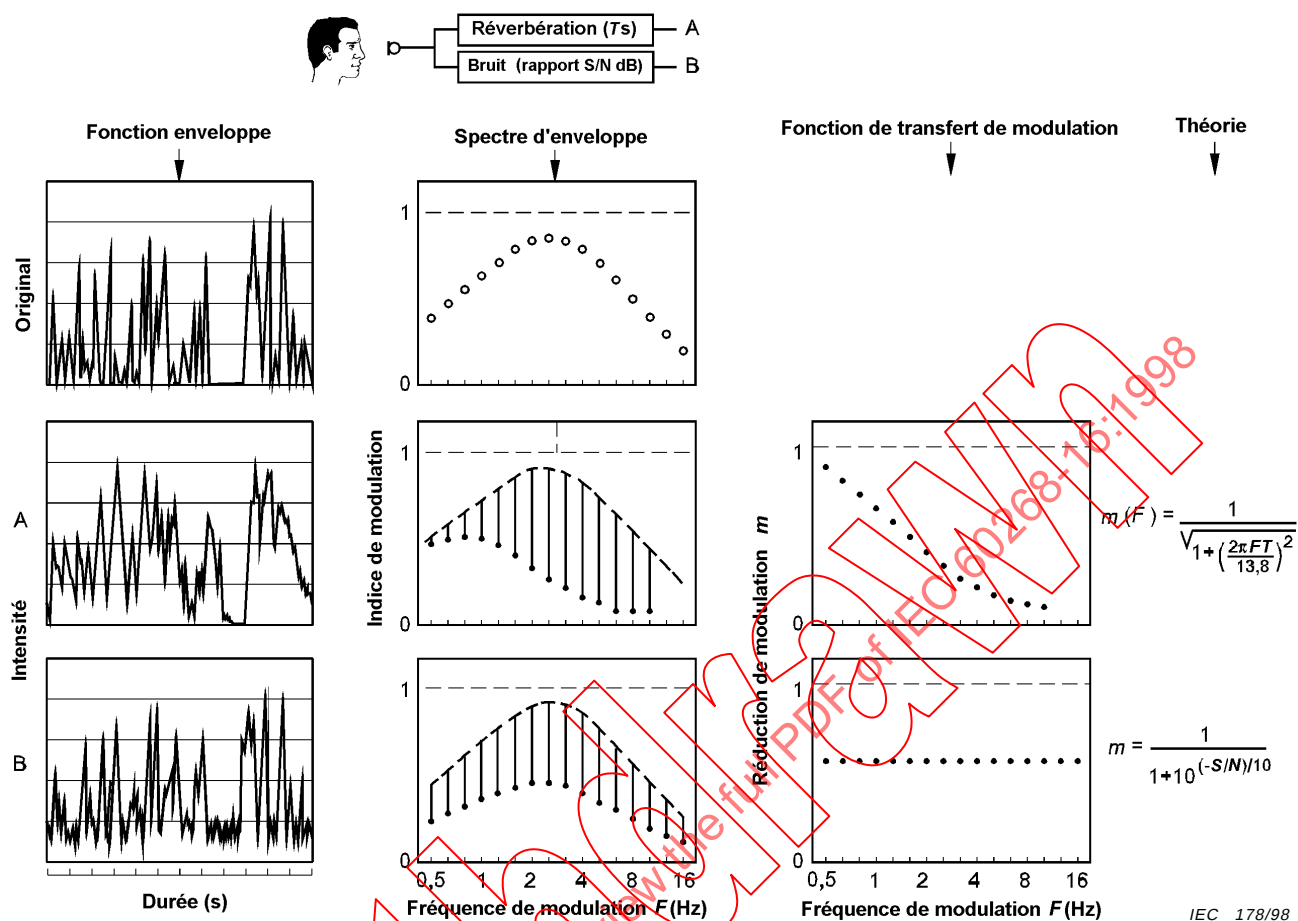
A.1.2 Modulation transfer function (MTF)

The rationale underlying the application of the MTF concept to studies of room acoustics has been described elsewhere [3,7,8]. The MTF quantifies the extent of the reductions in the modulations of the original material as a function of the modulation frequency. The modulations are defined by the intensity envelope of the signal, as it is in the intensity domain that interfering noise or reverberation will affect only the depth of modulation of a sinusoidal modulation without changing its shape. Figure A.2 illustrates this for the octave-band centred on 250 Hz for two simple transmission systems, one with reverberation only (case A; $T = 2,5$ s) and the other with only interfering noise (case B; signal-to-noise ratio $S/N = 0$ dB).

With reverberation, the MTF has the shape of a low-pass filter, the faster fluctuations being relatively more affected. In the theoretical case of purely exponential reverberation, the MTF can be derived mathematically (see figure A.2, case A) and the FT product (F = modulation frequency, T = reverberation time) determines the roll-off. For noise, the MTF is defined by the signal-to-noise ratio and is independent of the modulation frequency: the noise, by increasing the mean intensity, reduces the modulation index for all modulation frequencies.

A.1.3 Determining the MTF

The MTF of a sound transmission path can be determined in various ways, the principle being the derivation of the modulation reduction factor from the comparison of the intensity modulations at the output and the input to the path. Thus, a) speech signals, or b) the impulse response, or c) special test signals can be used.



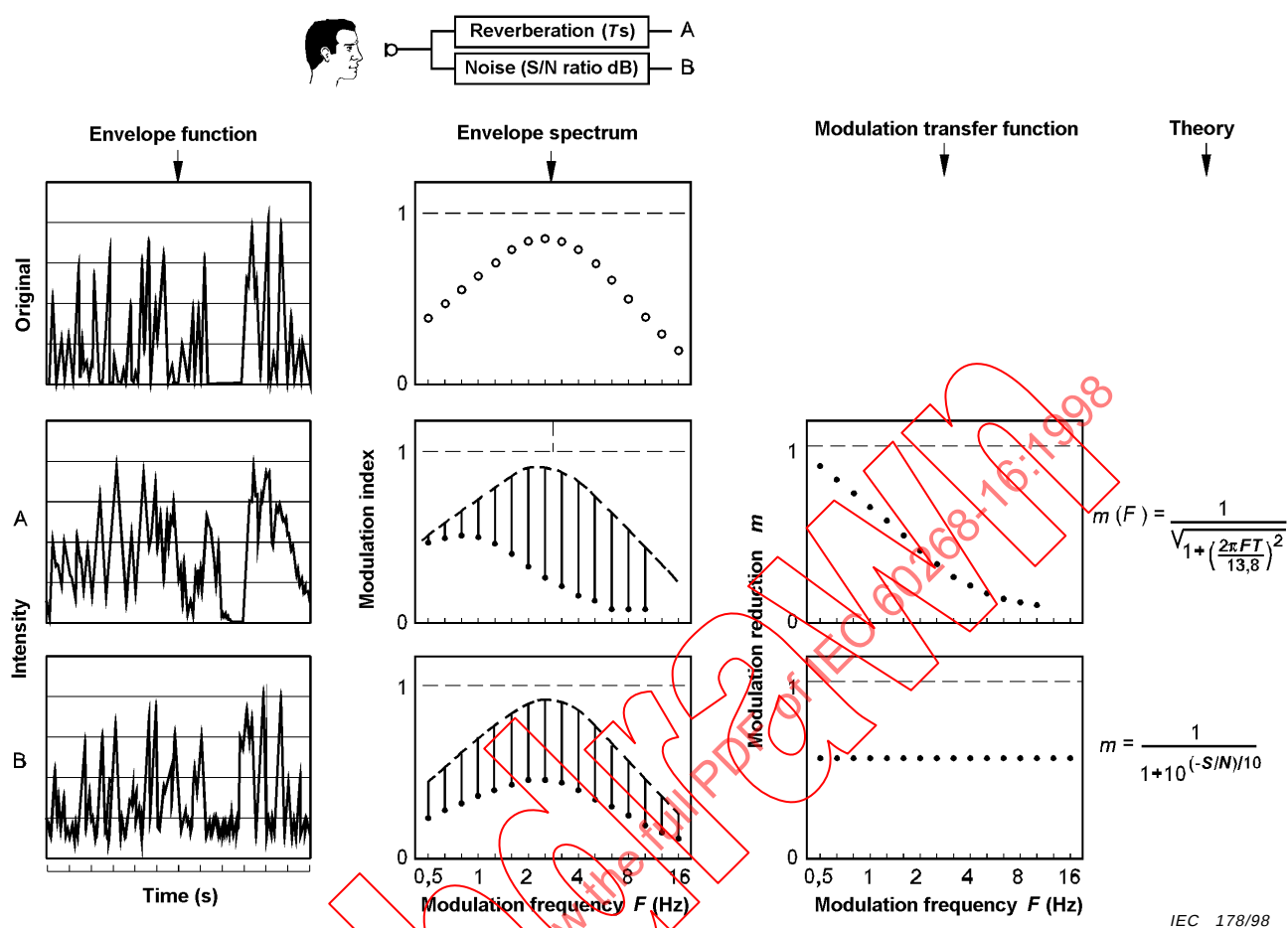
NOTE – Les pertes (spécifiquement les bandes d'octaves) de l'enveloppe d'un signal de sortie (A ou B) par rapport au signal d'origine peuvent être exprimées par la MTF. Les deux conditions considérées (réverbération, A, ou interférence de bruit, B) conduisent à caractériser les MTF conformément aux expressions théoriques illustrées ici.

Figure A.2 – Expressions théoriques de la MTF

- L'utilisation des signaux vocaux présente l'avantage de permettre la détermination de la MTF dans des conditions particulièrement réalistes, par exemple au cours d'une prestation en direct, mais en contrepartie la précision est moindre [9].
- La réponse impulsionnelle peut être utilisée pour déterminer la MTF ainsi que pour mesurer l'effet de la réverbération ou des échos.

NOTE – Les signaux impulsionnels ne sont pas adaptés quand le bruit de fond, la limitation de la bande passante ou la distorsion de non-linéarité (systèmes de diffusion sonore) sont importants car le spectre moyen de la parole et la distribution des niveaux ne sont pas représentés dans le signal d'essai.

- Des signaux d'essai artificiels appropriés permettent de déterminer le facteur de perte de la modulation pour chaque fréquence de modulation, les mesures étant effectuées pour les fréquences médianes de bandes d'octaves comprises typiquement entre 125 Hz et 8 kHz. Pour que la détermination soit représentative du rapport signal sur bruit, il convient que l'intensité moyenne du signal d'essai soit identique au niveau normal de parole à la position d'essai, c'est-à-dire que le niveau L_{eq} du signal d'essai est réglé en fonction du niveau L_{eq} typique de la parole en cours à cette position.



NOTE – The reductions in the octave-band specific envelope of an output signal (A or B) relative to the original signal can be expressed by the MTF. The two conditions considered (reverberation, A, or noise interference, B) lead to characteristic MTFs according to the theoretical expressions shown.

Figure A.2 – Theoretical expression of the MTF

- The use of speech signals has the advantage that the MTF can be determined under fully realistic conditions e.g. a live performance, but is less accurate [9].
- The impulse response can be used to determine the MTF and the effect of reverberation or echoes on it.

NOTE – Impulse signals are not suitable where background noise, bandpass limiting and non-linear distortion (sound distribution systems) are significant since the average speech spectrum and level distribution are not represented in the test signal.

- Appropriate artificial test signals allow the determination of the modulation reduction factor for each modulation frequency and measurements are performed for octave-band centre frequencies typically from 125 Hz to 8 kHz. For a determination representative of the signal-to-noise ratio, the mean intensity of the test signal should be equivalent to the normal speech level at the test position, i.e. the L_{eq} of the test signal is adjusted to the typical L_{eq} of on-going speech in that position.

A.2 La méthode STI

A.2.1 Généralités

L'indice de transmission de la parole (STI) est une mesure objective, fondée sur la contribution pondérée d'un certain nombre de bandes de fréquences contenues dans la gamme de fréquences des signaux vocaux, ces contributions étant déterminées par le rapport signal sur bruit réel. Sa description ainsi que les facteurs de pondération de bandes d'octaves sont indiqués en [3]. Si l'on choisit correctement la nature du signal d'essai, ce rapport réel signal sur bruit peut inclure et autoriser l'accès aux distorsions dans le domaine temporel, aux non-linéarités ainsi qu'au bruit ambiant, etc.

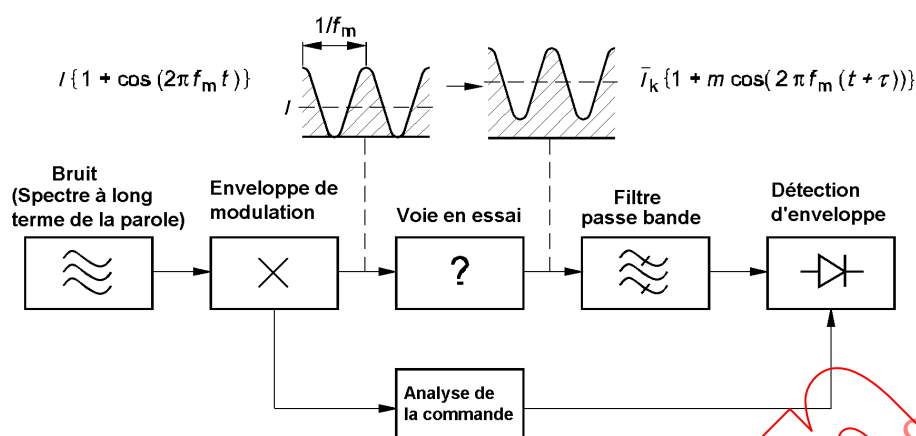
Il n'est pas exclu qu'une distorsion dans le domaine temporel (réverbération, échos et commande automatique de gain) dégrade le signal vocal modulé et réduise l'intelligibilité. Cela est pris en compte dans la procédure STI en déterminant la fonction de transfert de modulation pour la gamme de fréquences révélatrices, présentes dans l'enveloppe de signaux vocaux naturels. La gamme représentative pour ces fréquences de modulation s'étend de 0,63 Hz à 12,5 Hz en quatorze bandes de tiers d'octave. La figure A.3 illustre, schématiquement, une disposition de mesure dans laquelle la fonction de transfert de modulation, $m(F)$, est déterminée séparément pour chaque fréquence de modulation dans chaque bande d'octaves.

A.2 The STI method

A.2.1 General

The speech transmission index (STI) is an objective measure, based on the weighted contribution of a number of frequency bands within the frequency range of speech signals, the contributions being set by the effective signal-to-noise ratio. Its description and the octave band weighting factors are given in [3]. By proper choice of the form of test signal, this effective signal-to-noise ratio can include and allow for distortions in the time domain and non-linearities as well as background noise, etc.

Distortion in the time domain (reverberation, echoes and automatic gain control) may degrade the fluctuating speech signal and reduce intelligibility. This is modelled in the STI procedure by determining the modulation transfer function for the range of relevant frequencies present in the envelope of *natural* speech signals. The relevant range for these modulation frequencies extends from 0,63 Hz to 12,5 Hz in 14 one-third octave bands. Figure A.3 illustrates, diagrammatically, a measuring arrangement in which the modulation transfer function, $m(F)$, is determined separately for each modulation frequency in each octave band.



Bande d'octaves	125	250	500	1k	2k	4k	8kHz
$f_1 = 0,63 \text{ Hz}$	m						
$f_2 = 0,8 \text{ Hz}$							
$f_3 = 1,0 \text{ Hz}$							
$f_4 = 1,25 \text{ Hz}$							
$f_5 = 1,6 \text{ Hz}$							
$f_6 = 2,0 \text{ Hz}$							
$f_7 = 2,5 \text{ Hz}$							
$f_8 = 3,15 \text{ Hz}$							
$f_9 = 4,0 \text{ Hz}$							
$f_{10} = 5,0 \text{ Hz}$							
$f_{11} = 6,3 \text{ Hz}$							
$f_{12} = 8,0 \text{ Hz}$							
$f_{13} = 10 \text{ Hz}$							
$f_{14} = 12,5 \text{ Hz}$							
$\bar{I}_k$							

IEC 179/98

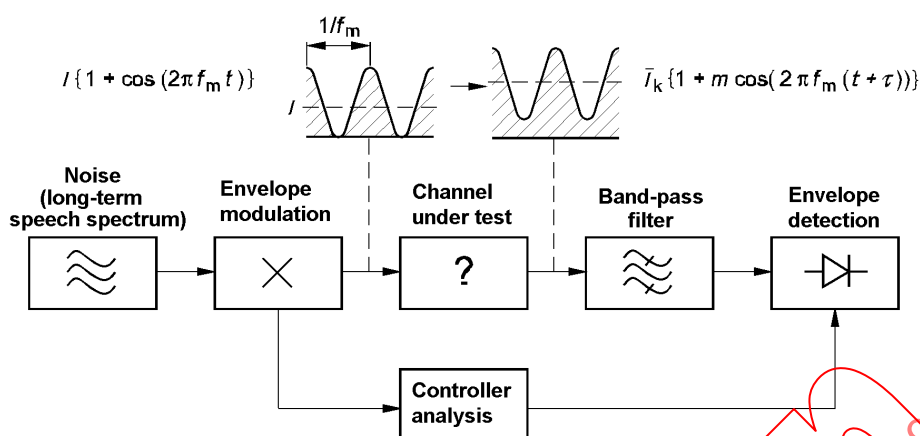
NOTE – La fonction de transfert de modulation (m) est déterminée pour toutes les cellules de la matrice des sept bandes d'octaves et des 14 fréquences de modulation. De même, les niveaux d'octaves $\bar{I}_k$ sont obtenus pour calculer l'étendue du masquage auditif.

Figure A.3 – Système de mesure et fréquences pour la méthode STI

A.2.2 Distorsions non linéaires

Un signal vocal qui traverse un système présentant une fonction de transfert non linéaire (par exemple avec écrêtage) acquiert des composantes harmoniques et d'intermodulation dans d'autres bandes de fréquences. Pour cette raison, il convient de ne pas moduler simultanément le signal d'essai avec la même fréquence de modulation pour toutes les bandes d'octaves car, dans ce cas, il ne sera pas possible de distinguer les composantes de distorsion non linéaire du signal d'essai modulé dans la bande de fréquences considérée.

NOTE – Les systèmes MLS qui ne modulent pas les signaux MLS ne sont pas soumis à cette limitation.



Octave-band	125	250	500	1k	2k	4k	8kHz
$f_1 = 0,63 \text{ Hz}$	m						
$f_2 = 0,8 \text{ Hz}$							
$f_3 = 1,0 \text{ Hz}$							
$f_4 = 1,25 \text{ Hz}$							
$f_5 = 1,6 \text{ Hz}$							
$f_6 = 2,0 \text{ Hz}$							
$f_7 = 2,5 \text{ Hz}$							
$f_8 = 3,15 \text{ Hz}$							
$f_9 = 4,0 \text{ Hz}$							
$f_{10} = 5,0 \text{ Hz}$							
$f_{11} = 6,3 \text{ Hz}$							
$f_{12} = 8,0 \text{ Hz}$							
$f_{13} = 10 \text{ Hz}$							
$f_{14} = 12,5 \text{ Hz}$							
$\bar{I}_k$							

IEC 179/98

NOTE – The modulation transfer function (m) is determined for all cells of the matrix of seven octave bands and 14 modulation frequencies. Also the octave levels $\bar{I}_k$ are obtained for use in calculating the auditory spread of masking.

Figure A.3 – The measurement system and frequencies for the STI method

A.2.2 Non-linearity distortions

A speech signal, that passes through a system with a non-linear transfer function (e.g. with peak clipping) acquires harmonic and intermodulation components in other frequency bands. For this reason, the test signal should not be modulated with the same modulation frequency simultaneously for all the octave bands, because in that case, the distortion components due to non-linearity cannot be distinguished from the modulated test signal in the frequency band under consideration.

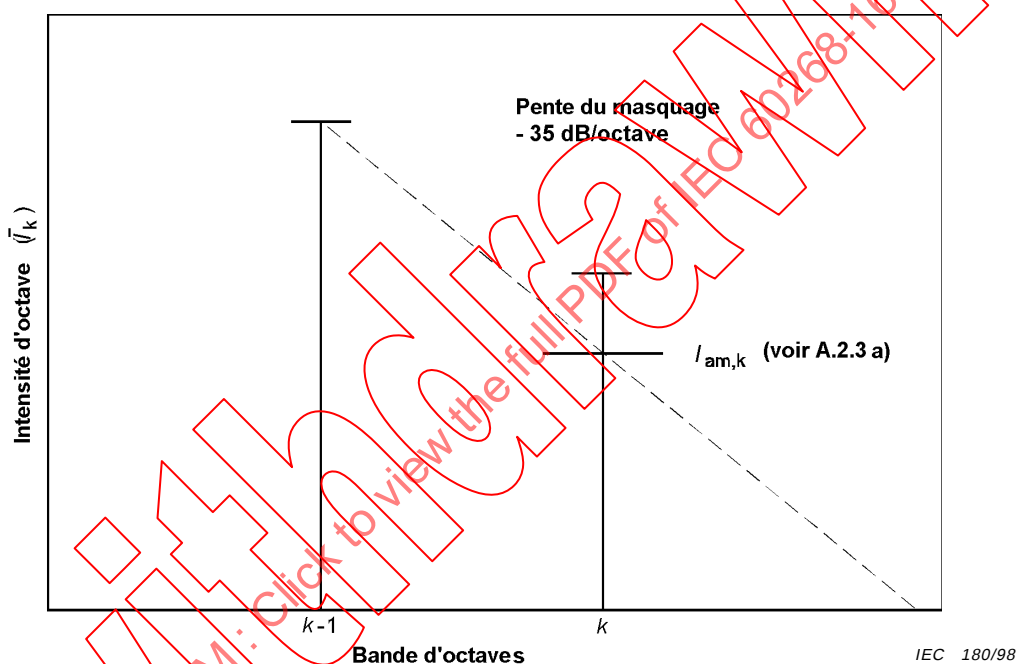
NOTE – MLS systems that do not modulate the MLS signal do not suffer this limitation.

Cependant, en modulant avec des signaux non corrélatifs toutes les autres bandes de fréquences à l'exception de celle testée, les composantes de distorsion non linéaire apparaissent également en tant que bruit, réduisant le rapport signal sur bruit réel, de la même manière que le ferait la distorsion linéaire (réponse de fréquence inégale).

A.2.3 Autres effets sonores

Le masquage auditif (perte de sensibilité due à un son plus fort et de fréquence plus basse) [10] et le seuil d'audition absolu sont pris en compte en modélisant des périodes de bruit appropriés et en les ajoutant, réduisant ainsi encore le rapport signal sur bruit réel et la fonction de transfert de modulation. Pour cela, il faudra tenir compte du niveau du signal dans chacune des bandes de fréquences.

La figure A.4 montre l'effet du masquage dans une bande d'octaves par la bande d'octaves immédiatement inférieure.



NOTE – La pente de –35 dB par octave est équivalente à un facteur de masquage auditif (amf) de 0,000 316.

Figure A.4 – Intensité de l'effet de masquage auditif d'une bande d'octaves ($k - 1$) sur la bande supérieure (k)

a) Modélisé de cette manière, l'effet est indépendant tant de la fréquence de bande que du niveau absolu du signal. On voit que la réponse de l'effet de masque est une ligne droite présentant une pente de –35 dB/octave, correspondant à un facteur de masquage auditif (amf) de 0,000 316. Ce facteur, appliqué à l'intensité dans une bande d'octaves, donne l'effet du masquage sur la bande d'octaves immédiatement supérieure, c'est-à-dire:

$$I_{am,k} = I_{k-1} \cdot amf$$

où

$I_{am,k}$ est l'intensité du masquage audio dans la bande d'octaves k ;

I_{k-1} est l'intensité du signal de masquage dans la bande ($k - 1$).

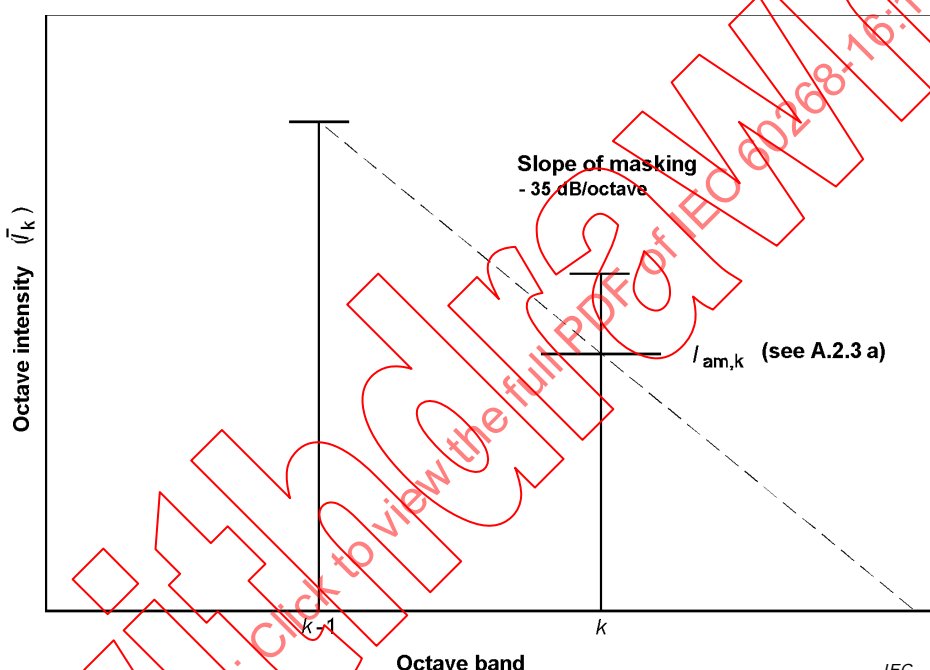
b) La modélisation du seuil d'audition absolu correspond à la limite inférieure du niveau de bruit de masquage à l'intérieur de chaque bande d'octaves ($I_{rs,k}$) et s'applique seulement si l'intensité, I_k , est le niveau opérationnel, c'est-à-dire le niveau (d'écoute) normal.

However, by modulating with uncorrelated signals all the other frequency bands except that under test, the distortion components appear as noise, reducing the effective signal-to-noise ratio, just as with linear distortion (uneven frequency response).

A.2.3 Other auditory effects

Auditory masking, (the reduction in aural sensitivity by a stronger, lower frequency sound) [10], and the absolute hearing threshold, are allowed for by modelling appropriate noise terms and adding them in, so reducing further the effective signal-to-noise ratio and the modulation transfer function. For this, the signal levels in each of the frequency bands have to be considered.

Figure A.4 shows the effect of masking in one octave band by the octave band immediately lower.



IEC 180/98

NOTE – The slope of –35 dB per octave equates to an auditory masking factor (amf) of 0,000 316.

Figure A.4 – Auditory masking strength of octave band (k – 1) on that above (k)

- a) Modelled in this way, the effect is independent of both the band frequency and the absolute signal level. The masking response is seen to be a straight line with a slope of – 35 dB/octave, corresponding to an auditory masking factor (amf), of 0,000 316. This factor, applied to the intensity in one octave band, gives the masking effect on the octave band immediately above, namely:

$$I_{am,k} = I_{k-1} \cdot amf$$

where

$I_{am,k}$ is the intensity of the audio masking in octave band k;

I_{k-1} is the intensity of the masking signal in band (k – 1).

- b) The modelling of the absolute hearing threshold is as the lower limit of the masking noise level within each octave band ($I_{rs,k}$) and only applies if the intensity, I_k , is the operational, i.e. normal (listening) level.

L'application des corrections dues au niveau du masquage et du seuil d'audition donne l'équation suivante:

$$m'_{k,f} = m_{k,f} \frac{l_k}{l_k + l_{am,k} + l_{rs,k}}$$

où

$m_{k,f}$ est l'indice de modulation pour la bande d'octaves k et la fréquence de modulation f ,

$m'_{k,f}$ est l'indice corrigé.

Le rapport signal sur bruit réel pour la bande d'octaves k et la fréquence f de modulation f devient alors:

$$SNR_{k,f} = 10 \lg \frac{m'_{k,f}}{1 - m'_{k,f}} \text{ dB}$$

Selon le concept STI, les rapports signal sur bruit dans la gamme -15 dB à $+15$ dB sont liés de façon linéaire aux contributions à l'intelligibilité dans la gamme 0 à 1 . Par conséquent, on convertit le rapport signal sur bruit réel en un indice de transmission correspondant ($TI_{k,f}$), spécifique à la bande d'octaves k et la fréquence f , selon l'équation suivante:

$$TI_{k,f} = \frac{SNR_{k,f} + \text{shift}}{\text{range}}$$

Cela conduit à $0 \leq TI_{k,f} \leq 1,0$ si la translation est de 15 dB et la gamme de 30 dB.

La figure A.5 en fournit une représentation graphique.

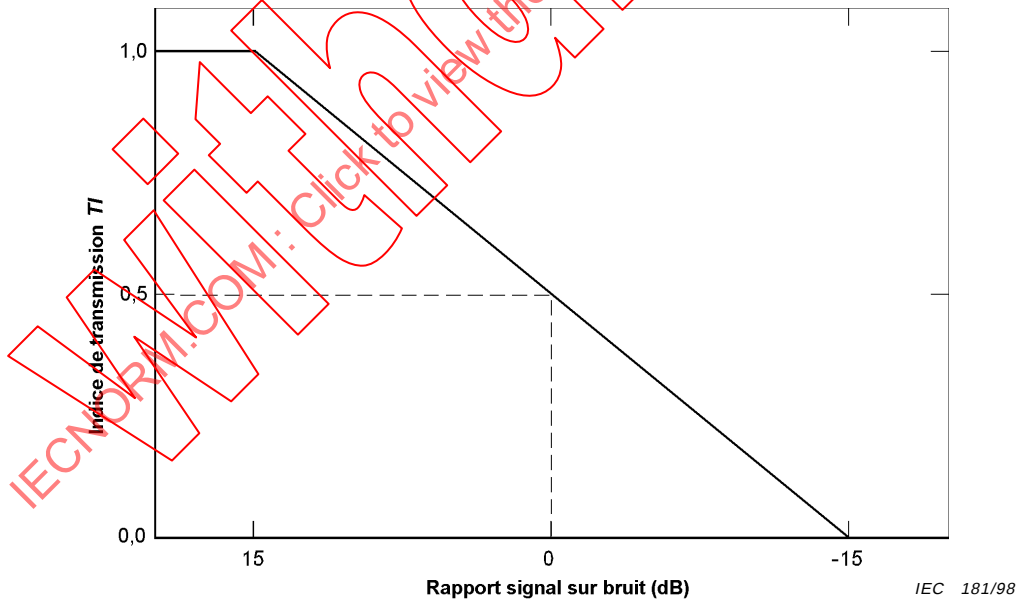


Figure A.5 – Relation entre le rapport signal sur bruit réel et l'indice de transmission pour une translation de 15 dB et une gamme de 30 dB

Les 14 indices de transmission pour les fréquences de modulation échelonnées de $0,63$ Hz à $12,5$ Hz sont tous obtenus pour chaque bande d'octaves et la moyenne est établie pour chaque bande afin d'obtenir l'indice de transfert de modulation (MTI_k) spécifique à la contribution de la bande d'octaves k . C'est-à-dire:

$$MTI_k = \frac{1}{14} \sum_{f=1}^{14} TI_{k,f}$$

Applying the corrections for masking and threshold gives:

$$m'_{k,f} = m_{k,f} \frac{l_k}{l_k + l_{am,k} + l_{rs,k}}$$

where

$m_{k,f}$ is the modulation index for octave band k and modulation frequency f ;

$m'_{k,f}$ the corrected index.

The effective signal-to-noise ratio for octave band k and modulation frequency f then becomes:

$$SNR_{k,f} = 10 \lg \frac{m'_{k,f}}{1 - m'_{k,f}} \text{ dB}$$

According to the STI concept, signal-to-noise ratios in the range -15 dB to $+15$ dB are linearly related to contributions to intelligibility in the range 0 to 1. Therefore the effective signal-to-noise ratio is converted to a corresponding transmission index ($Tl_{k,f}$), specific for octave band k and frequency f , by the equation:

$$Tl_{k,f} = \frac{SNR_{k,f} + \text{shift}}{\text{range}}$$

This makes $0 \leq Tl_{k,f} \leq 1,0$ if *shift* is 15 dB and *range* 30 dB

Figure A.5 shows this graphically.

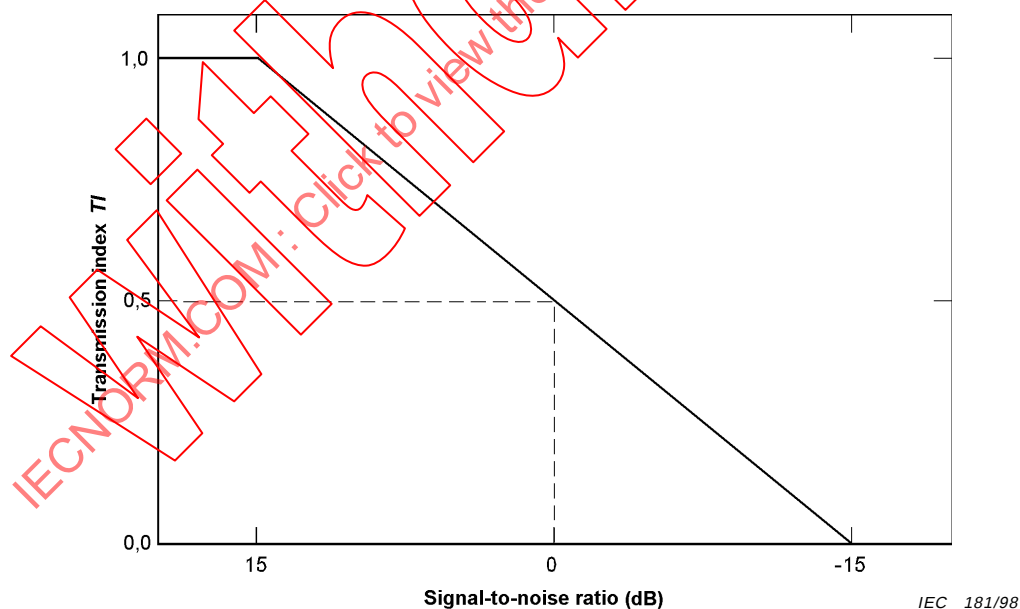


Figure A.5 – The relationship between effective signal-to-noise ratio and transmission index for a shift of 15 dB and a range of 30 dB

All 14 transmission indices for the modulation frequencies from 0,63 Hz to 12,5 Hz are obtained for each octave band and averaged for each band to give the modulation transfer index (MTl_k) specific for the contribution of octave band k , namely:

$$MTl_k = \frac{1}{14} \sum_{f=1}^{14} Tl_{k,f}$$

Enfin, l'indice révisé de transmission de la parole, STI_r , qui prend en compte le masquage auditif et le seuil d'audition absolu, est obtenu en effectuant une addition pondérée des indices de transfert de modulation pour les sept bandes d'octaves avec les corrections de redondance correspondantes.

$$STI_r = \sum_{n=1}^7 \alpha_n MTI_n - \sum_{n=1}^6 \beta_n \sqrt{MTI_n \times MTI_{(n+1)}}$$

où

les facteurs α_k sont les facteurs de pondération d'octave;

β_k sont les facteurs de correction de redondance.

Cette correction de redondance est liée à la contribution des bandes de fréquences adjacentes.

A.2.4 Facteurs relatifs à la parole masculine et féminine

Les facteurs de pondération optimums et les facteurs de redondance relatifs à la parole masculine ou féminine sont indiqués dans le tableau A.1, ainsi que les facteurs de correction de seuil.

Tableau A.1 - Facteurs de pondération masculin et féminin spécifiques à la bande d'octaves STI_r

Bande d'octaves Hz		125	250	500	1 k	2 k	4 k	8 k
Hommes	α	0,085	0,127	0,230	0,233	0,309	0,224	0,173
	β	0,085	0,078	0,065	0,011	0,047	0,095	–
Femmes	α	–	0,117	0,223	0,216	0,328	0,250	0,194
	β	–	0,099	0,066	0,062	0,025	0,076	–
Facteur de seuil d'audition	$I_{rs,k}$	46	27	12	6,5	7,5	8	12

A.3 Les signaux d'essai

A.3.1 Spectres des signaux STI et STITL

Les spectres des signaux d'essai sont spécifiés par niveaux moyens de bandes d'octaves (qui peuvent éventuellement être dépassés instantanément de 3 dB par un signal d'essai modulé). Les niveaux de bandes d'octaves sont normalisés au niveau de la pondération A de 0 dB. Cela coïncide avec le niveau nominal à long terme de parole pondéré A. Dans les signaux d'essai sont incluses sept bandes d'octaves pour la parole masculine et six bandes d'octaves pour la parole féminine.

Tableau A.2 – Niveaux de bande d'octaves (dB) par rapport au niveau de parole à long terme pondéré A

Bande d'octaves Hz	125	250	500	1 k	2 k	4 k	8 k	Niveau pondéré de A
Hommes	2,9	2,9	–0,8	–6,8	–12,8	–18,8	–24,8	0,0
Femmes	–	5,3	–1,9	–9,1	–15,8	–16,7	–18,0	0,0

Finally, the revised speech transmission index, STI_r , which takes account of auditory masking and the absolute hearing threshold, is obtained by a weighted summation of the modulation transfer indices for all seven octave bands and the corresponding redundancy corrections.

$$STI_r = \sum_{n=1}^7 \alpha_n MTI_n - \sum_{n=1}^6 \beta_n \sqrt{MTI_n \times MTI_{(n+1)}}$$

where

factors α_k represent the octave weighting factors;

β_k are the so-called redundancy correction factors.

This redundancy correction is related to the contribution of adjacent frequency bands.

A.2.4 Factors for male and female speech

The optimal weighting factors and redundancy factors for male and female speech are shown in table A.1 along with the threshold correction factors.

Table A.1 – STI_r octave band specific male and female weighting factors

Octave band Hz		125	250	500	1k	2k	4k	8k
Males	α	0,085	0,127	0,230	0,233	0,309	0,224	0,173
	β	0,085	0,078	0,065	0,011	0,047	0,095	–
Females	α	–	0,117	0,223	0,216	0,328	0,250	0,194
	β	–	0,099	0,066	0,062	0,025	0,076	–
Hearing threshold factor	$I_{rs,k}$	46	27	12	6,5	7,5	8	12

A.3 The test signals

A.3.1 Spectra of STI and STITEL signals

The test signal spectra are specified by octave band mean levels (which may instantaneously be exceeded by 3 dB by a modulated test signal). The octave band levels are normalized to an A-weighted level of 0 dB. This coincides with the nominal long term A-weighted speech level. For male speech, seven, and for female, six, octave bands are included in the test signals.

Table A.2 – Octave band levels (dB) relative to the A-weighted long-term speech level

Octave band Hz	125	250	500	1k	2k	4k	8k	A-weighted
Males	2,9	2,9	–0,8	–6,8	–12,8	–18,8	–24,8	0,0
Females	–	5,3	–1,9	–9,1	–15,8	–16,7	–18,0	0,0

A.3.2 Fréquences de modulation STI

0,63; 0,80; 1,00; 1,25; 1,60; 2,00; 2,50; 3,15; 4,00; 5,00; 6,30; 8,00; 10,0; 12,5.

Les 14 valeurs sont appliquées à toutes les bandes d'octaves (sept).

A.3.3 Forme de signal pour le STI (et STITEL)

La porteuse de bruit pour chaque bande d'octaves a une largeur de bande d'un demi-octave pour éviter un débordement des bandes adjacentes et tenir compte des filtres utilisés par l'analyseur. Par conséquent, échelonnées tous les octaves, il existe sept composantes d'un demi-octave pour la parole masculine et six pour la parole féminine.

L'enveloppe d'intensité de la porteuse de bruit pour chaque bande d'octaves est modulée sinusoïdalement avec la ou les fréquences de modulation. Au niveau de l'amplitude, la fonction de modulation est la racine carrée de l'enveloppe d'intensité désirée:

$$\sqrt{1 + \cos 2\pi \cdot f_m t}$$

A.3.2 STI modulation frequencies

0,63; 0,80; 1,00; 1,25; 1,60; 2,00; 2,50; 3,15; 4,00; 5,00; 6,30; 8,00; 10,0; 12,5

with all 14 frequencies applied to all seven octave bands.

A.3.3 Form of signal for STI (and STITEL)

The noise carrier for each octave band has a bandwidth of half an octave to avoid spillover from adjacent bands and to allow for practical filters in the analyzer. Therefore, for male speech there are seven, and for female six, half-octave components at octave intervals.

The intensity envelope of the noise carrier for each octave band is sinusoidally modulated with the modulation frequency or frequencies. In the amplitude domain, the modulation function is the square root of the desired intensity envelope:

$$\sqrt{1 + \cos 2\pi \cdot f_m t}$$

Annexe B (informative)

La méthode STITEL

Au lieu de 14 fréquences de modulation appliquées à toutes les bandes d'octaves (sept), comme c'est le cas dans le mode opératoire complet STI, la méthode STITEL applique uniquement sept fréquences de modulation, une à chacune des sept bandes d'octaves, comme illustré dans le tableau B.1.

Tableau B.1 – STITEL: fréquences de modulation pour les sept bandes d'octaves

Bande d'octaves Hz		125	250	500	1 k	2 k	4 k	8 k
Fréquence de modulation	Hz	1,12	11,33	0,71	2,83	6,97	1,78	4,53

Le signal d'essai comprend les sept bandes d'octaves modulées qui sont toutes analysées simultanément.

Annex B

(informative)

The STITEL method

Instead of the 14 modulation frequencies applied to all seven octave bands as is the procedure for the full STI, the STITEL method applies, uniquely, seven modulation frequencies one to each of the seven octave bands as shown in table B.1.

Table B.1 – STITEL: modulation frequencies for the seven octave bands

Octave band Hz	125	250	500	1 k	2 k	4 k	8 k
Modulation frequency Hz	1,12	11,33	0,71	2,83	6,97	1,78	4,53

The test signal includes all seven modulated octave bands, and these are all analyzed simultaneously.

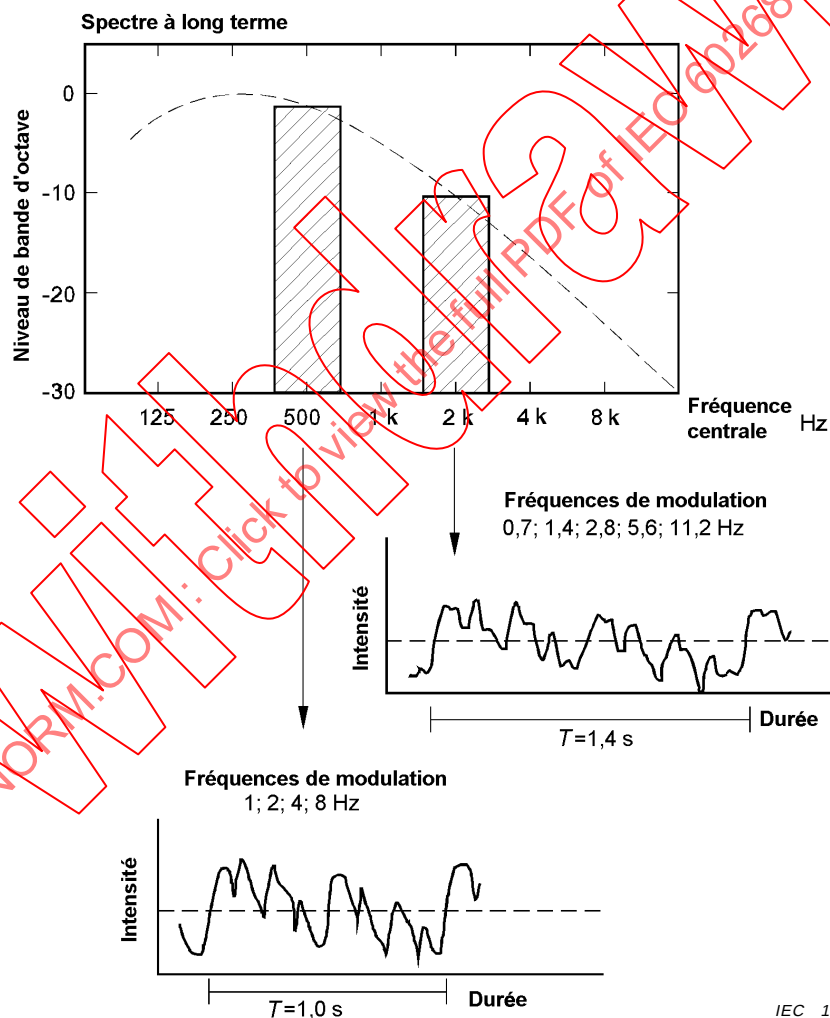
Annexe C (informative)

La méthode RASTI

Comme indiqué dans la figure C.1 on utilise un total de neuf fréquences de modulation:

- bande d'octaves 500 Hz: 1,0; 2,0; 4,0; 8,0 Hz
- bande d'octaves 2 kHz: 0,7; 1,4; 2,8; 5,6; 11,2 Hz

La pondération des fréquences est appliquée indirectement en utilisant quatre contributions pour la bande d'octaves 500 Hz et cinq pour la bande d'octaves 2 kHz, c'est-à-dire que les pondérations sont respectivement 4/9 (0,45) et 5/9 (0,55).



IEC 182/98

Figure C.1 – Illustration d'un signal d'essai RASTI en pratique

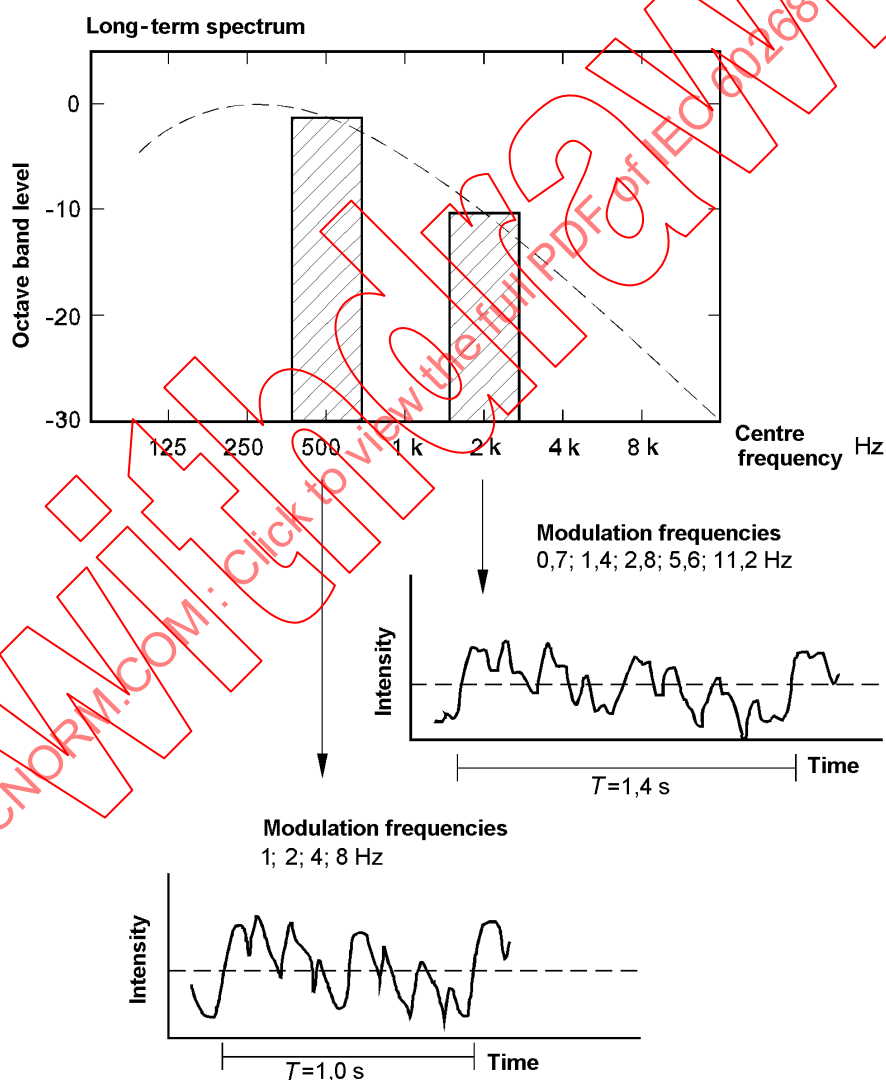
Annex C (informative)

The RASTI method

As shown in figure C.1, a total of nine modulation frequencies are used:

- 500 Hz octave band 1,0; 2,0; 4,0; 8,0 Hz
- 2 kHz octave band 0,7; 1,4; 2,8; 5,6; 11,2 Hz

Frequency weighting is applied indirectly by using four contributions for the 500 Hz octave band and five for the 2 kHz octave band, i.e. the weightings are $4/9$ (0,45) and $5/9$ (0,55), respectively.



IEC 182/98

Figure C.1 – Illustration of a practical RASTI test signal

Annexe D (informative)

Estimation de la correction dans le cas de fortes réflexions

D.1 Généralités

Les modes opératoires précisés dans cette annexe indiquent les moyens d'estimer la correction des valeurs STI obtenues par la méthode RASTI (voir 4.5.7); cependant, ils sont, en principe, également applicables aux autres méthodes STI.

D.2 Le principe

Les valeurs STI obtenues par la méthode RASTI sont calculées à partir des neuf fonctions de transfert de modulation (m), une pour chaque fréquence de modulation (f_m) indiquées dans le tableau réparties sur les deux bandes d'octaves.

Fréquence de modulation f_m Hz	Bande d'octaves Hz	
	500	2 000
0,7		m_1
1,0	m_2	
1,4		m_3
2,0	m_4	
2,8		m_5
4,0	m_6	
5,6		m_7
8,0	m_8	
11,2		m_9

Chaque fonction de transfert de modulation est le produit de deux termes:

$$m = \left(\frac{1}{\sqrt{1 + \left(\frac{2\pi f T}{13,8} \right)^2}} \right) \left(\frac{1}{1 + 10^{-\frac{S}{10}}} \right)$$

où

f est la fréquence de modulation en hertz;

T est le temps de réverbération en secondes;

S est le rapport signal sur bruit en décibels.

La plage de variation des deux termes est comprise entre 0 et 1.

Pour chaque bande d'octaves, le premier terme donné est une fonction de f , alors que le second est une constante pour un rapport signal sur bruit donné.

Annex D (informative)

Estimating the correction in the case of strong reflections

D.1 General

The procedures shown in this annex indicate the means of estimating the correction to STI values obtained by the RASTI method (see 4.5.7); however, in principle, they are also applicable to the other STI methods.

D.2 The principle

STI values obtained by the RASTI method are calculated from a total of nine modulation transfer functions (m), one for each of the modulation frequencies (f_m) shown in the table, spread across two octave bands.

Modulation frequency f_m Hz	Octave band Hz	
	500	2 000
0,7		m_1
1,0	m_2	
1,4		m_3
2,0	m_4	
2,8		m_5
4,0	m_6	
5,6		m_7
8,0	m_8	
11,2		m_9

Each modulation transfer function is the product of two terms:

$$m = \left(\frac{1}{\sqrt{1 + \left(\frac{2\pi f T}{13,8} \right)^2}} \right) \left(\frac{1}{1 + 10^{-\frac{S}{10}}} \right)$$

where:

f is the modulation frequency in hertz;

T is the reverberation time in seconds;

S is the signal-to-noise ratio in decibels.

The range of both terms is from 0 to 1.

For each octave band, the first term given is a function of f , while the second is a constant for the given signal-to-noise ratio.

Pour un emplacement sans anomalie acoustique notable ou évidente, le relevé de m par rapport à f est lissé et correspond à la réponse d'un filtre passe-bas.

Cependant si, par exemple, la réponse est distordue par une réflexion tardive, cela provoquera de la distorsion de m par rapport à la courbe f et une erreur potentielle dans la valeur STI déduite. Par exemple, le relevé pour un espace ayant une réflexion retardée d'environ 125 ms indique typiquement un plateau ou maximum secondaire à des fréquences de modulation aux environs de 8 Hz.

La correction est faite en ramenant visuellement la courbe par lissage à une forme approximativement normale et en utilisant alors les valeurs estimées de l'indice de modulation pour calculer de nouveau la valeur estimée STI.

NOTE – Cette procédure n'est pas adaptée pour corriger des courbes ayant d'importantes anomalies.

D.3 Calcul

Les étapes du calcul sont les suivantes:

- 1) Convertir chaque valeur m en un rapport signal sur bruit réel:

$$S_{\text{eff}} = 10 \lg \frac{m}{1-m}$$

- 2) Calculer la valeur moyenne du rapport signal sur bruit réel:

$$\bar{S}_{\text{eff}} = \frac{\sum S_{\text{eff}}}{n}$$

- 3) Calculer la valeur STI:

$$STI = \frac{\bar{S}_{\text{eff}} + 15}{30}$$

A titre d'exemple et en prenant le cas de l'espace ayant la réflexion tardive de 125 ms, les valeurs non corrigées et corrigées de m sont:

Modulation de fréquence Hz	Valeur non corrigée de m	Valeur corrigée de m
1	0,675	0,675
2	0,360	0,360
4	0,340	0,150
8	0,365	0,040
0,7	0,825	0,825
1,4	0,520	0,520
2,8	0,100	0,130
5,6	0,150	0,050
11,2	0,130	0,030
Valeur STI par la méthode RASTI	0,42	0,32

For a space with no noticeable or obvious acoustic anomalies, the graph of m against f is smooth and similar to the response of a low pass filter.

However if, for instance, the response is distorted by a late reflection, this will cause distortion in the m against f curve and a potential error in the derived STI value. For example, plots for a space with a reflection delayed by some 125 ms typically show a plateau or secondary maximum at modulation frequencies around 8 Hz.

The correction is made by smoothing the curve visually to approximately the normal shape and then using the estimated values of the modulation index to recalculate the estimated STI value.

NOTE – This procedure is not suitable for correcting curves with gross anomalies.

D.3 Calculation

The stages in the calculation are as follows:

- 1) Convert each m value to an effective signal-to-noise ratio:

$$S_{\text{eff}} = 10 \lg \frac{m}{1-m}$$

- 2) Calculate the mean value of the effective signal-to-noise ratio:

$$\bar{S}_{\text{eff}} = \frac{\sum S_{\text{eff}}}{n}$$

- 3) Calculate the STI value:

$$STI = \frac{\bar{S}_{\text{eff}} + 15}{30}$$

As an example and taking the case of the space with the late reflection at about 125 ms, the uncorrected and the corrected values of m are:

Modulation frequency Hz	Uncorrected value of m	Corrected value of m
1	0,675	0,675
2	0,360	0,360
4	0,340	0,150
8	0,365	0,040
0,7	0,825	0,825
1,4	0,520	0,520
2,8	0,100	0,130
5,6	0,150	0,050
11,2	0,130	0,030
STI value by RASTI	0,42	0,32

Annexe E (informative)

Bibliographie

- [1] Steeneken, H.J.M. and Houtgast, T., "Some applications of the Speech Transmission Index (STI) in auditoria", *Acustica* 51, 1982, p 229-234.
 - [2] Barnett, P.W. and Knight, R.D., "Some practical limitations of STI method", *Proc. Inst. Acoust.* 14.5, 1992, St. Albans, England.
 - [3] Steeneken, H.J.M. and Houtgast, T., "A physical method for measuring speech transmission quality", *J. Acoust. Soc. Amer.* 67, 1980, 31, p 318-326.
 - [4] Rife, D., "Modulation Transfer Function Measurements with Maximum-Length Sequences", *J. Audio Eng. Soc.*, Oct. 1992, Vol. 40, No. 10.
 - [5] ISO 9921-1:1996, *Evaluation ergonomique de la communication parlée - Partie 1: Niveau d'interférence sur la parole et les distances de communication pour des personnes ayant une capacité d'audition normale en communication directe (méthode SIL)*.
 - [6] American National Standard ANSI S 3.5, 1969, American National Standards Institute, New York, USA.
 - [7] Houtgast, T. and Steeneken, H.J.M., "The modulation transfer function in room acoustics as a predictor of speech intelligibility", *Acustica* 28, 1973, p 66-73.
 - [8] Houtgast, T. and Steeneken, H.J.M., "A Multi-lingual evaluation of the Rasti-method for estimating speech intelligibility in auditoria", *Acustica* 54, 1984, p 185-199.
 - [9] Steeneken, H.J.M. and Houtgast, T., "The temporal envelope spectrum of speech and its significance in room acoustics", *Proc. 11th International Congress on Acoustics*, Paris, 1983, Vol. 7, p 85-88.
 - [10] Zwicker, E. and Feltkeller, *Das Ohr als Nachrichtenempfänger*, Hirzel Verlag, Stuttgart, 1967, p 187-200.
 - [11] Steeneken, H.J.M. and Houtgast, T., "On the mutual dependancy of octave band contributions to speech intelligibility", *Proc. Eurospeech 91*, Genoa, 1991, p 1133-1136.
-

Annex E (informative)

Bibliography

- [1] Steeneken, H.J.M. and Houtgast, T., "Some applications of the Speech Transmission Index (STI) in auditoria", *Acustica* 51, 1982, p 229-234.
 - [2] Barnett, P.W. and Knight, R.D., "Some practical limitations of STI method", *Proc. Inst. Acoust.* 14.5, 1992, St. Albans, England.
 - [3] Steeneken, H.J.M. and Houtgast, T., "A physical method for measuring speech transmission quality", *J. Acoust. Soc. Amer.* 67, 1980, 31, p 318-326.
 - [4] Rife, D., "Modulation Transfer Function Measurements with Maximum-Length Sequences", *J. Audio Eng. Soc.*, Oct. 1992, Vol. 40, No. 10.
 - [5] ISO 9921-1:1996, *Ergonomic assessment of speech communication – Part 1: Speech interference level and communication distances for persons with normal hearing capacity in direct communication (SIL method)*.
 - [6] American National Standard ANSI S 3.5, 1969, American National Standards Institute, New York, USA.
 - [7] Houtgast, T. and Steeneken, H.J.M., "The modulation transfer function in room acoustics as a predictor of speech intelligibility", *Acustica* 28, 1973, p 66-73.
 - [8] Houtgast, T. and Steeneken, H.J.M., "A Multi-lingual evaluation of the Rasti-method for estimating speech intelligibility in auditoria", *Acustica* 54, 1984, p 185-199.
 - [9] Steeneken, H.J.M. and Houtgast, T., "The temporal envelope spectrum of speech and its significance in room acoustics", *Proc. 11th International Congress on Acoustics*, Paris, 1983, Vol. 7, p 85-88.
 - [10] Zwicker, E. and Feltkeller, *Das Ohr als Nachrichtenempfänger*, Hirzel Verlag, Stuttgart, 1967, p 187-200.
 - [11] Steeneken, H.J.M. and Houtgast, T., "On the mutual dependency of octave band contributions to speech intelligibility", *Proc. Eurospeech 91*, Genoa, 1991, p 1133-1136.
-

IECNORM.COM: Click to view the full PDF of IEC 60268-16:1998

Withdr2wn